Understanding the Challenges of Using Personal Data in Media Experiences

By Neelima Sailaja

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

August 2019

Dedicated to my Appuppa…..

Once my First Teacher,
Some times my Critique,
Many times my Philosopher,
Most times my Grandfather,
And Always my Fountain of Support,
Motivation and
Unconditional Love.

Abstract

This thesis explores the challenges associated with the turn to personal data in novel media experiences. Emergent media experiences, is turning towards using personal data as a resource to help enhance the possibilities for innovation in media service provision. But while the capabilities presented by personal data are manifold here, historically this shift has seen to introduce many socio-technical challenges that confront the use of these experiences. It is the study and understanding of these challenges, as manifest within the scope of media experiences leveraging personal data that this research turns to.

The research studies this problem from the perspective of two stakeholders involved in this scenario, the users and the service providers. Here, an overtly multidisciplinary approach is adopted, starting from the literature review which engages with previous work from the disciplines of media research, technology, digital economy, law and ethics. To do this, a range of methods which support qualitative research like informal interviews, focus groups, scenario based design, design fiction, thematic analysis, grounded theory and endogenous topic analysis are employed within three studies reported here.

The two formative studies reported seek to elicit user and service provider viewpoints on the challenges of using personal data in media experiences. This is followed by the co-design of a media experience that leverages personal data while including a ‘data dialogue’ that aims to respond to challenges previously uncovered. This design is presented to users and service providers to evaluate their response on this ‘data dialogue’ and to further probe the challenges of using personal data within the media experience.

The contribution of this work could be categorised into two, conceptual contributions and implication for design.

The conceptual contributions explicate the following challenges, as reasoned by both users and service providers. They present the practically grounded subtleties embodied by these challenges when considered within the context of media experiences leveraging user personal data. This is done by comparing the findings of the studies reported here to build upon and contribute to previous conceptualisations of these challenges within literature from multiple disciplines. These conceptual contributions are :

- Value
- Trust
- Privacy

- Transparency
- Control
- Accountability

The implications for design build upon these conceptual contributions to present some practically reasoned sensitivities to be taken into account when considering the design of media experiences that leverage personal data. These recommendations combine the viewpoints of the users and service providers to present design considerations that are sensitive to the challenges raised by both parties, to work towards responding to these challenges. These sensitivities are focused around the following challenges :

- Trust
- Privacy
- Transparency
- Control
- Accountability

The conceptual engagement with challenges here highlight the importance of enabling the users with a more central role in this scenario while the implications for design provide sensitivities that help realise this shift, to work towards alleviating the challenges of both the users and the service providers.

Keywords

Personal Data; Media Experiences; Human Computer Interaction; Value; Trust; Privacy; Transparency; Control; Accountability;

Acknowledgements

I am very grateful and thankful to everyone who has supported me through this beautiful journey of learning and discovery. Here, I name a few who have been an integral part of this process. I would like to extend my sincere gratitude to the many more individuals whose names I could not include in here.

Beginning with my PhD supervisors who have generously supported me through every phase of this PhD. I am very grateful for the mentoring they continue to give me, for the confidence they show in me, for the opportunities they gift me with and the generosity they show towards me in terms of their time and guidance. Thank You Andy, Phil and Mac for choosing me to do this research, helping me through every step of it and making this possible for me. I owe the successful completion of this work to your kindness.

Thank You to Horizon CDT, particularly Prof. Steve Benford, Dr. Sarah Martindale, Emma Juggins, Andrea Haworth and Felicia Black for giving me all the support I have wanted throughout these four years and making this PhD an pleasantly unforgettable experience.

Thank You to the BBC for the generous support, access, resources and opportunities offered to me. I would like to particularly thank Phil (again) for the continuous BBC support he provides me, always providing me with more than what is expected of an industry partner. Thank You to Ian Forrester for his on-going support and mentorship with this work. Thank You to Lianne Kerlin, Rhianne Jones, Adrian Woolard and everyone else from the BBC who has been a constant source of support and encouragement through this endeavour.

Thank You to James Colley, Lachlan Urquhart, Stuart Reeves, Martin Flintham and everyone else from the Mixed Reality Laboratory, who have been generous towards me with their guidance and support.

Thank You to all the other partners of the Living Room of the Future : Adrian Gradinar and Paul Coulton from the University of Lancaster, Roger and Tom from FACT, British Council U.K. and the University of York for the great collaboration and for making this incredible project a reality.

Thank You to all my study participants who kindly kept aside their valuable time and invested effort into this work.

Thank You to my parents for always believing in me and presenting me with the best opportunities a child could have asked for. Thank You to my grandparents for blessing me abundantly with their unconditional love, prayers and wisdom.

Thank You to all my friends from the lab, the CDT, Ancaster Hall and around the world for your support, encouragement and love.

Thank You my love, for your patience.

Thank You Devi,

Neelima Sailaja.

Table of Contents

List of Figures and Tables

Figures

Tables

List of Publications

Publications from the Content Reported within this Thesis

N, Sailaja, J, Colley, A, Crabtree, A, Gradinar, P, Coulton, I, Forrester, L, Kerlin and P, Stenton. "The Living Room of the Future". In Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video, ACM, 2019.

N, Sailaja, A, Crabtree, D, McAuley and P, Stenton. "Explicating the challenges of providing novel media experiences driven by user personal data." In Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (pp. 101-113). ACM, 2018.

N, Sailaja. "Building Trust in Personal Data Driven Media Experiences." 2018. Adjunct Publication of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video. Retrieved March 21, 2019 from **https://figshare.com/articles/DC_Building_Trust_in_Personal_Data_Driven_Media_Experiences/6527939/1.** (ACM Doctoral Consortium Award).

Coulton, P., Lindley, J. G., Gradinar, A. I., Colley, J., Sailaja, N., Crabtree, A., Forrester, I. & Kerlin, L. (2019). Experiencing the Future Mundane.

N, Sailaja, A, Crabtree and P, Stenton. "Challenges of using Personal Data to drive Personalised Electronic Programme Guides." Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, 2017.

Other Research Conducted in line with the PhD

N, Sailaja and R, Jones. "Industry Ideals and Barriers in Using Alternative Privacy Policies." In Proceedings of the 31st British Computer Society Human Computer Interaction Conference. BCS Learning & Development Ltd., 2017.

L, Urquhart, N, Sailaja and D, McAuley. "Realising the Right to Data Portability for the Internet of Things" Personal and Ubiquitous Computing 22.2 (2018): 317-332.

R, Jones, N, Sailaja and L, Kerlin. "Probing the Design Space of Usable Privacy Policies: A Qualitative Exploration of a Reimagined Privacy Policy and Data Exchange Model." In Proceedings of the 31st British Computer Society Human Computer Interaction Conference. BCS Learning & Development Ltd., 2017.

Definitions

Personal Data : "'personal data' (within the context of this thesis) means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;" (GDPR, 2018)

Media Experiences : Within the context of this study, media experience has been extended to be applied to the entire spectrum of both traditional and novel media experiences. While traditional media experiences include news, sports and other media content consumed on radio and television, more modern interpretations of media experiences are inclusive of digital platforms that include social media, video-on demand services, recommendations and advertisements on these platforms and other miscellaneous novel genres of media experiences that include AR (Damiani, 2017), VR (*David Bowie's "Heroes" Inspires Quest into New Technologies at Sundance 2017*, 2017), Interactive Media like physiologically driven stories (Pike *et al.*, 2016) and adaptive media (Gradinar *et al.*, 2015).

1. Introduction

This chapter introduces the context of the research undertaken here, beginning with the Research Area this thesis explores, followed by a statement of the Research Problem, the Aims and Objectives of the research, a brief overview of the chapters and a short account on the contribution of the research conducted.

1.1. Research Area

Modern society's most mundane everyday activities are now often deeply embedded within data rich environments, increasingly powered by user personal data. Most technological products and services of today are personalised to be sensitive to each individual's preferences, context and lifestyle. To uphold these expectations of customisation of the average user, there is a growing demand for user personal data, the fuel which drives this shift and the asset which is exchanged in return for innovation in services (Schwab et al., 2011).

One domain, that is recently demonstrating significant interest in leveraging user personal data owing to its' potential for experience enhancement in ways never witnessed before is **media** or **media experiences**. From tailoring recommendations on YouTube and Spotify to delivering curated News feeds and even adapting the content of a piece to reflect the audiences' context, media is both exploring and showcasing multi- dimensional ways of **exploiting user personal data to power innovation**.

While theoretically this formula for exchange of personal data for enhanced services seems like a straightforward one, when applied in the lived, real world, this glossed over abstraction starts stripping down to reveal a number of underlying social, technological, legal and ethical challenges that affect this exchange.

Disregarding this underlying flaw in the system, the collection and use of personal data and its extrapolation into inferences tend to continue at a rampant rate across disciplines as diverse as health to sports to energy services. Here, international communities like the World Economic Forum have looked into this trend, highlighting the opportunities and challenges presented by the use of personal data (Kalapesi, 2012; Rose, Rehse and Röber, 2012). While such reports help surface and bolster the underlying adverse impacts of this scenario, their efforts often stop at laying out statistics to highlight the significance of these implications and proposing very high level responses that demand further contextualised study to be practically adaptable in real-life contexts.

1.2. Research Problem Statement

As media turns towards user personal data as a significant contributor to its future enhancement, the shift simultaneously introduces challenges into this domain. Recent events like the Cambridge Analytica scandal (Cadwalladr and Graham-Harrison, 2018) and privacy incidents associated with Netflix's data use (Singel, 2009, 2010) are indicators of these challenges adversely affecting audience trust and confidence in this shift.

As previously stated, studies and reports have looked at these challenges but often from higher levels of abstraction and generalisation. However, to alleviate these challenges effectively, there is the need to study them closely, exploring and unpacking the subtleties associated with this shift within the particular context of media experiences.

It is towards an interdisciplinary study, exploration and response to this specific problem that this thesis turns. The overarching research question here focuses on **understanding and unpacking the challenges presented by the use of personal data in novel media experiences**. This effort is expected to help highlight the challenges adversely affecting this shift within the domain of media. It will present the nuances of these concerns framed within the context of media experiences and grounded within the practicalities of creating and consuming media experiences. This is expected to develop sensitivities that could inform the creation of future data driven media that would effectively respond to these challenges starting right from a very practical level, thereby sustaining both innovation and personal data leverage in future media experiences.

1.3. Aims and Objectives

The core aim of this research is to address the presented problem space which is : **to understand the challenges of using personal data in media experiences**.

This aim is achieved through three objectives, as follows :

> Elicit, identify and understand the challenges involved in the use of personal data in novel media experiences from the **user's viewpoint**

The audiences, being both the producers of the data and the consumer of the services, form a key stakeholder in this scenario. As part of the formative phase of the research, it is essential to understand the users' attitudes towards personal data being collected and used in association with the media services they consume and highlight challenges in this respect, if any.

Elicit, identify, study and understand the challenges involved in the use of personal data in novel media experiences from the **service providers' viewpoint**

Service providers who hold the power and discretion when it comes to handling of personal data (ranging from collection, processing, storage, sharing and design and dissemination of media experiences using personal data) form another major stakeholder in this domain. Handed with the capability to make changes in the data leverage mechanisms, understanding the challenges faced by service providers in this respect, as they participate in the turn towards personal data is essential to have a holistic understanding of this emergent shift.

To **design** responses that address the raised challenges to elicit **user** and **service provider feedback** and **response** to a media experience collecting user personal data while being sensitive to this process.

Once both the audience and service provider challenges are known, these learnings are bridged to design a media experience that leverages user personal data and includes responses to these challenges. This experience could be used to elicit feedback on the responses deployed and the use of personal data within the media experience once the users engage with them in an immersive, active manner. The same experience could be used as a probe to engage in discussions with the service providers to understand their perspectives on such a media experience leveraging personal data, the responses to data collection challenges used within it and the viability of inclusion of such responses in future data driven media experiences.

1.4. Chapters Overview

The thesis reports an overtly interdisciplinary research approach that inspired three studies. The reasonings around the various choices and decisions that led to the scoping and design of the studies are preambled by a detailed review of previous literature from various disciplines, followed by an explanation of the research methods and methodological configurations chosen along with the justification for the same. This is followed by the reporting of the studies themselves, followed by an extensive discussion that draws together the results of the studies and compares them with previous literature to highlight and present the key contribution of this research. An overview of these chapters are as follows :

Chapter 2 : Literature Review

The reporting of this interdisciplinary research endeavour begins with a detailed multi-disciplinary literature review. This chapter commences with an overview of the turn towards personal data in media experiences. Adopting a chronological approach here, the shift

is discussed starting from the earlier days of personal data use in advertisements to then move onto social media to recommendations and finally to the latest media forms where the content itself is customised using personal data. The review then presents the challenges of personal data use, as reported in literature from media research. Owing to the lack of significant critique in this scope, the review then turns to broader multi-disciplinary literature to engage with the different challenges and responses identified in the use of personal data by various domains. The generic nature of these interdisciplinary approaches to personal data challenges and the lack of critique contextualising these challenges within media experiences identifies a research gap, which is explored within the rest of the thesis.

Chapter 3 : Methodology

Once previous literature is reviewed to highlight the research gap to be explored, the thesis then turns to reporting the methodology adopted for the research. The research employs a number of methods, configured in ways that best fit the requirements of each study. The chapter begins by presenting the methods employed in the various studies, grouped as data capture methods and data analysis methods. This is followed by explanations about the methodological configurations of each study wherein the reasonings around the selection and use of each method and methodological configuration is provided.

Chapter 4 : Study I

This chapter reports the first study done as part of this PhD, which probed user viewpoints on the potential use of personal data in near-future media experiences. It employs scenario based exploration enabled by the use of mockups to convey the functionalities of data-driven personalised EPGs to users and to facilitate discussions around its potential use. The results of the study show that users preferred personalised EPGs over current popular EPGs but expressed a significant lack of trust in the personal data collection that drives personalisation. The results further unpack this finding to reveal user appreciation of the various functionalities afforded by personalisation of media while simultaneously presenting their apprehension regarding the implications of their personal data being collected, particularly focusing on concerns about their privacy, data collection in the context of their homes and the current lack of legibility and control.

Chapter 5 : Study II

This chapter details the second study done as part of this PhD which elicited service provider viewpoints on the potential use of personal data in near-future media experiences through an interview study of 20 media service providers within the BBC. The results of the study

identify a need for better interactions in the user-data-service ecosystem where trust and value are to be prioritised and balanced at all times. The major considerations and responses that help maintain this balance further unpack as the need for being legally compliant, going beyond just the mandatory by ensuring social accountability and exercising ethical responsibility as an organisation. The study also presents how technology itself is seen and used as a solution for overcoming some of these challenges to provide value while preserving trust within the personal data ecosystem.

Chapter 6 : Study III

This chapter presents a bridging study that utilises the learnings from the previous formative studies to co-design a media experience that leverages personal data from the audiences. This design fiction experience is then presented to the audiences and their feedback, particularly, their response to the experience itself, it's use of personal data and the 'data dialogue' interventions included is collected through post experience interviews. This experience is also used to engage in conversations with service providers to understand their perspectives on a media experience leveraging personal data, the responses to data collection challenges included within it and the viability of inclusion of such responses in future data driven media experiences.

The results of the deployment show that audiences found the LRoTF highly immersive, engaging and 'magical'. However, audience members were clear about the demarcation between enjoying a short-lived new media experience versus adopting it in their everyday lives. Here, they articulate a spectrum of adoption challenges that confront the uptake of new data-driven, immersive media experiences in everyday life, which should be taken into account for the sustainable adoption of media experiences that leverage user personal data. The service providers' responses focused on the importance of data legibility, control and the balance of provision of value to the users versus loss of trust. They contextualized their reasonings around these topics within the experience designed and their expectation of including similar responses in the data driven media experiences of the future.

Chapter 7 : Discussion

The discussion chapter brings together the findings of the three reported studies to present the challenges of value, trust, privacy, transparency, control and accountability, as reasoned by the users and service providers, within the context of media experiences leveraging user personal data. These challenges are explained in detail by comparing and contrasting them with the findings of the previously discussed literature to present the unique conceptual and design contributions this thesis makes.

1.5. Contribution Statement

The contribution of this work could be categorised as conceptual contributions and implications for design.

The conceptual contributions unpack the following challenges as elicited by both users and service providers to present the practically grounded complexities characterised by these challenges when considered within the context of media experiences leveraging user personal data. While findings from diverse disciplines may have considered these challenges in a non-contextualised, often abstracted and generic manners (detailed in Section 2.3) previously, the conceptual contributions of this thesis grounds these challenges within the scope of media experiences leveraging user personal data by looking into both user and service provider practicalities around these challenges and the conceptual manifestations of these practicalities that unpack as the contextualised subtleties of these challenges. Thus, the findings of the studies reported here are used to build upon, contextualise and contribute to the previous, more generalised and siloed conceptualisations of these challenges found in literature from multiple disciplines, discussed in Chapter 2. These conceptual contributions are :

Value

This research presents the challenge of value by grounding it within the scope of media experiences that use personal data through going beyond previously considered views of limiting value to financial returns to extending it to highlight and explicate the various complexities associated with this challenge.

Trust

The research extends the previously established ‘crisis in trust’ (Section 2.3.1) to describe its properties in legal, social and ethical terms, reflecting its’ practical manifestation within media experiences that leverage user personal data.

Privacy

This thesis extends previous discussions around privacy when considering personal data to explicate the nuances of ‘intangible risks’, concern around access and potential dystopias, characterised by privacy as a challenge when personal data is leveraged by media experiences.

Transparency

This work contributes to dialogue around transparency as a challenge by moving from generic abstractions to contextualised resolution of the details of its properties like consideration of data legibility and

communication of the ‘effect’ of the data on the experience, as present in media experiences that leverage user personal data.

Control

The challenge of control is again treated beyond generalised conceptualisations to present the various practically adaptable aspects of control within a media experience that leverages user personal data that include control over the media, control over the data, sense of autonomy through control of IoT devices and control of social scenarios.

Accountability

The research also presents the different aspects of the challenge of accountability manifest within media experiences that leverage user personal data by reflecting them back to the various siloed perspectives on accountability, as presented by literature from different disciplines.

The implications for design reflect upon the conceptual contributions by presenting some practically reasoned sensitivities to be considered by academics, researchers and practitioners during the future design of media experiences that leverage user personal data, so as to help alleviate the discussed challenges. These recommendations combine the viewpoints of the users and service providers to present contextualised design considerations that are sensitive to the challenges raised by both parties. These sensitivities are focused around the following :

Trust

The research moves from high level theoretical discussions around trust to present legal, social and ethical sensitivities that would help in responding to the challenge of loss of user trust by actively involving considerations that speak to this challenge.

Privacy

Sensitivities around privacy considerations that are contextualised within media experiences and which when prioritised in the future making of data driven media experiences would help in alleviation of the said challenge in this context is highlighted.

Transparency

Solid recommendations that focus on the various subtleties and opportunities around transparency within media experiences that leverage user personal data, which extends from just terms and conditions statements to the design of an effective ‘data dialogue’, the consideration of which would help alleviate the challenge around transparency are brought forth.

Control

The research unpacks the practically inspired nuances around control, which include not just control over the data but aspects of the media, devices and social settings, which help explain the different dimensions of this challenge, along with the paradoxical requirement for granular but minimal, quick and easy active user intervention, which would help alleviate the challenge around control.

Accountability

Accountability and its manifestations in different forms, in line with different multi-disciplinary perspectives on it is highlighted to suggest the careful consideration and response to these different perspectives when considering response to this challenge.

Here, the conceptual explication of challenges highlight the importance of enabling the users with a more central role in this scenario to work towards alleviating the challenges of both the users and the service providers identified here. The implications for design lay out sensitivities that help realise this shift by proposing measures that both support current practices while also recommending a shift from cloud based models to alternate methods of data processing that allow for increased user participation.

2. Literature Review

This chapter presents a review of literature and previous work around the challenges of using personal data in media experiences. The review begins with a presentation of the various aspects of media that have been increasingly leveraging personal data for service enhancement. Once the turn towards personal data in media is grounded through an elaboration of these novel media examples, the challenges of this shift, as outlined in media literature is reviewed. Owing to the lack of significant critique in this specific domain, given the emergent nature of this shift, focus then turns to literature from other multiple disciplines that have previously engaged with the challenges of personal data use, which while being broader in scope, is found to be generic and lacking contextual sensitivity.

Thus, the chapter is concluded with the identification of a research gap stemming from the drawbacks of these two sets of literatures considered. Here, the more general multidisciplinary discussions warrant more contextualized study focused on media while literature from media research calls for a more engaged critique on these challenges of using personal data in media experiences if the alleviation of these challenges is to be successful.

Thus, the review of these multi-disciplinary literatures unveils a **research gap** for the explication of the challenges of using personal data in media experiences.

2.1. The Turn to Personal Data in Novel Media

Personal data has been a contributor to several aspects, forms and derivations of media over the years. The following sections talk about this turn starting from advertisements that use it for optimising user response to the increasing number of emergent media genres that leverage personal data for enhancing the very content of the experience. While this section discusses the various expressions found by personal data in and around a media experience to demonstrate the impact of personal data in this domain, the focus of this thesis is on the use of personal data for experience enhancement, where the term experience ranges from the formative phase of discovery of media content of interest to the traditional sense of consumption of the very content of the experience.

2.1.1. Advertising

Advertising was one of the very first aspects of media that started leveraging user personal data through personalisation. Here, personal data was not used *in* the media content but rather *around* it, to improve its consumption and optimize audience response to it.

Customised adverts that reflect the context of user browsing, exploration and activity is now an everyday mundane occurrence. These adverts are not just on websites and computers but with the proliferation of technology, they have spread to mobile phones, tablets, smart watches and even Digital Televisions (Bozios et al., 2001).

Customised advertisements provide the advantage of serving material that is relevant to the users, thus eliminating the need for longer searches (Chellappa and Sin, 2005; Tam and Ho, 2006; Okazaki, Li and Hirose, 2009). Research has shown that providing advertisements that resonate more with the user's preferences could lead to increased purchase intentions (Franke, Keinz and Steger, 2009; Goldfarb and Tucker, 2011) whereas the opposite could lead to irritation (Thota and Biswas, 2009).

2.1.2. Online Media Platforms : Social Media and Video-On-Demand

Personal data also moved onto social media platforms, where again it found expression in a different aspect of media, helping customise feed and other particulars to suit user interests. Here, media is democratised, less about the platform, more about the content, mostly short-form and very much 'now'. Spaces like Facebook live, Twitch, and Vimeo have opened up the arena even for novel concepts like 'eating' as a genre of monetisable content (Hong, 2016).

On these platforms, personal data becomes almost a necessity for promotion of the most novel and enhanced experiences. The content on the social networks (eg., video, audio, livestreams, posts etc.) themselves form a collection of easily identifiable personal data. Here, personalisation often manifests as customisation of the users' feed, for e.g., content on regional festivals and events being promoted within the relevant area. Targeted advertising has also found its way into these popular platforms, leveraging the very personal data that these platforms furnish to infer the perfect context and timing to deliver themselves to the users on these platforms.

2.1.3. Recommenders on Online Media Platforms

The turn to personal data started influencing the media experience as online media platforms like Netflix and Youtube started personalising recommendations through personal data leverage. These recommender systems are used to provide personalised media suggestions from the wide range of potential content available in order to provide audiences with targeted content that is of interest to them. Recommender systems, when used within media experiences form an essential part of modern media consumption as it not only curates content but also shapes entire user journeys on digital media platforms by presenting and proposing content to the users. This makes the formative phases of the media 'experience', which involves the discovery of content, easier.

There are different types of recommender systems : collaborative filtering, content-based, demographic, knowledge based, context aware, ensemble, hybrid and social. These systems use a wide range of data including behavioural and contextual information to user preferences and attributes to process recommendations that reflect the audiences' interests and lifestyles. This inherently data driven nature of recommender systems have led to concerns over user privacy particularly around data collection, retention, sales, employee browsing private information, recommendations revealing information, shared device or service and stranger viewing private information (Jeckmans et al., 2013). These concerns around privacy in recommender systems have also spurred a range of responses which are targeted at dealing with the privacy concerns arising particularly within recommender systems.

One of these responses is anonymisation. This involves removing any identifiable information from the data while preserving other structures of interest within it. But, anonymisation has limitations, as only parts of data are often obfuscated while other parts are left intact for the dataset to yield sufficient results and inferences. This limitation was reflected in the re-identification of anonymised data published by Netflix as part of their recommender systems prize (Narayanan and Shmatikov, 2006; Singel, 2009, 2010).

Another option used here is the introduction of trusted agents (Cissée and Albayrak, 2007) to act as a relay and filter the data that is sent. Thus, the user interacts with the recommender system in an anonymous fashion, with the agent acting as a buffer with the expectation that the service provider is not able to link their information to them. This situation while seemingly more effective than just basic anonymisation techniques, calls for the user to place trust on the agents involved.

Similar to anonymization is randomization or perturbation where a degree of uncertainty is injected into the system by altering the data. There are various methods of perturbation, ranging from singular value decomposition predictor based on random perturbation of data (Polat and Du, 2005) to combinations of random perturbations which uses a peer-to-peer structure to create a form of dynamic random perturbation (Berkvosky et al., 2007). There are also methods of aggregation which differ from perturbation where aggregation happens between the users with minimal involvement from the recommender. Thus, the system will not be able to identify which information is part of the original user information and what is added by aggregation, thus adding a degree of uncertainty.

Differential privacy (Dwork, 2008; McSherry and Mironov, 2009) is another method for privacy preservation employed within recommender systems. It works towards obscuring the traceability between the single users' information in the input (the user's information) and output (the recommendation) by adding noise into

the inputs or outputs. The level of the noise is dependant upon how and how often the data is used and is balanced between accuracy of the output and input privacy.

Privacy preserving cryptographic protocols are also leveraged within this scope (Goldreich, 2006). Among them are secure multi-party computations, secret sharing, homomorphic encryption, and zero-knowledge proofs. Secure multi-party computations are a class of protocols that allow two or more parties to collaboratively compute a function based on input held by each of them. Homomorphic encryption allows one (or sometimes more) operation (for exam- ple addition or multiplication) on the encrypted values, by performing a corresponding operation on the ciphertexts. Zero-knowledge proofs allow a user to prove a property about a value, without revealing that value. For example, that a value is in a given range of possible values.

2.1.4. User Personal Data in New and Emergent Media

After advertising, social media and online recommendations personal data has now permeated its way to influencing the making of the content of media experiences through various emergent forms of interactive and adaptive media that uses personal data for experience enhancement in diverse ways.

In the interest of presenting the significance of the turn towards user data in new media, the next few sections go into the details of this more recent interest in leveraging user data to customise media content. They look into different genres of emergent media, most of which fall into the category of interactive media. This discussion is closed with a brief look into a relatively novel genre in this space, made possible by the availability of user personal data, that of perceptive media.

2.1.4.1. Interactive Experiences

The utilisation of personal data in new media has currently moved beyond even advertisements, social media and recommendations to influencing the very experiences itself. When content starts responding and reacting to user personal data, traditional forms of one-for-all broadcast media is redefined to deliver or ‘narrowcast’ user-specific, customised, unique experiences to audiences.

The area in media that is seeing much interest and progress in this respect is interactive media wherein the media content allows for direct user interaction and customisation through user data (Hanley and Viney, 2001; Ursu, Kegel, et al., 2008; Ursu, Thomas, et al., 2008; Cesar and Chorianopoulos, 2009; Obrist et al., 2015). The engagement of the audience member is so immersive that interactive films have almost shifted the role of the artist from one of possessing complete control over the experience to being the critical entity that crafts the story design with which the audiences ‘co-create’ (Some

examples., the International Emmy Award winner, the last Hijack (Last Hijack, 2015), The World in Ten Blocks (Chollangi, 2017), Room 202 (Room 202, 2016) and Possibilia (Sondhi, 2015)).

Interactivity in media has traditionally been a much-deliberated concept. Steur's definition of interactivity in a media experience was one where "users can participate in modifying the form and content of the mediated environment in real time" (Steuer, 1992). Markus's version of the same stated that an "interactive medium is a vehicle that enables and constrains multidirectional communication flows among the members of a social unit" (Markus, 1987). Rada added the capabilities of computer-mediated information sharing systems including multi-media and hypertexts (Rada, 1995). Today, personal data has emerged as an enabler that helps realise all of these definitions of interactivity to provide users with multifaceted interactive experiences that removes "old constraints" and offers users "new liberties of action" (Cherry, 1977), through interactivity that respects every individual user's requirements.

To showcase the diversity and spread of these 'liberties' allowed by personal data, the next few sections have been utilised to select, describe and explain through examples, the role of data in some of these emergent genres that are experiencing significant innovation through the use of data.

Mixed Reality/ Augmented Reality

Mixed Reality (MR) and Augmented Reality (AR) are technologies that overlay digital content on the physical world, thus exposing audiences to richer visual experiences, embedded in layers of information.

Since its introduction to the mainstream market, AR applications have been employed for a multitude of purposes like marketing (Blippar campaigns(Blippar, 2019)), cultural tours of places and museums (Ioannidis, Balet and Pandermalis, 2014) (Museum of London (Zhang, 2010)), viewing constellations (Google Sky (Google, 2019)), applying makeup (ModiFace(Constine, 2016)), educating children (Happy Atom (Goodner, 2016), Mardles (Crisp, 2017)), and understanding other languages (Google Translate (Gershgorn, 2015)). The popularity of this genre increased exponentially with the arrival of Pokemon Go (Grant, 2016; Si-soo, 2017) in July 2016 and follow ups like the Harry Potter: Wizards Unite (Etherington, 2017).

This trend has also propagated to novel media wherein AR is increasingly finding space in media experiences either through the audiences' smart tablets, phones, or head mounted displays like the Magic Leap, Microsoft Hololens, Meta, or Google Glass.

The Outthink Hidden (Robertson, 2017) application was such an initiative, based on the movie Hidden Figures, it allowed users to

discover and learn about black, female doctors, engineers, and scientists and other marginalized individuals who have shaped history. The reliance of the application on the antiquated QR code model brought it much critique but it helped expose the possibilities of using AR in introducing more dimensions to an otherwise traditional media experience.

Heroes: A Duet in Mixed Reality (David Bowie's "Heroes" Inspires Quest into New Technologies at Sundance 2017, 2017) is a multidimensional experience using 2-D, virtual reality, and augmented reality headsets that premiered at Sundance film festival in 2017. "The plot is to immerse the audience in a fiercely athletic and romantic dance that lets them chase the dancers around a pool and on stage, scale through a labyrinth of stunning architecture, and even reach out and hold the dancers in the palm of their hands." (David Bowie's "Heroes" Inspires Quest into New Technologies at Sundance 2017, 2017)

In the same year at Sundance was premiered another phenomenal Interactive AR experience, Journey to the Center of the Natural Machine (Damiani, 2017). This example moves further from usual HMD driven experiences which tend to be isolated and single user driven by introducing collaborative AR as two users are able to watch and interact with the same hologram of the human brain, thus influencing the experience 'together'.

As AR and MR becomes a promising avenue for future exploration in media experiences, it brings along the need for rich sets of user data fed in through user phones, tablets, cameras and other sensing equipments. Be it the need for extending beyond just QR codes in the Outthink Hidden app or the constant need for geolocation and interaction capture, recalibration and redrawing in the other two, we see the use of data shifting from a luxury to almost a primal need in this class of experiences.

Physiologically Driven Stories

Physiologically Driven Stories refer to emergent media that exploits the capabilities of the human physiology and the massive amounts of data it offers to be collected, processed and reflected back.

These include works like #Scanners (Pike et al., 2016) and the Moment Movie (Ramchurn et al., 2019), which alters the trajectory of the movie itself dependant upon variations in audience physiological data collected. This data is collected through devices capable of reading physiological parameters like fMRIs, GSRs, etc, adhered to audience bodies, thus reading their subtle bodily responses to the media they are consuming, and tailoring the media to suit this data.

Another class of Physiologically Driven Stories would be bio-responsive or bio-connected media that use biometric technologies to create connected or transcendent media experiences for audiences (*UKI*, 2009; *My Sky is Falling*, 2013; Henry, 2013). The latest in this arena being the concept of bioart where biological matter is used as a medium to perform software and hardware. An example of such work would be Gina Czarneck's project Heirloom (Clara Rodriguez Fernandez, 2017; Czarnecki, 2018) that 'grows' living portraits of the artist's daughter, from her own cells, cultured from a single sample taken from her mouth.

As media moves from screens to being ubiquitous to even being able to learn and influence the very bio-matter that makes up the human body, it is not difficult to realise the very heavy dependance on the exchange of data, data that delves into the very 'being' of a 'human being', this shift embodies.

Data Storytelling

A upcoming 'genre' of storytelling is data storytelling which employs data collections to produce 'stories' on and about the human experience. These stories often give data a visual interface, mostly at a high level, forming almost a 'bird's eye view' of human phenomena.

While such data visualisation might have begun with examples like the map of Napoleon's march (Cheng, 2014), they have come a long way with efforts like We Feel Fine (Harris and Kamvar, 2006), I Want You to Want Me (Harris and Kamvar, 2008), Artificial Killing Machine (Reich, 2015), Derive (Quevillon, 2011) and Dear Data (Silverberg, 2016).

With media makers exploring the possibility of marrying data analytics into VR and AR and the possibilities afforded by 'ambient user interfaces' (Bisson, 2017) that exploit mundane everyday Internet of Things connected objects to express data, data storytelling is expected to elevate its' game significantly in the days to come. Needless to say, the progress in data storytelling, which is primarily driven by large amounts of user data will only further strengthen the turn towards data driven innovation in new media experiences.

Geolocative or Geo-Aware Experiences

Geolocative Experiences contextualise the story to the surroundings and manage live interactions by leveraging GPS status on the audience's mobile devices. By tying the story to the immediate context of the real world, the audience is presented with elevated levels of immersion and engagement with the experience.

An example of such an experience is "The Silent History, a groundbreaking novel, written and designed specially for iPad and iPhone, that uses serialization, exploration, and collaboration to tell

the story of a generation of unusual children" (The Silent History, 2012). The story is available only to those whose GPS readings match the co-ordinates of a specific location. Hundreds of location based stories were dispersed around the world to enable access to audiences, irrespective of their geographic location on the globe.

Location aware music like the National Mall project (Holladay, 2013) and others that help with park/ museum experiences (Hazzard, Benford and Burnett, 2015) that activates selected bits of music depending on the location of a person is another form of geo-aware media that is getting much attention. This genre of media, which is primarily dependant on user geographic data will foster innovation only with the promise of sustained flow of user location data, further emphasising the shift towards personal data in yet another form of novel media.

2.1.4.2. From Interactive to Perceptive Media

Perceptive (or adaptive) media could be considered a step ahead of even interactive media. In interactive media, users create or influence a dramatic storyline through direct action (Riedl and Bulitko, 2013). 'Perceptive media' adapts a story to the audience using contextual information relevant to them, gathered through the use of a range of sensors and sensing technologies, and tailors the story within the predefined story arc set by the content creator. The goal of the interactive narrative is to immerse users in a virtual world such that they believe that they are an integral part of an unfolding story. In the case of perceptive media, the overall story arc does not change. It simply varies the more ambient parts of the narrative to create a more emotionally engaging experience for the audience by making the content more contextually relevant, thus approaching immersion from a liminoid perspective (Coulton, 2017).

Here, the data generated from the audience could therefore be considered as diegetic influencers in that it is both generated within, and has influence upon, the experience. The resulting experience is thus similar to how a person telling a story might adapt aspects of it to the particularities of the location and the group. They might make use of specific expressions, sayings or habits, and reference landmarks and places in the local vicinity to ensure the story resonates with the audience.

The most famous example of the power of such engagement is Orson Welles' contemporary retelling of H.G. Wells' War of Worlds, in a radio programme in 1938 which allegedly produced a wide scale panic amongst many of the show's listeners who believed it to be real. What this highlights is that by making content contextually relevant to an audience, the newly created media can produce levels of emotional engagement and immersion to such an extent that it can even override logical reasoning.

The Perceptive Radio (Gradinar et al., 2015) is one such example which is "able to play specially designed content that adapts to the physical and social context in which the radio resides....[...].... Through the use of a revised fable, the radio delivers an adapted narrative based on the location of the play, number of people in the room, time of day and weather data collected through the available sensors."

The work by Coulton (Coulton, 2015) is another example on the same lines that contextualises the story for the audience by augmenting broadcast using a number of IoT objects (lights, doll, olfactory device). These IoT objects were synchronised to particular on-screen events within a BBC documentary thus creating a broadcast experience which utilised household objects within the viewer's own context.

This compilation of genres of new media experiences is not comprehensive when considering the universal set of all emergent media. But, it is utilized here to present a substantive introduction to the major share audience data holds in this economy of innovation in new media powered by user data, particularly user personal data[1]. Having grounded the claim of a shift towards user personal data in new media through the elaboration of these emergent experiences and their dependance on user data, the argument now turns to exploring this shift further, particularly looking at previous literature around media that discusses the challenges and potential responses in this scope.

2.2. The Challenges of using Personal Data in New Media

Having looked at the multifarious ways in which personal data manifests enhancement in media experiences, the dialogue now turns to the discussion of challenges presented by the same in previous literature within media.

2.2.1. Privacy

Media literature has highlighted privacy as a challenge confronting the use of personal data at various occasions. Here, privacy has been talked about starting from research in advertising and social media to research that looked into interactive media experiences that leveraged user personal data, as explained below.

Advertising research explicitly states that the use of personal data, could be seen as intrusive by the consumers which in turn affects the user's cognitive processes, interrupting goal pursuit (Li, Edwards and Lee, 2002). In instances where user privacy is violated (Rapp *et al.*, 2009), wherein if the users feel like their privacy has been compromised by data use, they find it off-putting and creepy (Stone,

2010). Here, White et al. (White *et al.*, 2008) has found that while perceived utility helps reduce these negative responses to customisation, when highly distinctive personal information is used, this effect is actually lowered (Van Doorn and Hoekstra, 2013) sometimes even leading to reactance (Brehm, 1966, 1989; Clee and Wicklund, 1980) where the users start resisting the advertisement's appeal (White *et al.*, 2008).

User privacy has also been highlighted and studied as a case for major concern (Ramzan *et al.*, 2012) within social media. Research around the same starts right from the use of cookies (Pierson and Heyman, 2011), whose very existence users are often hardly aware of and extends to development of taxonomies around privacy concerns to help understand the same (Jamal, Coughlan and Kamal, 2013).

User privacy was surfaced as a challenge in scenarios of personal data use by online media platforms when Netflix was forced to cancel the sequel to their $1 million personal data driven recommendation algorithm improvement contest (Singel, 2010) following a warning from privacy advocates that notified the new dataset to be easily de-anonymised. Also highlighting the privacy concerns in media experiences using personal data was the instance of a suit being filed by a privately homosexual mother who was "alleging that Netflix violated fair-trade laws and a federal privacy law protecting video rental records, when it launched its popular contest in September 2006" (Singel, 2009).

Most research around media recommenders focus on the technical challenges surrounding information privacy associated with recommender systems, often focusing on "the inherent trade-off between accuracy of recommendations and the extent to which users are willing to release information about their preferences" (Xin and Jaakkola, 2014). What is left to be explored here are the range of socio-technical challenges of using data within media experiences, including media recommenders, that go beyond the technological underpinnings of information privacy alone. While there has been brief mentions of trust concerns stemming from exposure, bias and sabotage (Lam, Frankowski and Riedl, 2006) the detailed and contextualised exploration of such challenges of trust, accountability, transparency and control and their manifestation within this scope are avenues of further study in this scheme. Also, to be included on the agenda is the treatment of privacy itself as more than just information privacy, to be sensitive to the socio-technical subtleties associated with this notion which has traditionally been treated in multi-dimensional manners in diverse disciplines. To understand which of those definitions of privacy apply or do not apply here and/or what are the novelties introduced by this case of data driven media experiences, including recommenders when considering privacy, are all open to further study.

A good portion of early research on these lines was found often in association with studies based on interactive TVs. An example is the Social TV which allowed users to be connected during a consumption experience through a camera and microphone. Here, Cesar et al. (Cesar, Chorianopoulos and Jensen, 2008) found that when introducing social TV "privacy issues should also be taken into account." And Tullio et al (Tullio, Harboe and Massey, 2008) found that issues of privacy arose during Social TV use when the users did not know who their conversation partner was. Other issues relating to privacy were considered as de facto non-existent at the time as for example, users could not do anything about their data (like photos) when using the set-top box as a storage medium and the lack of personalisation options reduced the possibility of keeping information private.

Bernhaupt et al's ethnographic studies around Interactive TV (Bernhaupt *et al.*, 2007) focused on understanding user perspectives on security and privacy. This work, following the lines of ethnographic field reporting, details a range of user responses to privacy and security. It concludes that "the majority of the households were not concerned about security and privacy issues" because "Interactive TV [was] still seen as an extended version of normal TV, not posing any security risks". The study concludes by recommending "to allow people to see possible security and privacy risks, new interactive TV offers must carefully explain the connection of the TV to the Internet (in the case of IPTV offers), the possible connection of the (hybrid) set-top box to the Internet, and other networks and the problems in terms of security and privacy for user-generated content (access to the content on other set-top boxes is not secure, and additionally children might get access to content they would not have on the set-top box in the household)."

Trepkeviciute's (Trepkeviciute, 2017) more recent work on IoT driven media recommendations quite contradicts Bernhaupt's findings by stating that "It could be concluded that privacy is the most important element which should be taken into deep consideration" here. The work glosses over privacy as a "predominating issue" which should "ensure high data security standards". They conclude by stating that solutions here are "complicated" due to the involvement of personal data and that "the IoT TV should guarantee data security and privacy."

While privacy is discussed by a number of pieces of research around interactive media experiences that leverage personal data, these findings pose a couple of challenges themselves. Firstly, while the majority of these findings do mark privacy as a challenge in the use of personal data in novel media, Bernhaupt's findings starkly contradict this notion by stating how their audiences did not perceive privacy as an issue in media consumption scenarios due to the data collection being not apparent in these instances. Secondly, the

findings presented here are often either gross abstractions that stop at simply stating privacy, sometimes even limiting the scope to just information privacy, as a concern or makes a call for the challenge of privacy to be responded to without much effort to explicate this challenge enough within the scope of these media experiences, so that they can be responded to in future design. Thirdly, the interspersed nature of these studies and their findings, makes it difficult to effectively compare them to weave together a more coherent narrative around privacy in this scenario. These challenges highlight the need for studying this landscape to both explicitly present privacy as a barrier in the use of personal data in media (if that is the case) and unpacking its nuances to help develop practical sensitivities that would respond to it in future design.

2.2.2. Other Challenges

Along with privacy and security, media literature also mentions a few other 'values' that challenge the turn towards personal data in future media experiences. These include control, agency, transparency and trust.

The need for user control of their personal data (Culnan and Bies, 2003) and its potential role in alleviating user concerns has been presented within the context of advertisements (Tucker, 2014). Social media research has highlighted the potential of increased user agency within data collecting environments as a response to concerns (Kennedy and Moss, 2015). On more practical levels, agency through the perspective of user control and awareness was also considered. Here, "a negative association [was found] between information disclosure and perceived control over personal information, but a positive association was found with user awareness and security notices." (Benson, Saridakis and Tennakoon, 2015). Moving onto studies in interactive media, Trepkeviciute's (Trepkeviciute, 2017) recent work with IoT driven media recommendations tries to address the challenge of privacy by putting forth an abstract call for more control while the subtleties around its manifestation in media experiences is yet to be defined.

Trepkeviciute (Trepkeviciute, 2017) also recommends transparency as a response for alleviation of privacy concerns in this context, stopping at describing these solutions of control and transparency as 'complicated' while leaving the space for exploration of this 'complication' open to future research. Wijnand IJsselsteijn extends this list of values to include trust as a challenge confronting the use of personal data in novel media. He says that while personal data affords "new horizons for personalised, interactive and immersive entertainment as well as marketing...[…].... At the same time, some such proposals may be at variance with human values many of us hold dear, including privacy, trust, and control." (IJsselsteijn, 2017)

Hook's work (Hook, 2018) that explores the design space of data in interactive storytelling highlights the need for data usage to tell more engaging stories and the potential it creates for exploration in adaptive media which features non-linear pathways. Here, he introduces the concept of ethics and data ownership as yet another set of 'complex issues' that would be crucial in "creating positive audience perceptions and responses to interactive video content that is based on personal data".

While these recommendations help add more dimensions to this exploration of the challenges of personal data use in media, it is apparent that such considerations within media research is currently sparse and shallow. These findings often limit themselves to glossing over these proposed concepts of control, agency, transparency, ethics and trust with little effort rendered to unpacking their impact on the landscape, their manifestations in the context of media, the practicalities associated with it or potential response considerations for these challenges. Considering the emergent nature of the turn towards personal data in media experiences which justifies the lack of substantial critique surrounding the same within media research, this discussion now turns to literature from other disciplines that have managed to engage in more invested exploration of the challenges associated with personal data use.

2.3. Personal Data : A Global Phenomenon

Personal data has been a resource not just in media but long before media experienced the turn towards it, personal data had been a driver of innovation and cutting edge research in many domains. In 2009, the then European Consumer Commissioner, Meglena Kuneva stated that "Personal data is the new oil of the Internet and the new currency of the digital world"(Schwab et al., 2011). Here, the exact implication of this statement is contestable because while the prime position personal data holds as an asset class is undisputable (Schwab *et al.*, 2011), it also stands separate from traditional forms of assets like gold or oil due to some very unique properties its embodies.

These unique properties of personal data, as outlined by the World Economic Forum, are as follows :

> *"The intangible nature of personal data means it can be copied infinitely and distributed globally, thereby eliminating many of the physical barriers that exist for the trade of tangible goods.*
>
> *Data, unlike most tangible assets, is not consumed when used; it can be reused to generate value.*
>
> *Data grows ever more connected and valuable with use. Connecting two pieces of data creates another piece of data*

and, with it, new potential opportunities (as well as new potential harms).

The role of the individual is changing. Individuals are no longer primarily passive data subjects. They are also increasingly the creators of data. In addition, personal data is intimately linked with an individual's background and identity, unlike interchangeable commodity goods."(Kalapesi, 2012)

The power of personal data being an asset class combined with these unique properties makes it the center of the broad technological transformation that is termed the Fourth Industrial Revolution which is redefining the very manner of interaction of individuals to institutions to industries to governments. On this scale, the opportunities it affords extrapolate way further from technological advancements like making better recommendations and predictions (Amatriain, 2013) to national level efficacies that saved the US $700 billion in health cost (Kalapesi, 2012) or responding to global crises like measuring human impact of crises (UN, 2018) or detecting pandemics (Dugas et al., 2012).

While previously personal data might have been associated to computer terminals alone, with the wide scale adoption of smart phones and the increasing popularity of the Internet of Things devices, personal data is being created, collected, stored, shared, processed, analysed, transferred, and copied at an ubiquitous scale. This situation is only further amplified with the projection of 50 billion internet connected devices by the year 2020.

As these billions of data sources go online and mine for more data, there is the explicit call for proper management of this overwhelming new 'ecosystem' (Kalapesi, 2012). Owing to the magnanimity of the situation, this call is put forth to a number of stakeholders simultaneously.

- To the governments and regulatory bodies to help build in rules and laws that establish standards for proper management of personal data
- To industries to abide by these set standards while building in measures that further cater to their specific domains
- To institutions to actively subscribe to data management measures that uphold their customers' best interests
- For the end-users to become more informed and active in this ecosystem that is fuelled by the very data that is created by them.

But for these ideals to be practicalised there are a number of challenges confronting their effective realization. Owing to the highly complex nature of this landscape that forms the confluence of multiple disciplines with investments from various stakeholders, the study of these challenges are often siloed and spread across various disciplines.

The rest of this chapter lists and elaborates these challenges of trust, privacy, security, transparency, control and accountability, using interdisciplinary literature to effectively frame previous research considerations around them.

2.3.1. 'A Crisis in Trust'

In 2012, the European Justice Commissioner, Viviane Reding claimed that "72% of European citizens are concerned that their personal data may be misused, and they are particularly worried that companies may be passing on their data to other companies without their permission" (Reding, 2012). This percentage explicitly showcases the lowered levels of confidence and trust users show in the service providers collecting personal data from them and the processes they employ for data centric endeavours.

A Boston Consulting Group Research Report on the internet economy helps show the impact of such lowered user trust. In 2012 it predicted online retail to grow to US$ 2 trillion in the next four to five years. But this estimate was heavily influenced by consumer perception of trust in the underlying personal data collection and use. While with enhanced trust this growth would be faster (US$ 2.5 trillion), with eroded trust it would be significantly slower (US$ 1.5 trillion). (Dean *et al.*, 2012)

The World Economic Forum report of 2012 (Kalapesi, 2012) which details a 'crisis in trust', elaborates on these findings to help give better perspective to the impact this challenge would have on a holistic level. "Given that this US$ 1 trillion range is from just one small part of the broader personal data ecosystem, it provides an indication of the magnitude of the potential economic impact when other sectors (health, financial services, etc.) are considered – potentially in the tens of trillions of dollars." (Kalapesi, 2012)

Thus, we see how the loss of trust associated with personal data use could have significant economic impact (of a few trillion dollars), if not managed diligently. While fiscal factors are a major driving force in the current society, loss of trust also encumbers other potential socio-technical implications. Lowered trust could lead to tampering of the reputation of service providers, lesser user engagement with data driven technologies and barriers to adoption of emergent technologies, all of which could slow down the very progress rate of innovation significantly. Also, to be considered, representative of the general landscape of personal data use, and as

in the case of privacy, trust is also context dependant. For example, consumers are on average 30% more willing to share data with e-commerce companies, cable operators and automobile manufacturers than with Web 2.0 communities (Rose, Rehse and Röber, 2012). Hence, while the philosophical underpinnings of these hypotheses seem obvious, their empirical substantiations are to be evaluated in various contexts, in accordance with the appropriate variables and scenarios involved.

Responding to this crisis in trust cannot be abstracted into simple steps. Instead, for effective alleviation of the challenge, the contributors to this loss of trust should be identified and responded to. Weaving together literature from diverse disciplines, the following challenges could be seen as concerns eroding user trust in personal data driven technologies.

- Security Breaches
- Privacy Concerns
- Lack of Transparency
- Control
- Accountability

Apart from dialogue around these challenges and responses to them, literature from **law** and **ethics** have also seen interest in the challenges posed by personal data use, particularly responses to these challenges. The section also looks into a brief presentation of these literatures so as to provide a more comprehensive set of perspectives around the challenges posed by personal data use.

It is towards unpacking these interdisciplinary challenges and potential responses that we now turn.

2.3.2. Security Breaches

Along with the rise in importance of the use of personal data there have simultaneously been a number of popular security breaches, identity thefts and fraud, all owing to some form of compromise in using personal data, which has led to widespread user discomfort.

One of the earlier examples of this is the 2011 Sony security breach where names, addresses and even credit card details of more than 100 million people were compromised (Arthur, 2011; Baker and Finkle, 2011; Miller, 2011). This resulted in a monetary loss of around $1-$2 million dollars for Sony but the greater cost being that of customer distrust. The year 2011 took a double hit with respect to fostering user disconcert in organisations leveraging personal data as hackers broke into the Epsilon database, which was responsible for handling

40 billion marketing e-mails on behalf of 2,500 customers, including such companies as Best Buy, Disney and Chase (Lennon, 2011).

Due to the widespread coverage of such incidents and their impact on user attitudes towards personal data driven technologies, research in security responses has seen phenomenal advancements recently. To state some of the latest advancements, Intel's Authenticate (Hachman, 2016) integrates authentication into the very hardware of a system wherein a simultaneous combination of hardware enhanced factors are required to validate the user's identity. This level of security with hardware authentication becomes particularly crucial when the upsurge of Internet of Things devices that are trying to go online has to be carefully authenticated. The use of User Behaviour Analytics (UBA) to trigger red flags when a deviation from expected user behaviour is experienced is another avenue that is seeing promise in the security domain. Thus, if an account is compromised, while in the traditional scenario a hacker could just log in and engage in malicious behaviour, UBA could help flag these deviant actions and look into appropriate security measures (Green, 2018).

While these are more back-end security innovations, there is also a lot of interest in improving data security measures by making them more user centred and useable. For example, the work by Bursztein et al. (Bursztein et al., 2014) looks into the effectiveness of CAPCTHAs, the puzzles used on websites to control abuse caused by computer algorithms posing as human beings. While CAPTCHAs might be effective with this purpose, they also cause inconvenience to real users, negatively contributing to the website's usability and discouraging its proper use. Ur et al looks into the design and evaluation of a data driven password meter that "provides accurate strength measurement and actionable, detailed feedback to users" (Ur et al., 2017) through the use of neural networks and numerous combined heuristics. Their research resulted in "more secure, and no less memorable, passwords than a meter with only a bar as a strength indicator." (Ur et al., 2017) Sawaya et al, in their work conducts an extensive cultural study of the universal security defences by applying the Security Behavior Intentions Scale (SeBIS) (Egelman, Harbach and Peer, 2016) to 3,500 participants from seven countries. Their results showed that "what people think they know affects their security" (Sawaya et al., 2017) more than what they actually know.

Thus, innovation in data security and making that security increasingly useable for the users is considered a major contributor while responding to security challenges raised by the use of personal data. By re-assuring the users of the enhanced security measures they are provided with and making these measures quick and simple to engage with, users are also given the opportunity to place more trust in the services by witnessing that the service providers are contributing more to securing their personal data.

2.3.3. Privacy

Responses to user concerns surrounding privacy has been an area of much interest in technology sectors in the recent times. When discussing responses to privacy, often privacy and security are talked about interchangeably and this ambiguity warrants clarification. While *privacy and security might have overlaps, they are fundamentally different*.

Security involves protection of user data from "unauthorized access, use, disclosure, disruption, modification or destruction" while privacy deals with "managing the risks to individuals associated with the creation, collection, use, processing, storage, maintenance, dissemination, disclosure or disposal of personal data." (World Economic Forum, 2018) Thus, the scope of privacy is well beyond that of security and could continue to be in a vulnerable state even after the data is secured.

The right to privacy is referenced in the constitutions of over 150 different countries. This notion of privacy "offers protection against outside intrusion into people's homes, communications, opinions, beliefs and identities" (World Economic Forum, 2018). These rights manifest in diverse manners including the 'right to be let alone' (Warren and Brandeis, 1890) to 'secrecy, solitude and anonymity' (Gavison, 1980) to the right to self-determination and control of information (Westin, 1968).

In the context of personal data, privacy has been a long-and-often debated topic because personal data, as the name rightfully states is personal and thus has the capability to expose someones most intimate habits, interests and preferences, compromising their privacy. Owing to this nature of personal data, there is a whole range of research that looks into the sensitivity of personal data (Bing, 1972; Al-Fedaghi, 2007), the details of which are out-of-context for this discussion.

The ability to produce data regarding most aspects of an individual has given rise to the concept of digital identities for physical humans which is shaped solely on the basis of their personal data :

"'Digital identity' is the sum of all digitally available information about an individual. It is becoming increasingly complete and traceable, driven by the exponential growth of available data and the big data capabilities to process it." (Rose, Rehse and Röber, 2012)

With the creation of such identities that are almost like a virtual reflection of one's real world image and that in many ways represent a person in the digital dimension, consumers are growing increasingly concerned about the creation and management of these profiles. The Eurobarometer study of 2011 ((Ec), 2011), a major study which spanned almost the whole of Europe showed that 70% of the users were concerned about their digital identity management by companies

through the collection and use of personal data. The Boston Consulting Group digital identity survey of 2012 (Rose, Rehse and Röber, 2012) further backed this result as it showed that 88% of users consider atleast one industry to be a threat to their privacy.

What the results of these independent studies show is severe consumer concern regarding privacy when using data driven technologies. This highlights that access to such sensitive personal data demands responsibility and accountability on various levels. On a governmental scale, space for privacy through data protection offered by different governments depends upon the very nature of that government and their view of personal autonomy, which then filters down to the industries run under the rules overseen by them and then into the institutions in these sectors and finally the end users. The varying sensitivities of these hierarchies along with the diversity among the various notions and the very context dependant nature of privacy wherein its contextual integrity has been spoken about for years in research (Nissenbaum, 2004), makes the concept of provision of privacy a very complex challenge to overcome, where a generalised approach might not necessarily be effective.

Klaus Schwab, Founder and Chairman of the World Economic Forum said “One of the greatest individual challenges posed by new information technologies is privacy. We instinctively understand why it is so essential, yet the tracking and sharing of information about us is a crucial part of the new connectivity. Debates about fundamental issues such as the impact on our inner lives and of the loss of control over our data will only intensify in the years ahead.” (Schwab, 2016)

Thus, privacy, within the context of personal data driven technologies moves from being a philosophical value to a practical need which, if not addressed in the appropriate manner, poses a significant challenge to the very existence and growth of data driven services, putting forth the call for immediate, appropriate and substantial responses.

The **response** to concerns of privacy in the personal data ecosystem calls for measures that extend beyond just the security measures explained above. While security deals with a portion of the technological aspects of privacy, there is also the need to go beyond the technical to the socio-technical underpinnings of this issue. Historically, privacy has been viewed from many different perspectives : socially, philosophically and legally (Solove, 2008). Human values like dignity (Brownsword, 2008), integrity (Clarke, 2006) and autonomy (Bernal, 2014) have been discussed in line with privacy. To encompass the multi-dimensional nature of this challenge of privacy, there is the need to ground the study of privacy not just around technologies but around the users who bind together the different dimensions of privacy. Studying the users, their culture, their context and their wants and needs, is the method for uncovering the challenges of privacy and moving towards solutions that alleviate this challenge.

In user-centred fields of study like Human Computer Interaction, recent increase in interest around privacy has been exponential (Ackerman and Mainwaring, 2005). In the early days studies had revealed the 'privacy paradox', the gap between user attitudes and behaviours when it came to data privacy management (Norberg, Horne and Horne, 2007). This proved to be a key factor in the design of privacy responses later on, by biasing the design process towards simpler and quicker data settings interfaces. More recent research explores this gap further to understand the subtler motivations behind this gap (Phelan, Lampe and Resnick, 2016) to precisely understand user attitudes and the motivations that lead to their behaviours so that the 'gap' could be overcome through future privacy responses.

Understanding this 'gap' had led to more onus on furthering user-centered privacy responses that are increasingly design-led, thus placing lesser effort and time on the users. An example of such an effort is useable privacy (Iachello and Hong, 2007), an entire sub-domain of study which looks precisely at making privacy more in line with the users' expectations. Applications of useable privacy range from emails (Ruoti et al., 2016), social media (Wang et al., 2014; Usmani et al., 2017), mobile applications (Shklovski et al., 2014; Almuhimedi, 2015) and ubiquitous technologies (Egelman, Kannavara and Chow, 2015) to pervasive displays (Davies et al., 2014) and even children's toys (McReynolds et al., 2017).

It is not just the application areas that are diverse here, but useable privacy also reflects similar expanse and range in terms of the questions they pose and the recommendations they make. This includes exploring user attitudes and behaviours with respect to privacy and security (Egelman, Harbach and Peer, 2016), looking into children's privacy preferences about their content being socially shared by parents (Moser, Chen and Schoenebeck, 2017), exploring digital divides and socioeconomics (Redmiles, Kross and Mazurek, 2017) and youth perspectives on data literacy (Hautea, Dasgupta and Hill, 2017).

Useable privacy is not the only user-focused line of study in privacy. Other design-led privacy iterations that have been of interest lately are Privacy Enhancing Technologies(PETS) (Camp and Osorio, 2003), Privacy Impact Assessments (PIAs) (Wright and Hert, 2011), particularly for Radio Frequency Identification (Spiekermann, 2012), privacy engineering (Spiekermann and Cranor, 2009; Dennedy, Fox and Finneran, 2014; Oliver, 2014) and Privacy by Design (Cavoukian, 2009; Spiekermann-Hoff, 2012), all of which respond to the call for enhanced privacy measures when using personal data, in varying capacities.

Thus, these wide range of responses that are being studied to respond to the varying types and levels of privacy concerns, especially with special regard to the role and implications on the users are expected to alleviate user concerns of privacy in the personal data ecosystem.

2.3.4. Transparency

The challenge of lack of transparency in data driven technologies has led to the importance of the provision of transparency being discussed in industry reports (Kalapesi, 2012; World Economic Forum, 2018), academic research papers (Krishnan, 2016) and in legal frameworks like the GDPR (EU GDPR, 2016). One major area of focus when it comes to lack of transparency is consent mechanisms like Terms and Conditions statements and Privacy Notices (Jones, Sailaja and Kerlin, 2017; Sailaja and Jones, 2017).

Users effectively enter into legally abiding contracts when they are passively forced to accept these long and legalese documentation that is presented to them (McDonald and Cranor, 2008; Cate, 2010). These documents are often framed in manners that make them practically incomprehensible to the average man. Research has shown that if one were to read all the Terms and Conditions statements presented to them in a year, it would take an average of 244 hours (McDonald and Cranor, 2008). Such convoluted mechanisms of consent that present little legibility, understanding and freedom marks an inauspicious beginning to the process of sharing personal data further leading into a relationship embedded in doubt and distrust. An ongoing 'strained' relationship further exacerbated by events like Cambridge Analytica (Cadwalladr and Graham-Harrison, 2018) has only contributed negatively to user trust in personal data driven technologies, by almost reinforcing their fears.

According to European Data Protection directives (EU, 1995) it is mandatory to provide users with privacy policies to ensure ethical exchange of data. These documents are expected to serve multiple and contradictory roles (Schaub et al., 2015). The first is to offer legal protection (through proper documentation) to companies in instances where compliance to the terms agreed upon are contested (FTC, 1998). The second is to notify users about the data exchange agreement they are entering into, thereby providing more transparency of the data exchange.

Design-orientated research is actively seeking alternative methods of presentation and communication to make this goal of provision of transparency effective in reality. These **responses** include labeling (Kelley et al., 2009) or provision of concise and standardised summaries to help users find information quickly and accurately; privacy warnings and identification of benefits and risks (Bauer *et al.*, 2013; Schaub *et al.*, 2015) to help support easier user assessment of privacy risks; iconography (McDonald and Cranor, 2008) to help visualise the data practices in more user friendly manners; the importance of timing when studying transactional notices (Good et al., 2007) and layered and contextual privacy notices that provide users with varying levels of information detail, to suit their interest and context (Pinnick, 2011).

From design solutions discussed here to tools (Luger et al., 2015) and best-practice recommendations (Schaub et al., 2015) that foster the design of improved consent mechanisms all show active interest in realising practical and effective solutions that provide efficient methods that respond to the challenge of lack of transparency in data practices.

2.3.5. Control

One of the major challenges feeding the loss of user trust is the lack of control users face when it comes to their personal data. Consumers are often expected to share their personal data and relieve all control they could exercise over it.

This becomes a counterintuitive situation also when working towards alleviating the challenge of privacy as studies have shown that consumers who are able to manage and protect their privacy are up to 52% more willing to share information than those who are not (Rose, Rehse and Röber, 2012). Currently there is an explicit asymmetry when it comes to control over personal data between the end users and the service providers and this asymmetry could be considered as significantly disturbing the balance in the personal data landscape.

This asymmetry is reflected upon by Jaron Lanier in his book "Who owns the future" (Lanier, 2014) where he speaks about how people often don't even understand that they are not involved in a fair trade when they exchange their personal data for services. They consider themselves to be 'first-class participants' in the trade when they are not because they do not have the same ability to exercise control over their personal data as the corporations to make a bargain, which puts them in "a structurally subordinate position" (Lanier, 2014) from where they are forced to accept whatever is handed down to them.

This situation of imbalance calls for solutions. A unique added necessity for solutions of control here to be effective is the need to also address the gap between user attitude and behaviour when it comes to exercising control over data. This gap, praxeologically termed the 'privacy paradox' stands for users expressing the need for control in privacy preserving situations but showing behaviours of low engagement when presented with such alternatives (Norberg, Horne and Horne, 2007).

This gap highlights the need not just for control mechanisms, but for effective control mechanisms that are simple, easy and quick to learn and use. This is supported by the results of a Boston Consulting Group survey (Rose, Rehse and Röber, 2012) which showed that "individuals who had a relatively easy time managing their privacy were 37% more willing to share [data] than those who had a relatively difficult time."

And once they are empowered with such effective control mechanisms, users are moving on from being passive observers of the digital economy to being actively engaged with service providers in a collective dialogue (Levine et al., 2009) over the use of data created by and about them. This kind of engagement has birthed a new class of personal data termed volunteered personal information or VPI which includes updated contact details as they happen, reasons why an action was made, future plans and preferences. It is projected that the accuracy and inherent opportunity of such voluntarily shared data is so high that it could reach approximately $32 billion in value by 2020 in the UK alone (Kalapesi, 2012).

Thus, enhanced user control over personal data, is currently a challenge that needs to be overcome. But once worked upon successfully, it offers the possibility of improving user concerns over their privacy, increasing user engagement with the digital economy and the technology and upgrading the current state of affairs in innovation of personal data driven services by improving the accuracy and quality of the data itself.

Traditional **responses** to this challenge of control has been quite varied, often depending upon the perspective of the origin of the response. For example, Hann et al. (Hann *et al.*, 2002) defines the promise of control as "allowing information to be disclosed only with the subjects' permission"

Culnan and Armstrong (Culnan and Armstrong, 1999) suggest the importance of giving the consumer "voice and control" over actual outcomes but limit their definition of control to people being able to object to uses of their personal data when it is used for purposes other than what it was collected for. (Culnan and Armstrong, 1999)

Malhotra et al.'s (Malhotra, Kim and Agarwal, 2004) use of social contract theory in the context of personal data control also limits its scope in terms of the existence of voice or exit such as opt–out.

Stewart and Segars (Stewart and Segars, 2002) also acknowledge the need to give consumers control over their data by stating that even a "small degree" of control over their personal data would result in the consumers viewing an information practice as less privacy invasive. This "small degree" of control is manifest as asking consumers for permission to use their personal information for secondary purposes and providing access to personal data to verify its accuracy.

Hui et al. (Hui, Hai Teo and Tom Lee, 2007) suggests that control over personal data, is restricted to the actions of the service provider organisation with little role from the users' side once the information is disclosed. Van Dyke et al. (Van Dyke, Midha and Nemati, 2007), by introducing the notion of "privacy empowerment", presents reason for shifting from this model to one that gives more control over personal data to the users by showing that "those firms which meet

the demand for control through empowering the consumer are rewarded with lower levels of privacy concern and increased trust".

Expanding on Van Dyke's notion of 'privacy empowerment', the state of the art in personal data control research has recently been witnessing the rise of a class of technology termed Privacy Information Management Systems (PIMS). PIMS suggest a central role for the users in the personal data ecosystem by helping them maintain a more active and engaged relationship with their personal data. Users retain control over data sharing, oversee who accesses their data, and have increased transparency around purposes of data use.

PIMS can challenge the dominant business models of cloud-based systems where users 'pay' for services with their own personal data and can enable a shift away from services aggregating and analysing data to derive value from these. Instead, with PIMS users can permit third parties use or access to their data through machine readable consent terms, exercise choice over running personal analytics services, and maintain oversight of their data through dashboards, thus giving the users a chance to participate in the digital economy in a more active manner. (Urquhart, Sailaja and McAuley, 2017)

Examples of such PIMS or PDVs include openPDS (de Montjoye et al., 2014), Mydex (Mydex, 2017), Higgins (Eclipse wiki, 2009), OwnCloud (OwnCloud, 2017), LockerProject (Project, 2012) and DataBox (Amar, Haddadi and Mortier, 2016). While, a detailed comparison of these technologies are beyond the scope of this thesis, it is safe to claim that they all, while working towards the goal of improved user control and management of user personal data, are all distinct in the affordances they offer and limitations they pose (detailed comparisons in our PUC paper (Urquhart, Sailaja and McAuley, 2017)). But, the wide breadth of research in such an alternative class of privacy preserving technology helps to highlight the promise it shows in future endeavours involving personal data control, while leaving space open for further study and evaluation of their effectiveness in practical contexts.

2.3.6. Accountability

"Data protection must move from 'theory to practice' … In the discussions on the future of the European and global data protection framework, accountability based mechanisms have been suggested as a way of encouraging data controllers to implement practical tools for effective data protection." (Opinion 3/2010 on The Principle of Accountability, 2010)

Within legislation accountability is viewed through one of two lenses, that of 'internal' accountability and 'external' accountability.

“Internal accountability refers to the policies and procedures that a ‘data processing’ organisation– e.g., government department, commercial entity or charity – puts in place to demonstrate to itself that its data processing operations comply with the requirements of proposed data protection legislation” (Crabtree et al., 2016)

The EU proposal (European Commission, 2012) lays down a number of legally binding requirements (Articles 22 to 43) for ‘data controllers’ to comply with. These mandates stipulate detailed documentation of data processing operations, security provisions, and the risks that attach to the categories of data being processed.

U.S. legislation does not put forth any such mandate, instead it works with a voluntary code of conduct enforced through Federal acts prohibiting unfair or deceptive practices. However, it does suggest the use of Privacy Impact Assessments as an effective means of conducting ‘structured assessments’ for potential privacy issues (House, 2012).

While the mechanisms of the proposals put forth by the US and the EU are dissimilar, their agreement on the need for rigorous internal accountability is obvious.

“External accountability refers to the policies and procedures a data processing organisation puts in place to demonstrate to others outside the organisation that its operations are in compliance with the requirements of proposed data protection legislation” (Crabtree et al., 2016).

Here, ‘others’ point primarily to either the regulatory bodies or the users whose data is leveraged by the organization. In proposed EU legislation these policies fall under ‘consent’, included in Articles 6 to 20 whereby it says that for data processing to be lawful the data subject has to give consent to the processing of their personal data (European Commission, 2012).

Accountability in a scenario as complex as the personal data landscape deserves to be discussed more than from just a legal persepctive. Computer Science research talks about computational accountability (Crabtree et al., 2016) which refers to the legibility of systems to end users. This notion is popular in the HCI and CSCW circles where the challenges of making the behavior of a system ‘observable and reportable’ or in other words enable human computer interaction is leaned upon (Dourish and Button, 1998). Within CSCW, the very core concepts of the field is expected to promote design that negotiates the boundaries between (networks of) objects and people, making them transparent, understandable and hence accountable (Robertson and Wagner, 2015).

Yet another notion of accountability has been spoken of recently, that of social accountability (Nilsson *et al.*, 2019). This dimension of

accountability extends beyond computational accountability to being accountable to the broader social context within which they were prospectively situated. Here, standards include more than just the legibility of a computational system to the expectations concerning the appropriateness of autonomous behaviour within a social setting, which includes reasonings around acceptability, agency, trust and potential risks.

Responses to this call to improved accountability have been of various kinds. Enabling improved transparency and control of the data practices are some of the premier among these. These responses have been discussed in detail in Sections 2.3.4. and 2.3.5. Responses to privacy challenges (discussed in Section 2.3.3.1.) are also expected to alleviate concerns around accountability. The same is the expectation around privacy management systems like PIMS (discussed in Section 2.3.5.).

Crabtree et al (Crabtree et al., 2016), in their work explicates how DataBox, one such example of PIMS builds accountability into the personal data eco-system through responding to the issues of transparency, control, privacy concern etc.

DataBox is a personal networked device (and associated services) that collates and mediates access to personal data (Amar, Haddadi and Mortier, 2016). The DataBox would exist as a physical object in the home which enables the users with the possibilities of physical interactions with it. It is a solution which could help data to be collected, combined and processed in ways that would let the users explore workflows and manage their own data and data involving other stake-holders (through ad hoc social interaction).

DataBox ensures client-side processing of user personal data, thereby ensuring that all data stays on the box, within user homes, where it would be processed by one or many service provider applications that run on the box (much like a smart phone with its multiple applications). These DataBox applications would provide the users with SLAs that allow users to see and understand what types of data they would be exchanging, provide space for negotiation of exchange of this data and allow for explicit control over this data as well.

Thus, the interaction mechanisms provided by this system make data processing ‘observable and reportable’ and thereby computationally accountable. This allows the users involved in the data processing to articulate not only their own actions but also, in surfacing relevant behaviours of computational systems, the actions of underlying machines. In thus “articulating the actions of actors and machines the IoT-Databox model reflexively provides the transparency and consent, granular choice, access and data portability required to make data processing accountable to the data subject and enable individuals to control the flow of data from IoT devices into the digital economy. In doing so it also enables data controllers to demonstrate compliance

with the external data subject accountabilities required by proposed legislation." (Crabtree et al., 2016) Thus, the Databox becomes beneficial to both data subjects and data controllers who leverage personal data, thus allowing them to benefit from the use of personal data while at the same time managing the widespread threat to privacy occasioned by it by employing improved accountability.

2.3.7. Legal Responses

The international regulatory landscape concerning personal data is complex to state the least. There are more than 120 different national laws pertaining to the same with more new laws being worked upon in the pipeline. Owing to the massive scale of this situation, this section examines the case of two major entities, the U.S. and Europe, two significant players with quite different perspectives towards data use, when viewed at a very high level.

The United States government has put forth a comprehensive online Privacy Bill of Rights that is intended to help individuals exercise more control over their personal data on the Internet (The White House, 2012). Its major goals include security, user control, access, accuracy, accountability, transparency and focused collection. In this context, the US Federal Trade Commission (FTC) is the delegated enforcement authority who oversees that stakeholders successfully translate the goals of the Bill to practical measures that help the everyday users.

In Europe, the same is approached from the perspective of protection of fundamental rights (European Union, 2018) in the Data Protection Directive of 1995 (EU, 1995). It shares the ideals of greater user control and trust with the U.S. as it expects Internet-based organisations to obtain explicit consent from their users for personal data collection and use. This legal framework was recently replaced by the General Data Protection Regulation(EU GDPR, 2016) in May 2018 wherein the law is the same for all EU members, without need for domestic regulation. The GDPR, being framed in the era of extensive personal data use, as opposed to the pre-Internet DPD, is much more thorough in its' efforts to uphold trust in the digital economy. It's major principles include lawfulness, fairness and transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality (security) and accountability and provides more grounded and comprehensive ways of achieving these principles.

In Europe, outside of the GDPR, there is more policy support that helps with the proper navigation in the personal data landscape. For example, Information Privacy by Design is a tool that encourages designers to consider privacy risks of their data driven technologies as early as possible, thereby providing safeguards that ensure privacy by default. This is in line with the GDPR which asks for safeguards during processing (EU GDPR, 2016), to be considered at the 'time of

the determination of the means for processing and at the time of the processing itself' that ensure that by default the processing is :

- For 'personal data which are necessary for each specific purpose of the processing';
- Controlling the 'amount of personal data collected, the extend of their processing, the period of their storage and their accessibility'
- That 'personal data are not made accessible without the individual's intervention to an indefinite number of natural persons' (EU GDPR, 2016)

Thus, these legal responses to the use of personal data by technologies can be seen as very much active in current times. But while legal frameworks propose concepts like accuracy, fairness and transparency, the gap between the proposal themselves and their effective practical realisation is still pronounced. This calls for technology designers to step up their roles and act more like regulators through their design itself. Black proposes a similar notion through the possibility of 'decentered regulation' (Black, 2001) wherein the designers can enable regulation through their work. Here, Privacy by Design is a good example of such an initiative. Another one being Luger's (Luger et al., 2015) work with ideation cards designed around privacy regulations, which gives designers a more streamlined manner of approaching the matter. But, while such line of thought is gaining momentum, it still requires much dialogue and study that help form concrete guidelines for designers to confidently step into the space of policy, regulation and law, while also overcoming issues like lack of authority and legitimacy (Urquhart, 2016) that they face when engaging in such regulatory design.

2.3.8. Ethical Responses

Ethics and philosophy have been discussing the concept of privacy for decades. While accounts on privacy term it as 'vanishing', 'diminishing' and 'evaporating', Professor Deborah Nelson puts it as "Privacy, it seems, is not simply dead. It is dying over and over again" (Nelson, 2001). To overcome this 'death' of privacy, there has been many proposals on how to view privacy, so that it can be renovated and restored. An example of this is Daniel Solove's work which conceptualises privacy based on (1) method, (2) generality, (3) variability, and (4) focus (Solove, 2008). Here, method refers to privacy, consisting of "many different yet related things"(Solove, 2008), in line with Ludwig Wittgenstein's notion of family resemblances (Wittgenstein, 2009). Generality refers to "privacy should be conceptualized from the bottom up rather than the top down, from particular contexts rather than in the abstract." "Regarding variability, a workable theory of privacy should account for the differing attitudes toward privacy across many cultures." And,

“Finally, an approach to conceptualizing privacy must have a focus.” As “otherwise it merely picks at privacy from many angles, becoming a diffuse and discordant mess” (Solove, 2008).

While Solove’s work is just one instance of the dialogue around privacy in philosophy, technology, particularly privacy and even personal data use is seeing an increasing interest in ethics, morality and philosophical underpinnings. These include but are not limited to Nissembaum’s work on the contextual integrity of privacy (Nissenbaum, 2004), Ryan Call’s work on the boundaries of privacy harm (Calo, 2011), Paul Peter Verbeek’s work that looks at the morality of technology (Verbeek, 2011) and the mediated value of privacy (Kudina and Verbeek, 2019), Urquhart’s work that marries technology and regulation to ethics (Urquhart, 2018) and Rosner’s exploration of “Who owns your data?” (Rosner, 2014).

Such considerations around ethics extends the possibility of design of ‘ethical’ systems which support the users beyond their intentional actions to assuage the implications of even their involuntary behaviours with a technology. Weiser and Brown discussed the role of such design as a core step in the realisation of ubiquitous calm technology (Weiser and Brown, 1997) which in turn becomes a necessity with data driven technologies being part of the most mundane activities of a user’s life.

All of this thought and dialogue around the ethics of personal data use by technologies has resulted in ethics being used as a measure for responding to the risks posed by personal data use. In June 2018, the European Data Protection Supervisor (EDPS) launched a public initiative on Digital Ethics which was “aimed at exploring the consequences of new digital technologies [like personal data driven services] on society and their implications for data protection and the right to privacy.

The consultations began with a published Opinion (EDPS, 2015) in 2015 which openly urged the EU and other prominent stakeholders in the global personal data landscape to actively employ and encourage ethical approaches to the development of novel technologies. This was followed by the setting up of the Ethics Advisory Group (EAG) who produced the report “Towards a New Digital Ethics” (European Data Protector Supervisor, 2015) which conducts an ethical reflection of the digital age and highlights the need for foundational values like dignity, autonomy, freedom, solidarity, equality, democracy and trust, to be included in the personal data revolution.

Thus, ethics is employed by philosophers to regulatory entities to help better understand and assuage the concerns of privacy and trust brought up by personal data. The 2018 Public Consultation on Digital Ethics (EDPS, 2018), which involved international stakeholders from various sectors, shows that more than 80% of the respondents affirmed that ethics concerning novel technology is very much on

their agenda, describing it as ‘important’ ‘extremely relevant’ or even ‘mandatory’. This goes to show that ethics is no longer a priority for academics, philosophers and high-level policy makers, but is becoming a potential active response path for service providers as well.

Thus, the turn towards ethics as a medium for response to the challenges in the personal data eco system seems almost inevitable. But here again, just as in the case of the legal responses, but even more so, there is a pronounced gap between the theories proposed and conclusions made, which often abstract away from the real world need for practical solutions.

2.4. Importance of Context Specific Study

The importance of context when studying information flows was discussed by Nissenbaum (Nissenbaum, 2004) where she discussed the ‘contextual integrity’ of individual privacy. Here, she says that context plays an integral role when evaluating the appropriateness of information flows. Appropriateness is no longer abstract but well defined, based on the experience of the users. Solove’s philosophical reasonings around generality, variability and focus of privacy also echos the same ethos of respecting individual contexts as the priority here is to consider the variances in individual attitudes in diverse contexts when trying to effectively conceptualise privacy. A report by the World Economic Forum (Kalapesi, 2012) also highlights a very similar characteristic being mimicked in scenarios of personal data exchange and user trust associated with it whereby the importance of context when considering user personal data is repeatedly talked about.

This calls into question the need for exploration of concepts of personal data exchange, privacy and trust in depth for different disciples and contexts. Section 2.3. shows, literature from various disciplines that have indulged in conversations around the challenges of using personal data. But oftentimes these discussions are generic, putting forth a call for context specific study of these challenges, so as to enable effective responses that help alleviate their implications.

As Dourish and Anderson (Dourish and Anderson, 2006) has pointed out in the case of privacy, the way to understanding it is not by making assumptions from an external perspective but by observing the interactions and negotiations of users with technology and studying what emerges from it. Thus, in order to gain a holistic understanding of the challenges of using personal data in media, there is the need to consider the context of media and the role of personal data within it.

2.5. Research Gap

This literature review begins by looking into the turn towards personal data use in media experiences, followed by the challenges of this shift, as engaged with in media research. Due to its emergent nature, the focus here is often on the media experience itself and the innovation facilitated by personal data. This priority exposed a lack of substantial critique around the challenges of using personal data in media experiences whereby privacy was often cited as a challenge while 'values' like trust, ethics and transparency received sparse mentions.

Thus, the review then turned to other literature that engaged more fluently in conversations around the challenges of using personal data. This multi-disciplinary review of previous work highlighted trust as an overarching challenge with others like privacy, transparency, control and accountability feeding into it. Literature from Law and Ethics, which also engages with various challenges of personal data use, often with the view of responding to them has also been presented here.

These literatures showed depth and breadth of scope, but they were often generic discussions that warrant closer inspection of particular contexts. Such close contextualised study is required to explicate the nuances associated with these high level challenges within the scope, uniqueness and practicalities of a considered context, so that the dialogue around these challenges could move from high level philosophical debates to practically applicable responses.

Also to be highlighted here is the lack of active involvement of the users in many of these studies. Most of the challenges and responses mentioned here, while subscribing to the symbolic presence of the user as an important stakeholder, has not actively involved the users to take their perspectives into account. When shifting from generic high-level abstractions to closer, practically rooted study of the challenges of using personal data in media, there is the explicit call to involve those holding significant stake in this economy. Within the context of data driven media experiences this points to both the users whose data is leveraged and the service providers who leverage this data.

Combining these calls for context specific study of the challenges of using personal data, the aforementioned current lack of critique on these challenges within media literature and the potential of studying this context through active user involvement surfaces **a research gap** for the exploration and explication of the challenges of using personal data in media experiences. The rest of this thesis reports three studies that synergise to respond to this gap by exploring user and service provider viewpoints on the challenges of using personal data in media experiences.

3. Methodology

This chapter outlines and elaborates the methodology employed for the research presented in this thesis. The studies, which explore the challenges associated with personal data use in novel media employs a mixed set of methods that are selected and leveraged to support the particular goals of each study reported here. The first part of this chapter elaborates the various methods used within the studies, starting with the **data collection methods**, followed by the **data analysis methods**. The second part of the chapter presents the **methodological configurations** used within each study, along with the reasoning and justification for this selection.

3.1. Introduction

The research, which was conducted through three studies, used a mix of methods to accommodate the requirements and goals of each study. Traditionally, the mixed methods approach has been seen to have various connotations in research where the triangulation of data from a mix of qualitative and quantitative data collection methods (Halcomb and Hickman, 2014) is the most common one. The approach referred to here simply refers to the selection and leverage of a certain set of research methods employed to help design, build and conduct the studies in order to elicit the appropriate kind of data from them.

Given the emergent nature of the turn towards personal data in media experiences and as seen from the research gap highlighted in the previous chapter, a call for a holistic view of the challenges facing the use of personal data in novel media experiences is made apparent. To respond to this call for a more 'holistic' and comprehensive view of the challenges, the use of a mix of methods was chosen to help effectively understand the perspectives of different stakeholders within this landscape. This approach to research, of looking at a problem through different lenses is inspired by the concept of viewpoints within software engineering where the importance of considering various perspectives is advocated for (Sommerville, 2011).

While there are several players or stakeholders within a personal data ecosystem (which includes the users, service providers, third party agencies dealing with the data, advertisers, data brokers, regulators etc.), the research, reflecting upon the current lack of user perspectives (identified in the previous chapter : Section 2.5) within research in this domain, chose to begin exploration by studying this population. Thus, the first study aimed at understanding the challenges from the users' perspective, where a study that piloted the idea of a media experience leveraging user personal data was presented to the users to elicit their response.

Once the user viewpoints around these challenges were identified, the next step was to begin looking into alleviation of these challenges. But, in order to work towards effective responses to these challenges, a good understanding of the challenges faced by stakeholders providing these data driven media experiences was seen as a priority. Hence, a second study was planned and executed, this time involving media service providers from various teams within the BBC, to understand the challenges associated with using personal data in the media experiences they create.

Studying the challenges from both the users' and service providers' perspectives helped reinforce the findings from both studies by highlighting the core challenges, while unpacking the subtleties associated with each of the populations studied. While the research provided the option of extending to more studies involving viewpoints from other stakeholders, the data from the first two studies showed ample overlap, indicating enough saturation to effectively use it to design responses that would help alleviate the challenges for these two major stakeholders. This could then be extended through future work to include feedback and perspectives from other stakeholders. Thus, the third study was conducted, which bridged the findings of the first two studies to design a provocative data driven media experience that embodies responses to some of the challenges uncovered. This media experience was presented to both the audiences and the service providers to engage with, give feedback to and think more about the challenges and responses, in a first hand, immersive manner.

These studies employed a number of research methods to realise their motives. These methods could be categorised as **data capture methods** which were used to engage, probe and elicit data from the users and **data analysis methods** which were adopted as per the demands of the data collected, to analyse it in the most optimum manner that responds to the research questions. The data capture methods used here are scenario based design, focus groups, interviews and design fiction. The data analysis methods used are thematic analysis, grounded theory and endogenous topic analysis.

These methods are configured in different combinations to respect and enable the research goals, questions and agenda. The rest of this section begins with descriptions of these methods followed by explanations of the methodology configurations of the three studies which includes the configurations chosen, how they are situated within the studies and the reason for choice of the same.

3.2. Data Capture Methods

The following section presents and describes research methods used in the studies to capture data from the participants.

3.2.1. Scenario Based Design

Scenario based exploration is an approach that could be used to elaborate a conceptual model of a future system, "a description of the proposed system in terms of a set of integrated ideas and concepts about what it should do, behave and look like, that will be understandable by the users" (Rogers, Sharp and Preece, 2002). Scenario-based design shifts the focus of design by placing the highest priority on the user by " defining system operations (i.e., functional specification) by describing how people will use a system to accomplish work tasks and other activities."(Rosson and Carroll, 2002) Thus, it is built on the foundational belief that "computers are more than just functionality" (Carroll, 1999) as they often restructure human activities themselves.

Therefore, software systems could be considered to be socio-technical interventions that are embedded in different social contexts and hence shape how people 'work' (conduct their jobs, leisure, play, etc.) within them. Understanding the social context in which computers are embedded is essential to do effective systems development as "each context in which humans experience and act provides detailed constraints for the development and application of computer technologies." (Carroll, 1999) Therefore, scenarios are a key means by which understanding of context could be developed and the design itself could then be adapted to suit this context of use.

3.2.1.1. Different Types of Scenarios

The Scenario Based Design Framework defines a series of different scenarios which support different design tasks within research. These are as follows :

"A **problem scenario** is a narrative of current practice that synthesizes actors, themes, relationships, and artefacts discovered in the field work" (Rosson and Carroll, 2002). They are developed through various elicitation techniques like interviews, diaries, fieldwork etc. and is a way of doing requirements analysis.

Activity scenarios go beyond user requirements specification to include systems requirements specification by identifying user needs and technology to support them. They have two key aspects, the first is that of analysing current practice, by providing an overview of the variety of goals, actors and sequences of actions involved in current activities that require support. The second is that of projecting future practice as mediated through design by elaborating how technology might support these activities.

Information scenarios focus on defining the information a system needs to provide to meet user needs. It takes the research further into the design process, goes beyond requirements to system modeling and

can be used as a resource for defining information objects and their attributes.

Interaction scenarios focus much on interface design. They define what the user will see and how they will interact with the new system. They include detailed definitions of input/ output mechanisms and how user tasks are actually supported. They form a key resource in evaluating design proposals.

3.2.1.2. Characteristics of a Scenario

Scenarios are explained by Carrol (Carroll, 1995) as "stories". Thus, they are narrative descriptions where the story depends on the kind of scenario they represent. Thus, unlike process models or use cases, they are not visual representations but tools that could be used to develop visual representations. Every scenario is defined through four "characteristic elements". These elements are setting, actors, goals and plot.

The **setting** is where the story is situated. It could be a place (e.g., the home) or can be a situation (e.g., watching a movie). In both cases, the setting implicates people (e.g., residents of the home and guests), resources (e.g., television and internet), and particular episodes of work (e.g., searching for the next piece of content to watch on the television).

The **actor** is a person or a thing who is a doer of an action by taking an active role in the setting. Here, the scenario must define one or multiple relevant parties in the setting who would orient the observer/ researcher to the active participation of someone in an episode of work, thus contributing to building of the scenario.

Scenarios involve **goals** that describe why the actors are doing typical and significant activities in the settings they are located. -"Every scenario involves at least one agent and at least one goal" (Carroll, 1995). Thus, a scenario may describe more than one agent and more than one goal. But minimally one of each.

Reflecting the nature of all stories, scenarios embody a **plot** as well. The plot of a scenario is unique as in "they include sequences of actions and events" (Carroll, 1995). Thus, the plot here refers to how actors achieve their goals within the settings. Scenarios thus describe current or future activities in a setting, including the collaborations involved and the resources used.

When effectively developed with the proper implementation of these properties, a scenario becomes "a concrete design proposal that a designer can evaluate and develop." (Carroll, 1995), while eliminating the need to build anything to conceptualise what a system must do at a deep level. It provides "minimal contexts for developing user-oriented design rationale [so that] a given design decision can

be evaluated and documented in terms of its specific consequences within particular scenarios." It " keep[s] the whole enterprise focused on past and future situations of use [and] they help keep designers focused on what matters most" (Carroll, 1995) which are "the needs and concerns of the people who will use the system being designed" (Carroll, 1995).

3.2.2. Focus Groups

When the research demands richer insights and multiple perspectives that might otherwise be missed by methods like questionnaires and surveys, focus groups are utilised. While providing a platform for eliciting more in-depth data, focus groups eliminate the intensive effort of individually meeting and talking to each participant involved, as demanded by interviews. It provides the attractive alternative of meeting several participants at once, thus presenting itself as a reasonably effective and inexpensive tool for gathering varying perspectives.

The size of a focus group often ranges. while some propose 8 to 12 in a group (Robson, 2011), others suggest fewer members of 5 to 7 to promote effective, in-depth conversations (Krueger, 2014). More than one focus group is suggested to be included in the research as using just a single focus group session could result in the data being unrepresentative or the group being unresponsive. Increasing the number of groups is expected to increase the success of the research(Krueger, 2014).

Participation of a number of individuals in the research is expected to easily surface similarities, differences and variations in perspectives and viewpoints which might otherwise be more challenging to map out. A controlled amount of debate and disagreement would help the conversations flow and highlight new avenues for further exploration in future research (Brown, 1999). Focus Groups also tend to be less structured in nature, helping extend the data to involve broader concerns grounded in the mundane, real-world, everyday activities of the users, which could be overlooked in a fully structured data elicitation setup.

Focus groups also eliminate some of the disadvantages of other methods like interviews where if the interviewee is not talkative, the researcher-respondent dynamic could deteriorate, adversely affecting the data. In focus groups, discussions tend to be more interactive and hence there are higher tendencies of participants balancing each other where they tend to encourage and probe each other to speak both implicitly and explicitly, in support of or in opposition to previous statements. The dynamic nature of these discussions often tend to stimulate respondents to surface issues that might have been overlooked in instances like a one-on-one interview.

Focus groups do pose the disadvantages of managing especially talkative or opinionated participants, limitation on the number of issues or questions that could be covered in a certain time and the need to carefully negotiate and manage differences and sometimes even conflicts within a group. Also, when selecting the participants of a group, there is the need to pay attention to consider the homogeneity/ representativeness, familiarity of the respondents so that the selection benefits the goals of the research agenda (Lazar, Feng and Hochheiser, 2017).

3.2.3. Informal Interviews

When rich and deep user insights become a priority in research, interviews are employed as the method for eliciting data. Here, direct conversations with participants help provide perspectives that might otherwise be missed by methods like questionnaires and surveys where the focus is on quantity rather than the subtleties involved. "By asking questions that explore a wide range of concerns about a problem and giving interviewees the freedom to provide detailed responses, researchers can use interviews to gather data that would otherwise be very hard to capture" (Lazar, Feng and Hochheiser, 2017).

Interviews are used for three key applications within HCI research. The first is initial **exploration** where the focus is on investigating new possibilities and gaining a better understanding of the current practices within an area. Here, interviews are often not focused on specific questions themselves but rather on the needs and challenges presented in the responses of the participants. The second is **requirements gathering** for the design of a new tool where the interviews are still fairly broad, often focusing on the user goals, how they are met by current tools, frustrations etc. (Lazar, Feng and Hochheiser, 2017). Here, a broader outlook is demanded as focusing on a single tool or certain tasks or capacities would narrow the responses to focus on just those, eliminating broader concerns which might be affecting the users' in their everyday lives. Finally, **evaluation interviews**, which are used to elicit user response and feedback about a technology or prototype, helping evaluate various aspects of the technology. Here, if the research demands focused evaluation of certain tasks, tools or design elements, the interviews would be designed to elicit specific responses around those. But if the research agenda is to allow the users to voice their unique and specific reasonings around the technology presented, without the researcher honing in on a specific design aspect, the interviews would be left with enough space to accommodate for the same.

To allow for these different contexts and motives that drive a particular interview study, interviews are presented with different structures, any of which could be adopted to suit the requirements of the research. On one end of the spectrum there are **fully structured** interviews which use a series of set questions put together in a

specific order. These questions are presented to the participants in the order defined by the script and this schema presents very little scope for exploration beyond the script developed, even if the participants' dialogue and situation demands thus. **Semi-structured** interviews help overcome these restrictions of fully structured interviews by allowing for follow-up questions, addition of more questions and comments, if required. Here, the possibility of opportunistically allowing for deeper exploration of the respondents' comments is possible through re-ordering the questions on the script or addition of more questions. **Unstructured** interviews work on similar premises but 'takes this idea to its logical extreme' (Lazar, Feng and Hochheiser, 2017). Rather than a structured script, unstructured interviews consist of an interview guide (Robson, 2011) that would have a number of seeding topics or questions. The interviews start off with an initial question and then the conversations are allowed to flow as per the choosing of the respondents. If the conversation stalls, pauses or digresses beyond the scope of the research, the researcher intervenes by introducing another topic of interest into the conversation. These interviews are considered informant in nature as the interviewee is seen as more in control of the conversations (Lazar, Feng and Hochheiser, 2017). Here, despite the conversation not having a pre-planned structure imposed by the researcher, the dialogue could yet be structured by the respondents themselves, the presence or lack thereof of a structure is dependent upon the viewpoint that is being considered (Robson, 2011).

3.2.4. Design Fiction

Design Fiction derives from an area of design practice termed Speculative Design, where the focus is on utilising design to probe and ask questions (Coulton, Lindley and Cooper, 2018) rather than solve existing problems. This process engages in the creation of prototypes. But rather than study the prototypes to then be put into production, the focus is on encouraging people to engage critically with the potential issues that the design embodies.

The exact definition of Design Fiction is yet under debate, with various researchers taking several different approaches towards it. Bleeker articulates Design Fiction through a set of ideas (Bleecker, 2009) : the intentional use of diegetic prototyping, a means for exploring 'unknown unknowns', a way of looking at pre-existing artefacts to derive insights from them, a making practice, an analytical practice, a subject of and a method for doing academic research or even a service to be sold to clients. Design Fiction is also seen as a method for looking at things (Tanenbaum, Tanenbaum and Wakkary, 2012) or even a construct which can be used in various ways to explain, legitimise or frame ideas (Markussen and Knutz, 2013). Lindley (Lindley, 2018), in his doctoral thesis, while affirming these varied perspectives through elaborate literature review and study, subscribes to the notion of Design Fiction as World Building

to be the approach that coheres with most of the currently held different perspectives on Design Fiction. The same is reflected in the words of Bruce Sterling (the person usually credited with inventing the term Design Fiction). He said, “It’s not a kind of fiction. It’s a kind of design. It tells worlds rather than stories” .(Sterling, 2012; Coulton, Lindley and Cooper, 2018)

Design Fiction is not the first instance of world building in the present day. Coulton, Lindley and Cooper (Coulton, Lindley and Cooper, 2018) point out a range of the same by putting forth diverse examples of world building that leverage varied forms of media in current popular culture. From soap operas like Eastenders that use the screenplay and acting to build a world, to alternate histories like Philip K Dick’s The Man in the High Castle that is built through a novel and a TV adaptation to fantasy worlds like the Middle Part created by J.R.R. Tolkien in his novel and its corresponding Film adaptations, all create worlds that immerse people into another potential reality.

Thus, it is obvious to see how this world building approach renders itself to designing ‘things’ that produce the impression of a future world, in the field of technology. Here, the very essence of the world helps prototype the things in it and the things in turn prototype the world itself, forming a reciprocal prototyping loop (Coulton, Lindley and Cooper, 2018). This loop helps ask important questions like what it will be like to live with future technologies (Lindley, Sharma and Potts, 2014). This is done by creating a fictional version of the future world which embodies the considered technologies, that people can move in and out of so as to experience and understand it and in turn help shape the real world that current technologies are creating (Coulton, Lindley and Cooper, 2018).

3.3. Data Analysis Methods

The following section presents and describes the various research methods used for analysing data collected through the methods detailed previously.

3.3.1. Thematic Analysis

Thematic analysis is a “process for encoding qualitative information” (Boyatzis, 1998) that requires an explicit code. Simplistically speaking, the process is the sequence of recognition of a pattern, followed by encoding and interpretation.

Thematic analysis is used in fields as diverse as literature, psychology, sociology, cultural anthropology, economics, astronomy mathematics etc. (Miller and Crabtree, 1992; Silverman, 2006; Denzin and Lincoln, 2011). Its popularity is partially owed to its capacity to allow easy communication between variant disciplines (Miller and Crabtree, 1992; Denzin and Lincoln, 2011). While

disseminating research to a diverse set of audiences, each of whom bring in their own unique epistemological priorities and orientations, thematic analysis forms a “bridge” (Boyatzis, 1998) that makes this process more accessible to everyone involved. The same reasoning could be translated to HCI, a domain which has traditionally fostered interdisciplinary research and hence thematic analysis is one of the most recommended techniques in HCI to analyse qualitative data.

There are various versions of thematic analysis. Within the scope of this research, we consider data driven thematic analysis as outlined by Boyatziz (Boyatzis, 1998). This consists primarily of three stages : sampling, code development and validation. For the purpose of this discussion, we focus on code development where majority of the data analysis is conducted. This phase begins with the selection and reduction of raw information from the samples where the researcher paraphrases and summarises the data, retaining only material that the researcher considers important and applicable to the research. The second stage is where themes are identified by spotting similarities in the data within subsamples and by comparing them across subsamples. Here, a theme is defined as “a pattern found in the information that at the minimum describes and organises the possible observations or at maximum interprets aspects of the phenomenon” (Boyatzis, 1998). Through these stages the researcher continuously inhibits any form of interpretation by investing all into the development of the themes. The next stage is the development of the code (“a model with themes, indicators and qualifications that are causally related” (Boyatzis, 1998)) which is the process of “seeing as” as explained by Wittgenstein (Creegan, 1989) where a link is provided between newly discovered patterns and previously explored patterns. Here, the themes are filtered to select the minimum number of themes that help maximise the clarity in communication of the results (e.g., those that show an explicit distinction between subsamples) while preserving the phenomenon. Codes for these themes are then developed. These codes are applied to the rest of the data and the themes then clustered either conceptually or empirically to form theme clusters or hierarchies.

3.3.2. Grounded Theory

Grounded Theory concerns the systematic discovery of theory from qualitative data and was developed as an alternative to previously used logico-deductive methods. Here, theory is “a strategy for handling data in research, providing modes of conceptualization for describing and explaining. The theory should provide clear enough categories and hypotheses so that crucial ones can be verified in present and future research” (Strauss and Corbin, 1990).

Research here uses the Constant Comparative (as detailed in (Strauss and Corbin, 1990)) method of grounded theory which is used to generate theories from qualitative data through explicit coding (using Thematic Analysis). The sampling for this process could be done by

employing theoretical sampling where the researcher simultaneously collects, codes and analyses the data and makes decisions about further data collection, thus, allowing the data collection to be 'controlled' by the emerging theory.

The Constant Comparative method consists of four steps. Firstly, "comparing incidents applicable to each category" (Strauss and Corbin, 1990) where the data is coded into categories. Here, coding is done in a simple manner by noting the categories and the comparison group into which the data falls. When an incident for a previously discovered category is coded, it is compared with other incidents in the same and other groups within the category. This continuous comparison of incidents generates theoretical properties for the category. Thereby arises a thorough understanding of the "full range of types or continua of the category, its dimensions, the conditions under which it is pronounced or minimized, its major consequences and its relation to other categories" (Strauss and Corbin, 1990). This step also ensures diversity in terms of the similarities and differences compiled within the data. Secondly, "integrating categories and their properties" (Strauss and Corbin, 1990). As the theory starts evolving at the end of stage I, the comparisons of incidents with incidents would slowly be replaced with comparisons of incidents to properties of the category. At this stage, diverse properties and categories also begin integrating with each other forming a coherent and logically grounded theory. Thirdly, "delimiting the theory". Delimiting features are imposed on two levels, theory and categories, to avoid overwhelming the researcher. Theory is delimited by applying "reduction" where underlying uniformities are generalised, conceptualised and abstracted into higher level concepts. As the theory becomes more solidified and the researcher's commitment to it increases, there is space and opportunity to reduce the number of categories and apply this smaller set of categories to the rest of the data for a more focused analysis. Fourthly, "writing the theory" (Strauss and Corbin, 1990). Here, the coded data and the theory are combined to form a systematic, substantive theory.

3.3.3. Endogenous Topic Analysis

"[The] immortal ordinary society … is only discoverable. It is not imaginable. It cannot be imagined but is only actually found out, and just in any actual case. The way it is done is everything it can consist of and imagined descriptions cannot capture this detail." (Garfinkel, 1996)

Garfinkel describes the importance of 'discovering' activities, objects and connections deeply rooted in the everyday mundane rituals of individuals. Many of the currently popular sociological qualitative data analysis methods, including thematic analysis and grounded theory has the tendency to overlook these intricacies by considering social settings as abstract and generic rather than being sensitive to

the particular distinctiveness and specificity of the setting (Randall and Rouncefield, 2012).

Endogenous topic analysis is an ethnomethodologically inspired alternative to the presently popular qualitative data analysis methods with the view of alleviating these concerns, to help explicate everyday life efficiently. It aims to explicate rather than explain, to remove subjectivity from the scheme by ensuring that the concepts, theories and phenomena generated are analysed and developed in a data-driven manner rather than reduced, abstracted and transformed to fit previously generated theoretical frameworks or audience expectations. Harvey Sacks (Sacks, 1992) also has explained the importance of explicating everyday life where he, instead of reducing, abstracting and transforming, unpacks the everyday 'real world' through close observation of its own organisation in the data.

"I want to encourage the sense that interesting aspects of the world, that are as yet unknown, are accessible to observation" (Sacks, 1992).

This ethnomethodologically inspired method, stresses upon following the phenomenon as opposed to identifying concepts that are interesting to the researcher. Thus, all topics talked about are valued and included in the analysis, irrespective of the frequency of discussion or its adherence to theoretical models. Thus, data that otherwise would have been discarded as outliers, distractions or noise would be included in the results, adding to the richness of the analysis, helping reflect the mundaneness of everyday life.

"No inquiries can be excluded no matter where or when they occur, no matter how vast or how trivial their scope, organization, cost, duration, consequences" (Garfinkel, 1996).

Here, there is significant deviation from thematic analysis where the focus is on finding similarities in the data through focusing on the frequency of occurrence (Lazar, Feng and Hochheiser, 2017) of a reasoning to result in a theme. In order to preserve the subtle nuance that holds the analysis close to the data, ETA choses to adopt the documentary method of interpretation which in Mannheim's words is a "search for an identical homologous pattern of meaning underlying a variety of totally different realizations of that meaning." (Manheim, 1952) Here, an actual appearance is considered "as the document of", "as pointing to" or as "standing on behalf of" an underlying pattern (Garfinkel, 1996). The generic nature of the documentary method has resulted in it inspiring a number of data analysis techniques, including thematic analysis and grounded theory. But while these trade upon this method to reduce and abstract, ETA looks at it from an ethnomethodological perspective by disposing practices of abstraction and reduction and utilizing the documentary method's capability to iteratively build up topics through continuous 'reinterpretation' of the topic based on 'individual documentary

evidences' gained from the data (Benson and Hughes, 1983). Thus, irrespective of whether a topic or an aspect of it is mentioned once or more, in an obvious or subtle manner, it is given due value and considered in the rest of the analysis.

This consideration is particularly important when dealing with qualitative data where respondents from varying backgrounds and fields with varying perspectives introduce data variations that is deeply embedded in highly contextualised practical reasonings. This data might often seem variant or discordant on a superficial level while actually embodying rich subtleties that call to be uncovered. Discarding this data as digressive because of their apparent incompatibility to the rest of a theme or theory due to lack of sufficient 'evidence' would be erroneous removal of quality data that represents the true nature of the 'real world'.

To overcome these shortcomings, the focus of analysis in ETA is on respecting the participants' perspective of the experience and the unique practical reasonings they bring forth. Everyday activities are often demarcated as obvious and unremarkable and hence often not given the attention they require. ETA advocates for overcoming this limitation by considering context, settings, activities and motivations for the topics discussed, as drivers for the analysis thereby allowing the process to stay close to the data, rather than use them as claims for reducing the data. Thus, rather than the researcher's subjective orientations being superimposed on the data and its results, the participants' accounts become objects of enquiry and reasoning. This priority of the process stems from the phenomenological approach of bracketing, which adopts a 'natural view' of the world. Here, the aim is to study phenomenon on the fly, without resorting to any external resources, preconceived notions or evaluations about some activity, outside of the situation itself, so as to be able to study it in its own terms, as it is actually accomplished (Given, 2008).

The result of this analysis would have enhanced nuance that has real world coherence which highlights connections, processes and understandings which could be easily filtered out by formal frameworks and models. Thus, rather than settling for a set of disconnected topics which might have academic standing but reflects the comprehensiveness of the real world poorly, ETA brings forward the subtleties and multidimensionality of the data in hand.

3.4. Methodology Configurations

This thesis reports three studies, each with its own unique methodological configuration occasioned by the need to engage with different stakeholder perspectives. This section presents, details and justifies these configurations to explain why and how the methods come together to build each study.

3.4.1. Study I

Study I explores user attitudes towards the shift in novel media to personal data leverage through scenario based design of a future data driven Electronic Programme Guide. The developed scenarios are presented to the users gathered in focus groups, through mockups of the EPG. Post presentation of the scenario, discussions are enabled within the groups. The data from these conversations are then analysed using thematic analysis and grounded theory.

Given the emergent nature of this domain, particularly during the course of this study in 2015/ 2016, the concept of using personal data to drive media services was not one that was popular or acclimatised to the users. Hence, in order to understand their viewpoints in this regard, there was a need to envision the future while concretizing the technology in a way that people could experience it without much difficulty.

With this view in mind, **scenario-based exploration** was chosen as the approach for this study. It is a method that allows elaboration of the conceptual model of a future system, thus aligning perfectly with the goal of the research. It helps describe "the proposed system in terms of a set of integrated ideas and concepts about what it should do, behave and look like, that will be understandable by the users" (Rogers, Sharp and Preece, 2002).

As explained previously (Section 3.2.1.), scenarios present several possibilities simultaneously. Firstly, by definition, "a scenario is a concrete design proposal that a designer can evaluate and develop." (Carroll, 1995) Secondly, they help maintain focus on "future situations of use [and] they help keep designers focused on what matters most" (Carroll, 1995) which are "the needs and concerns of the people who will use the system being designed." (Carroll, 1995) Thirdly, the characteristics of the scenario (setting, actor, goal and plot) help understand user social context in which these future technologies are embedded to move beyond just considerations of technology development to study the socio-technical implications, challenges and priorities of a technology without the immediate need to engage in the deeper systems development requirements for building them.

Because of these varying possibilities presented by scenarios, they allow the users the opportunity to experience a future technology in their unique everyday social contexts while simultaneously providing the researcher with solid grounding to develop a design and research agenda that evaluates the consequences of user engagement with future technologies within specific contexts or domains.

Thus, given the onus that scenario based exploration places on contextualising future technologies for users and, prioritising and understanding user viewpoints, particularly associated with the

everyday socio-technical implications of the service, it proved to be an efficient method to present a near-future data-driven technology to the users.

The study employs **focus groups** to elicit user viewpoints on the use of personal data in media experiences. This goal of understanding the socio-technical implications of using personal data to drive future media demanded a less structured research design that allowed the users to express and elaborate their everyday concerns and reasonings. Given its' flexible structure and the sensitive yet not blatantly intrusive or controversial nature of the topic, a group discussion was seen as an effective method to encourage the respondents to voice their opinions, probe the other participants and provide rich, insightful data that represents different user perspectives in this scenario.

Another reason for selecting user focus groups was to promote dialogue among users who were yet to be acclimatised to data driven media in their everyday mundane lives. Hence, the respondents were encouraged by dialogue within the room to support or oppose the views expressed by their fellow participants, an advantage posed by the very inherent social nature of focus groups. Focus groups also offered a logistically faster method of surveying this emergent scenario while effectively underpinning the challenges faced by the turn towards personal data in novel media experiences.

While interviews also presented the opportunity to collect user reasonings, the relatively novel and future-oriented changes introduced in the designs would have been more challenging to elicit responses from the respondents in an interview where the interviewer would be seen as someone who knew more about the technology, thus effecting the researcher-respondent dynamic and the resulting data. But when presented to a group of 'peers' who are introduced to the technology at the same time, the conversations and the probing is more natural, easier and comfortable for all parties involved, while simultaneously resulting in rich data from active discussion, agreements and disagreements among the respondents.

The data analysis of the study uses **Grounded Theory** wherein **Thematic Analysis** is used in the first stage to enable coding of the data into themes/ categories. These traditionally accepted and popularly used methods proved to be a fail-safe mechanism to surface the underlying themes in the data thus helping highlight and present explicitly, user viewpoints on the challenges of using personal data in novel media experiences.

3.4.2. Study II

The second study looks at the service provider viewpoints in this scenario, using interviews with media service providers serving in varied capacities and working on diverse types of media services

within the BBC. The data collected from the interviews are analysed using the ethnomethodologically inspired method of Endogenous Topic Analysis.

Interviews are used here to study and understand service provider viewpoints regarding the turn towards user personal data in novel media experiences. Here, the goal was to engage with service providers from diverse work teams, serving diverse roles and having varying levels of experience within the organisation, to uncover, unpack and present in detail their responses to map out service provider challenges associated with leveraging user personal data. To accommodate for the diversity in the respondents and to honour the exploratory nature of the research at its formative stage, the interviews were designed to be unstructured, containing a number of seeding topics that were presented to the participants. The conversations were allowed to flow and fit to the particular contexts of each of the participants with the researcher getting involved in the dialogue whenever intervention was required (e.g., a pause, break or digression from the topic). Thus, the focus here was on exploration of the notion of personal data leverage while allowing for grounding in everyday practicalities of the service providers, which was promoted by the use of an unstructured, informant interview format.

Although Focus Groups were another potential alternative for data collection in this study, the diversity of the participants recruited, where they served in vastly different roles and catered to media services whose priorities and audiences varied introduced the possible risk of not finding a common ground and a lack of similarity or overlap in perspective. Since there were no previous studies of a similar nature conducted within the organization, it was not possible at the time to gauge the probability of this risk, which would make conversations difficult and tense, leading to the expression of everyday practicality and nuances being suppressed, thus adversely affecting the data (Krueger, 2014). Here, interviews offered the elimination of this risk, while allowing the researcher the time and space to engage with each participant in an individual manner, providing the opportunity to follow-up and unpack concepts to procure data that is deep and rich, resulting in a clear mapping of the space.

Reflecting Sacks' analytic commitment (Sacks, 1984, 1992) we believe 'that the warrant for any given assertion should be visible in the data' (Randall, Marr and Rouncefield, 2001). Hence, the analysis of transcribed data here was done in a qualitative data driven manner.

Owing to the open ended and unstructured nature of the conversations and the varying contexts addressed by the participants, the topics of discussion and the data formed a rich spectrum of perspectives, often rooted in the practicalities of the real world. Therefore, we adopted **endogenous topic analysis** (Randall, Marr and Rouncefield, 2001) through close reading of the transcripts to identify discrete topics that

had manifested within the participants talk that demonstrate how they reason about personal data. Here, the onus is not on the questions, rather on the data itself as it churns out the topics and unpacks the objects, members and connections, it has embodied within it.

This ethnomethodologically inspired method (Benson and Hughes, 1983) also helped accommodate the heterogeneity in the topics discussed, due to the diversity in participants' roles associated with personal data. It ensured that the very unique practical reasonings presented in the data were not discarded because of lack of overlap or absence of patterns (cf. traditional thematic analysis (Braun and Clarke, 2006a)). In contrast, here the contextual distinctions help form, support and add to the richness of the emerging topics, helping map out the various service provider viewpoints associated with personal data in a rich and detailed manner. The analysis here is sensitive to the situational variations embodied within the data, which might otherwise be considered outliers or distractions, due to lack of overlap.

The following account is a detailed explanation of the shift from Thematic Analysis and Grounded Theory to choose ETA for the data analysis of this study:

Endogenous Topic Analysis was applied on qualitative data made available from Study II, held within the BBC. The study was an initiative to understand service provider viewpoints of the challenges of using personal data in media experiences.

Participants : Participants for the interview was recruited from various teams within the BBC, ranging from Sports, News, Children's content, Education, Video-on-Demand platforms, Radio to Research and Development, Audience Platform, Marketing and Audiences and Data Management, working in diverse roles including researchers, project leads, software engineers, data analysts, user experience designers etc. The common connection that served in their selection for participation was their interest in the use of personal data.

Study Design : The interviews were unstructured, consisting of a set of topics that helped guide the conversations if they ever stalled. The three areas of interest that were used as seeds to drive the conversations were the current collection and use of personal data, the benefits and risks of using personal data, and the future of personal data in new media.

The conversations flowed where the respondents directed them, making the method "informant" (Lazar, Feng and Hochheiser, 2017) in nature. In instances where topics that were aimed to be discussed were ignored, the researchers seeded them into the conversations through questions, follow up questions, comments or queries. Given the highly complex nature of the topics discussed wherein the attitudes towards personal data were varying and highly contextual

the exploratory approach helped accommodate the complexity of the data.

Challenges posed by the dataset : The data that emerged was deeply rooted in the participants' everyday practical reasonings. While the data did not show explicit overlaps, it did expose a systematic variance in perspectives between the teams, depending upon the services they offered and the audiences they catered to. Thus, while the data did not show shocking contrasts or surprising similarities, it exposed the everyday practicalities faced by teams and individuals who work with user personal data. This information built up to be a spectrum of opinions and priorities that were heavily context dependent but unassumably relevant when studying the adoption and use of personal data use in media organisations.

A trial using thematic analysis : A trial to analyse this data using thematic analysis and grounded theory led to atrophy of a considerable amount of data that did not follow the codes built and the theories that emerged based on patterns and similarities. The slight variations in opinions meant that a lot of information did not 'qualify' to be part of further analysis as per the guidelines provided by the codes. The codes in this case proved to be a constraint and a filter rather than an enabler for the analysis. Due to the concurrent nature of the research and the magnitude of its applicability, the notion of reducing, abstracting and transforming the data to form generalised versions of the results with little relevance to any specific party involved seemed a futile effort.

Shift to Endogenous Topic Analysis : Therefore, ETA was chosen to analyse and present results ensuring they were rooted in everyday examples and reasonings that were relatable to individuals seeking to derive value from the research. So, we resorted to applying endogenous topic analysis where the process was entirely data-driven and everything that was spoken about was interwoven as part of the bigger narrative rather than be discarded on the claim of being a misfit. What resulted was a set of major topics, each with a number of subtopics that meticulously accounted for a range of systematically variant perspectives.

Advantages of using Endogenous Topic Analysis : Here, while a set of different topics emerged from the analysis, one concept, trust, had representation over most of them. Trust was discussed in varying capacities, in varying contexts and with varying priorities and needs. The role of trust in the data was dominant to such an extend that it was considered one of the key results of the study. Here, the emergence of trust as one of the key results of the study despite it not becoming a topic itself is owed to the application of endogenous topic analysis.

Firstly, the diverse and mostly subtle ways in which trust was included in the conversations meant that with other forms of data

analysis it would have either been overlooked as a given at each instance of its mention, with the researcher focusing on the more tangible effects of trust, that were often spoken at great length. Also, the diverse contexts in which trust was mentioned increased the chances of it being filtered out due to lack of importance or ability to form patterns. These effects combined together would have resulted in trust being lost in the layers of reduction-abstraction-transformation of the data analysis, leading to the loss of the primary finding here.

Secondly, the data driven nature of the analysis helped preserve the concept of trust and not let it dissolve into the various privacy implications that were often linked to trust. With privacy being one of the biggest practical issues faced by organisations working with personal data and with research and previous literature dominating focus on it, it is easy for any researcher to get lost in the dialogue about privacy, losing out on the aspects about trust that were briefly spoken about but often formed the driver for conversations on privacy.

This barrier was overcome through the use of endogenous topic analysis by continuously keeping in touch with the data and ensuring that all aspects of it are included in the analysis, without filtration and making sure that the information that the data embodied was accurate, minimizing subjectivity from researcher orientations.

Thus, the use of endogenous topic analysis helped in the successful analysis of a dataset that was systematically variant in nature, by preserving the richness and diversity of the results and enabling 'real world' coherence. It also, through its various stages of analysis ensured a bottom-up emergence of a highly dispersed but significant concept which might have otherwise been reduced-abstracted-transformed to fit previous theories and researcher pre-conceptions or even completely filtered out.

3.4.3. Study III

This study used **design fiction** to create the Living Room of the Future[1], which was a room filled with smart objects and where the user interaction data within the room was used to adapt a movie and the holistic media experience within the room in real time. The room, which collected user personal data to adapt the experience also presented several active responses to the data collection process, by bridging the learnings of the previous studies. The experience was used to gather feedback on these initiatives and to further probe user thoughts on the use of personal data in novel media experiences through a tangible experience that immersed the audiences in a potential data driven future. The participants were served with a pre-experience questionnaire. The data from these questionnaires were used to understand the mix of demographics of the audiences and their motivations around media and technology. This step seemed

necessary as the experience was run ‘in-the-wild’ in a public space and no specific screening was imposed on the participation as opposed to the participants being recruited through pre-screening processes in a lab based study. The primary source of data in this study which involved both the users and service providers was exit **interviews** and this data was analysed using **Endogenous Topic Analysis** (used for the same reasons as described in Section 3.4.2).

The World Building approach of Design Fiction was adopted in the making of the Living Room of the Future in this study. Given the emergent nature of personal data use in media, there was the call for an experience that delivered this future reality to the people in a tangible manner. Through Design Fiction, this version of the world was designed, built and delivered to the audiences to experience, so that they were provoked to engage in active thought and dialogue around this scenario.

When building a world, the importance of clearly communicating the change to the participants is talked about by Coulton (Coulton, Lindley and Cooper, 2018). Thus, “when building a fictional world, everything stays the same unless you indicate that it has changed. If it is important that you change something, then you need to communicate it to people.” This premise fit in perfectly with the third study which aimed to bridge the first two studies by drawing out inspiration from the learnings of the first two studies to build responses that explicitly highlighted the use of personal data in the experience, to seek both user and organisational response to such initiatives. Such responses, which strayed away from current popular methods of data management sought the creation of an alternate reality for two reasons. One, to present the impact of data collection and use to an effective degree called for the creation of an alternative-verse that was situated in ‘disruptive’ technologies like the IoT, which transformed everyday mundane objects into data collection points. World Building using IoT, while giving a flavour for living in a future that is IoT-driven (Coulton, Lindley and Cooper, 2018), also makes the notion of data use by technology more prominent and thus worth active consideration by the users. Secondly, to leverage this impact created in the minds of the participants, by simultaneously presenting alternative responses aimed at helping alleviate user concerns around data use by making them a central part of this future world that is created. Thus, this world would help situate the users in this alternate reality that explicitly points out the data collection to provide responses that are meant to alleviate user concerns. This immersion into a tangible, alternate, possibly-future world is expected to probe their thoughts on living in such a future, eliciting rich responses to the design decisions adopted while also encouraging more thought and dialogue around the world itself.

The data gathering protocol utilized here were **informal interviews** with both the users and the service providers. While the goal of the

research was to elicit feedback, given the emergent nature of the domain, there was the need and scope for exploration of the scene to understand the broader concerns surfaced by a media technology that is data driven. This led to the interview shifting from the usual fully structured/ semi-structured evaluation interview formats that would have focused on certain aspects of the experience to an unstructured, informal one where the participants were given the space to respond to any part of the intervention that spoke most profoundly to them. The conversations then flowed from there to further probe and unpack the respondents' thoughts around the various seeding topics, which were introduced into the conversations by the researcher.

4. Study I

The following chapter details the first study done as part of this PhD which probed **user viewpoints** on the potential use of personal data in near-future media experiences.

The study practicalises this goal by using the example of a personalised Electronic Programme Guide, which tailors media recommendations based on user personal data to explore user attitudes towards personal data driven media services. It employs scenario based exploration enabled by the use of probes to convey the functionalities of data-driven personalised EPGs to facilitate user discussions around its potential use.

The results of the study show that users preferred personalised EPGs over current popular EPGs but expressed a significant lack of trust in the personal data collection that drives personalisation. The results further unpack this finding to reveal user appreciation of the various functionalities afforded by personalisation of media while simultaneously presenting their apprehension regarding the implications of their personal data being collected, particularly focusing on concerns about their privacy, data collection in the context of their homes and the current lack of legibility and control.

4.1. Motivation

A research gap for the exploration and explication of the challenges of using personal data in media experiences was identified by the Literature Review chapter of this thesis (Section 2). As a first step of understanding these challenges from the users' perspective, a study that piloted the idea of a media experience leveraging user personal data was presented to the users to elicit their response.

Here, scenario based exploration was used to present novel versions of an Electronic Programme Guide that was customised to be more engaging and relevant to the users through the use of their personal data. These scenarios were presented to focus groups of users who discussed the use of personal data within media experiences using these scenarios as probes. The rest of this chapter reports the design, practicalities and findings of this study.

4.2. Electronic Programme Guides

The PhD research being partially supported by the BBC, the study was done in close collaboration with the organisation. At the time of the study, the Research and Development sector of the organisation had just begun looking into alternatives for iPlayer functionalities that could be made more relevant and supportive of the user's lifestyle through utilisation of user personal data.

Here, one very traditional media service that still followed archaic norms was the iPlayer's Electronic Programme Guide. An Electronic Programme Guide or EPG is the electronic version of the conventional TV listings magazine. It is available on the internet or on a TV service like Freeview, satellite or cable. Its purpose is to give the user a list of programmes available at the present moment and later. Most modern EPGs also allow reminders to be set to trigger the beginning of chosen TV shows and also record shows if necessary. "Like a TV listings magazine they also provide extra information about each programme like plot summaries and cast lists." (BBC, 2014)

While the rest of media technology was moving from traditional broadcasting to narrowcasting highly curated and personalised content, there was a distinct call to update current, popular EPGs from the traditional model of providing generic lists of TV schedules to a contemporary one that is more engaging, relevant and reflective of the modern user's media consumption trends.

With the EPG being personalized, it would no longer be a simple list of programmes aligned against a timeline, as chosen by the broadcaster. It would be an intelligent mix of both live broadcast and video on demand (Inclusion of on-demand services in EPGs was part of the Amendment to Communications Act 2003 (DCMS, 2013)), laid out against a timeline, featuring content that would be of interest to the user, reflecting the user's lifestyle and TV viewing habits.

Recent research on EPGs has shown interest in the technology of systems which recommend TV broadcast through adapting the EPG, for example, Personalised Television Listings (Smyth and Cotter, 2000), TV Predictor (Krauss, George and Arbanowski, 2013) and the work by Haesung Lee and Joonhee Kwon that defines an application architecture for "personalized TV content recommender service[s]" (Lee and Kwon, 2013).

Research has also looked into design architectures for EPGs to enable users with the most efficient and easy interfaces for EPGs, like the work by Bonnicci (Bonnici, 2003) which outlines "a design model for an EPG interface [which] has been developed as a framework for designers to understand and produce EPG's […which…] take into consideration viewer interaction and interface issues" (Bonnici, 2003). Another closely related piece of research is the work by Cotter and Smyth (Smyth and Cotter, 2000) which details the system specifics of an adaptive EPG on a very technical level.

While there has been interest in alternative ways of personalising and presenting information on EPGs most of this research is yet to be transferred to large scale industry systems, that users engage with on a daily basis. This could be mostly because this research often focuses only on the technological specifications of such systems and thus

often stop short of understanding the socio-technical implications of introducing such changes to the service. Hence, while the technical considerations of the work is comprehensive and detailed, it fails to address the impact of the technology in a user-centered manner for an activity which depends primarily on the user's preferences.

Study I of this thesis exploited this very gap by studying the socio-technical implications of personalising the EPG. Here, in order to address the requirements of the broader research agenda of this thesis, the iPlayer EPG was used as a vehicle for demonstrating the innovation produced by personal data in a media experience while provoking user response to elicit the challenges presented by such a shift.

4.3. Design of the Personalised EPG

Design of the personalized EPG began with the study of previous work on EPG usability. These have shown that users "found it overwhelming and disliked [....] confusing ordering of channels" (Jones, 2000) and that the "users have fundamental difficulties navigating in this way with the first generation of UK services" (Jones, 2000). To overcome these navigation fatigue issues, our design aims towards minimum navigation by optimising screen space through replacement of content of disinterest with content of interest, enabled by filtering and customisation.

This design decision is further supported by the findings of the study by Taylor and Harper (Taylor and Harper, 2002) where they state that it is recommended to reduce "the cognitive demands associated with using an EPG" (Taylor and Harper, 2002) by "simplifying the decision making process" (Taylor and Harper, 2002). One of their recommended methods for this was "by limiting the number of channels", which, again, our proposed EPG does in a less constraining manner by filtering out content that is not of interest to the user.

There has also been some research on user preferences but it has been around understanding and calculating user decisions surrounding content preference. An example of this is the work by Sullivan, Smith and Wilson(O'Sullivan, Smyth and Wilson, 2003) where they look at "the type of implicit indicators that can be identified within the DTV domain and the extent to which they can accurately reflect a user's true preferences" (O'Sullivan, Smyth and Wilson, 2003).

This study is quite different from most of such previous research discussed here wherein the focus is not on the study of usability of a personalised EPG. But it leverages the current lack of innovation in this scheme to present the innovation possible by the introduction of personal data into an otherwise mundane everyday media experience in a stark manner to probe user thought and dialogue.

Thus, the design of the personalised EPG in this study builds on these previous EPG usability studies to ensure the new version of the EPG will overcome the drawbacks of the previous versions to deliver a more productive and enjoyable experience to the user, as a result of personal data use. While accuracy of personalisation and relevance of the recommended content to the user is of importance when tailoring the EPG, the specifications of the same is out of scope for this study and hence the assumption that all recommendations made are ones that are reflective and supportive of the user's life and media consumption habits is made to eliminate that variable.

Once inspiration from previous work was considered, the next step was to draw out the deficiencies of the current popular EPG that could potentially be overcome through personalisation.

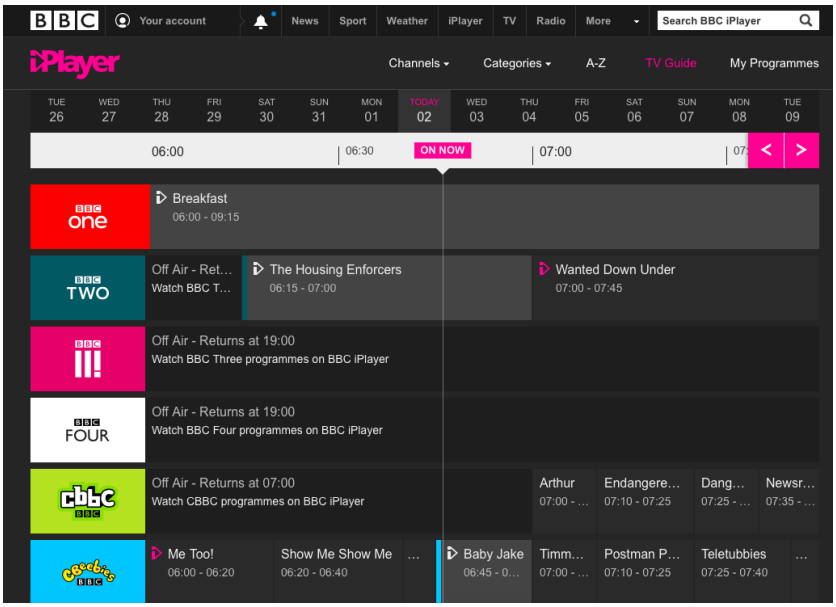


Figure 1. A Screenshot of the BBC iPlayer EPG.

This image is a screenshot of the current BBC iPlayer EPG. It shows that certain channels are not airing content at the present moment. Also, if there are no children in the household or if none of the adults are interested in children's content, the relevance of the rows showing CBBC and CeeBeeBees shows is very low. Thus, the screenshot displays two major flaws, one of screen real estate wastage and the second, the demand on the user to scroll and search for content that is relevant to them, leading to time wastage, user fatigue and frustration.

Owing to the obvious design improvements presented by this current EPG and deriving inspiration from research that studied TV recommendations (Smyth and Cotter, 2000; Krauss, George and Arbanowski, 2013; Lee and Kwon, 2013), shifting from this generic list of all-TV-programmes to a personalised screen space that presents

the user with optimised opportunities for media consumption was chosen as a design decision to be applied to the personalized EPG.

4.3.1. Four EPG versions

Four versions of personalized EPGs were designed and presented to users as part of this study. The rationale and design of each is detailed in the following account.

4.3.1.1. The Recommender

Deriving insights from the previous useability work around EPGs and the flaws highlighted in the current iPlayer EPG, the first mockup of a fully automated, Recommender EPG was produced.

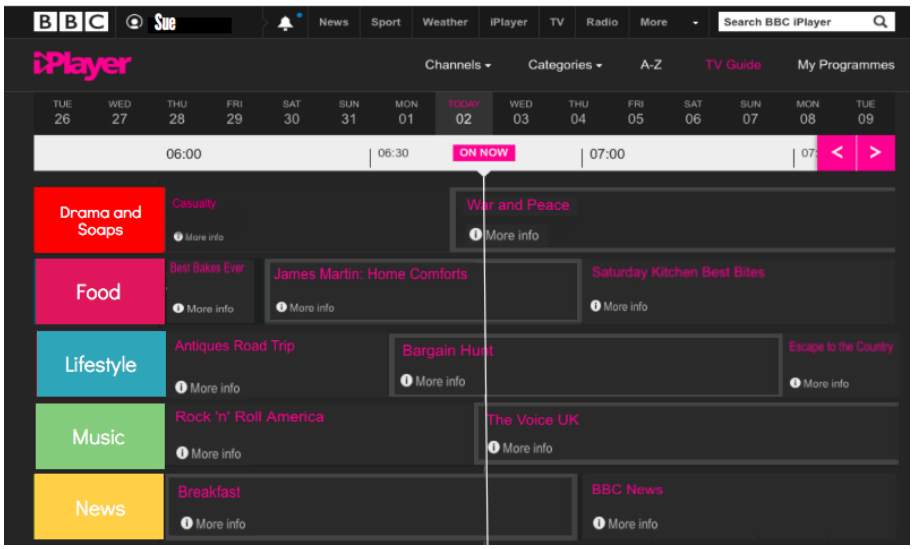


Figure 2. The Recommender EPG.

This EPG was always intelligently fully-filled with live, broadcast and video-on demand media options that reflect the user's actual media consumption habits. The concept of channels are replaced by categories/ themes. Here, programmes from different channels are brought together to cluster within these themes that are of interest to the user, irrespective of whether they are being aired in the moment or have been aired previously. Content that is not of interest to the viewers are filtered out and the functionality of being accessed from and used on multiple devices is provided, making the application responsive. Also, the big shift of going from a universally delivered generic website to a personalised experience, where the user has a profile and is logged into the system, with data collection and processing happening in the back-end is initiated here.

4.3.1.2. Calendar Metaphor

The second version of the EPG proposed was one that was quite different to the recommender version where a fully filled EPG was frontloaded and delivered to the users. This design, is that of an EPG which follows a calendar metaphor. It would start out empty like any schedule and the user could pull in content of their choice to populate it. It would also be grouped by categories, in this case, the user gets to create and customize their own categories. This includes content from different channels and it would be available on multiple devices. This design, whose idea was initiated by BBC R&D, was realized by the inclusion of two EPG interfaces : one that showed an empty EPG, followed by a populated EPG, which shows an example of how the EPG interface would look like once it is 'used' by a viewer.

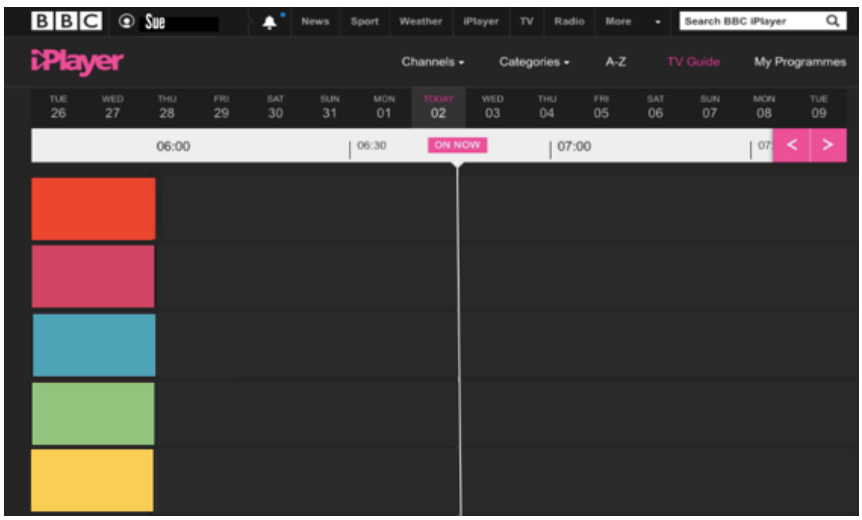


Figure 3. The Calendar EPG (Yet to be Used).

Other extra functionalities were also made available to this version. For example, social groups which work like private groups on Facebook or WhatsApp, where users share information between members of the group. Previous studies have shown that 77% of users find new programmes through recommendations from friends (Lull, 1980), hence the EPG allows users to form groups through their EPG profiles and send show recommendations to the group, which would populate the EPG of the members of the group. Also included were variable length videos where even if the time slot available is shorter than the length of a show, it could still be accommodated intelligently into the slot by the user just dragging and dropping it into the slot in the EPG. The video will resize itself to the new time duration. If there is a programme the user does not want to miss, like their favourite sports team playing live, it could be assigned highest priority and an alert would pop up irrespective of what is scheduled on the EPG. And

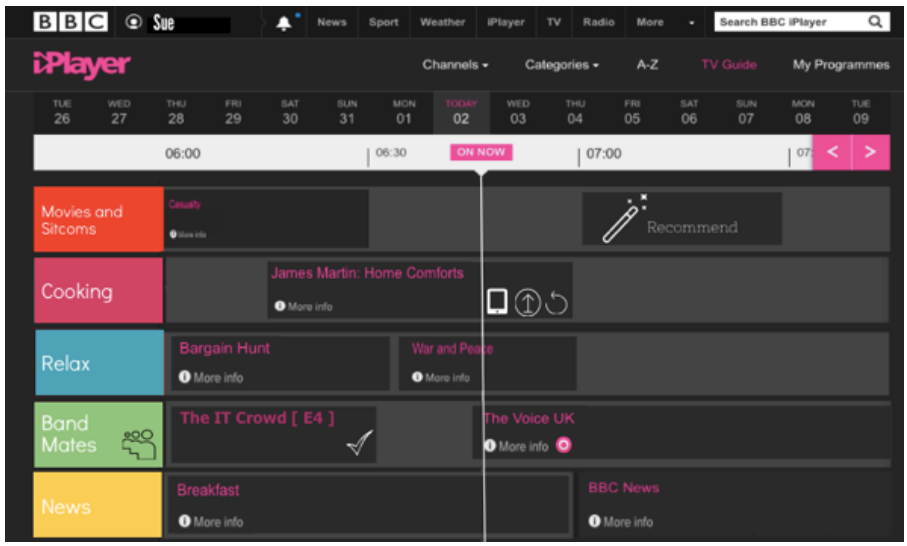


Figure 4. The Calendar EPG (Filled in by the User).

if the user does not want to actively choose something for a particular time slot, they could select that slot and ask the system to recommend.

These various functionalities are represented in Figure 4. using visual cues, later communicated explicitly by the researcher to the research participants during it's presentation.

4.3.1.3. Accomodating for social scenarios

This version of the EPG takes into account the very mundane everyday social nature of TV viewing, where user's tend to watch TV with others (friends, family, colleagues etc). It is assumed that the type of content they watch with these different groups would also be different.

In order to accommodate this situation, the EPG would give collaborated profiles where all the content watched with a specific group or person will be available in a group profile that is associated with that group or person. To demonstrate this case with clarity, two very different use cases were chosen and implemented as mockups.

Figure 5. is an example of two university students who live together. When together, they tend to watch music videos, movies, series, sitcoms and sports. Hence their profile reflects categories related to that kind of lifestyle.

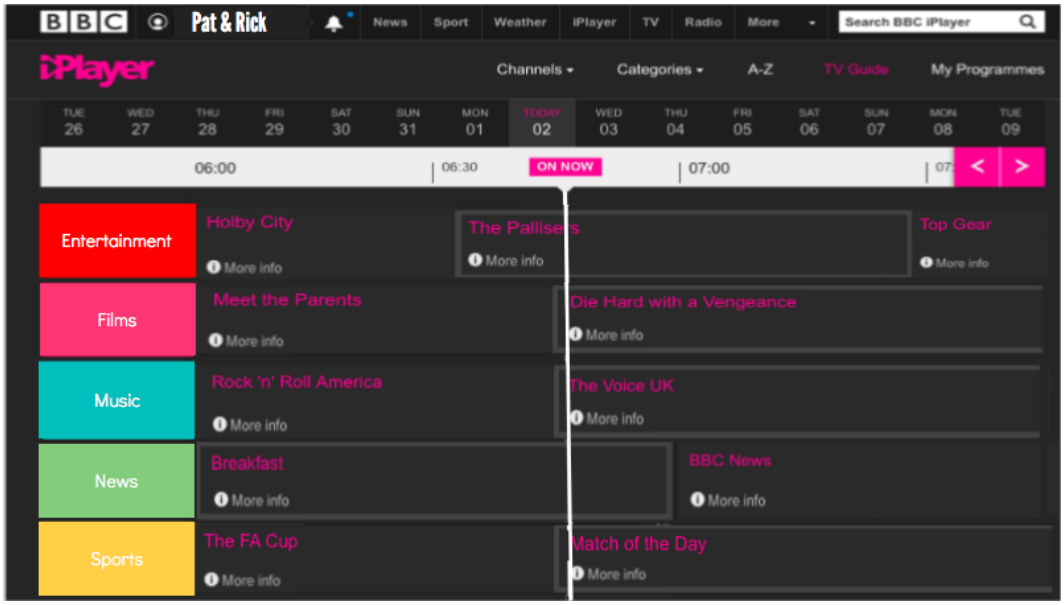


Figure 5. The Collaborative EPG for two University Housemates.

Figure 6. shows the profile of a father and two children who often watch TV together. Their recommended content is vastly different, showcasing animation movies, cartoons and nature shows, the kind of media they collaboratively consume.

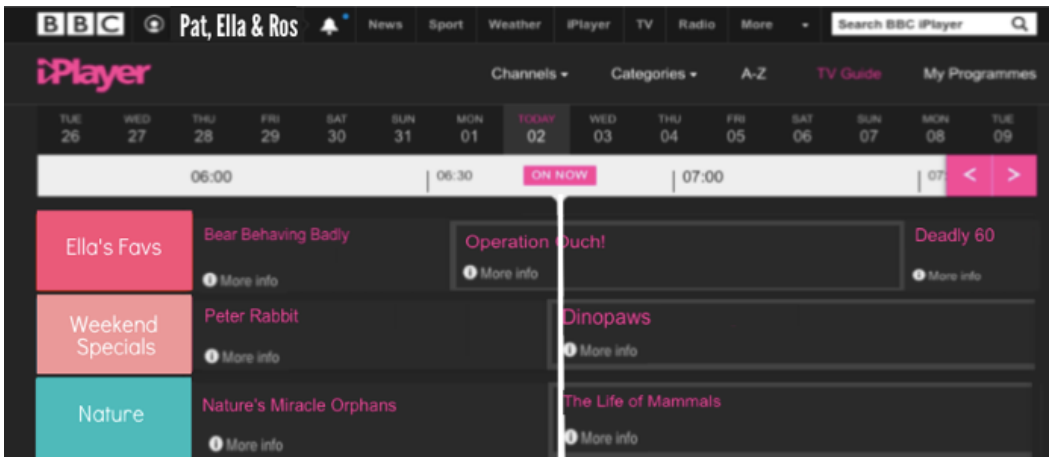


Figure 6. The Collaborative EPG Case for a father and two children.

4.3.1.4. Introducing IoT for provocation

The three previous designs help communicate the innovation provided by personal data use in media experiences. This is done by presenting these possibilities by making significant changes to an everyday media experience and grounding its consumption projection in current mundane media consumption scenarios. All the while, it is simultaneously communicated to the users that these changes are

possible through the collection of user personal data, either through interface affordances like the presence of user profiles or by the researcher explicitly pointing this out while describing the designs presented.

With the fourth design, the Internet of Things is added into the equation, making the data collection involve everyday mundane objects within the home and showing their capability to collect personal data on a granular and ubiquitous manner to power the media experience presented.

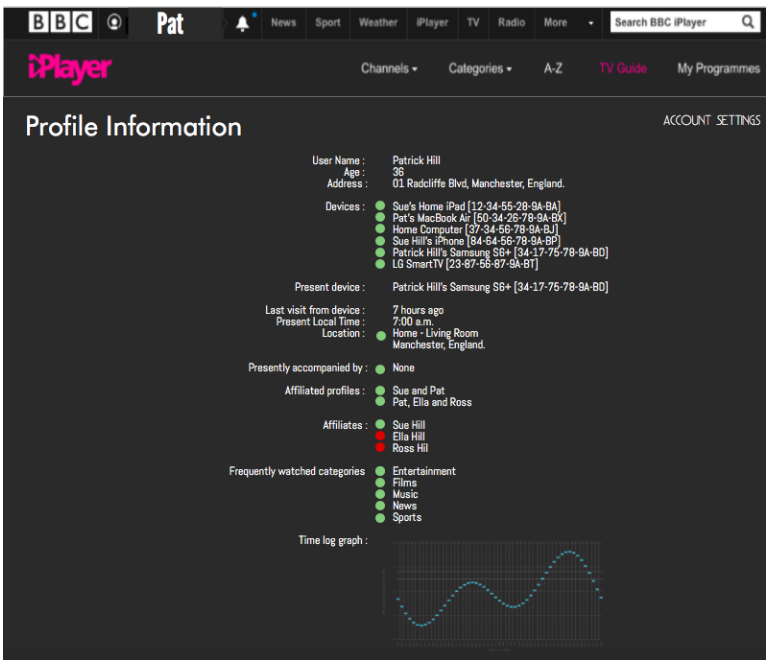


Figure 7. Image showing granular IoT Data Collection.

This addition is communicated to the users through a mockup that shows the different bits of data that could be collected about a single user and others associated with them. It also shows the possibility of forming patterns from this data using a simple visualisation of consumption patterns along with introducing the concept of granular control over the data collected as an option to be engaged with by the users.

Such representation of the granular collection of personal data and the inferences derived from them for a media experience was in line with Bernhaupt et al.'s (Bernhaupt *et al.*, 2007) work where the challenges associated with the use of personal data in a media experience was not considered by the users because the involvement and implications of data use was not apparent to them. Thus, an explicit visualization of the data collection was included to overcome this barrier to communicate to and provoke the users' thought and dialogue around media experiences leveraging personal data and elicit the challenges associated with this shift.

4.4. Approach

Given the emergent nature of this domain, particularly during the course of this study in 2015/ 2016, the concept of using personal data to drive media services was not one that was popular or acclamatised to many users. Hence, in order to understand their viewpoints in this regard, there was a need to envision the future while concretising the technology in a way that people could experience it without much difficulty.

With this view in mind, scenario-based exploration was chosen as the approach for this study. It is a method that allows elaboration of the conceptual model of a future system, thus aligning perfectly with the goal of the research. It helps describe "the proposed system in terms of a set of integrated ideas and concepts about what it should do, behave and look like, that will be understandable by the users" (Rogers, Sharp and Preece, 2002).

While the use of scenarios is nebulous enough to give the users space to experience a future technology while rooted in their unique everyday social contexts, it simultaneously provides the researcher with solid grounding to develop a design and research agenda that evaluates the consequences of user engagement with future technologies within specific contexts or domains. Firstly, because by definition, "a scenario is [also] a concrete design proposal that a designer can evaluate and develop." (Carroll, 1995) Secondly, because they help maintain focus on "future situations of use [and] they help keep designers focused on what matters most" (Carroll, 1995) which are "the needs and concerns of the people who will use the system being designed." (Carroll, 1995).

Thus, given the onus that scenario based exploration places on contextualising future technologies for users and, prioritising and understanding user viewpoints, particularly associated with the everyday socio-technical implications of the service, it proved to be an efficient method to present a near-future data-driven technology to the users.

In this study, the scenarios were used as probes that encouraged user thought and dialogue about the use of data-driven future media experiences. They were not fashioned against the models of cultural, technology or informational probes discussed by Gaver (Gaver, Dunne and Pacenti, 1999), Hutchinson (Hutchinson et al., 2003) or Crabtree (Crabtree et al., 2003), instead they were put to practice in the most mundane sense of the word wherein they were intended to 'poke' the users' and open up the space to enable deeper and broader thought on data driven media technologies, through the possibilities presented by the scenarios.

The scenarios were presented using mockups of the EPG interface (Section 4.3) . These mockups were lightweight with respect to their design and development cost and time, but their level of detail and familiarity with current, similar technologies made them efficient probes that helped efficiently present and explore the scenarios presented.

The scenarios were presented to focus groups and discussions were enabled as part of the data gathering protocol of the study. The data was audio recorded, transcribed and analysed qualitatively using thematic analysis and grounded theory.

A brief summary of the four scenarios that resulted follows :

Scenario 1, Recommender : This case delivers an automated fully filled EPG based on the users' TV viewing history. All content that is not watched would be intelligently filtered out, channel wise grouping of content (in the current EPG) would be replaced with category wise grouping, where content from different channels would be grouped to fill up the various categories.

Scenario 2, Calendar metaphor : This EPG follows a calendar metaphor and is built by the user by pulling in content of their choice. It would also be grouped by categories (which the user creates and customizes). This version also introduces extra functionalities like social groups, where users can form groups with other remotely located users and share show recommendations between members of the group, use of variable length videos, setting priorities, alerts etc.

Scenario 3, Collaborative Profiles : This case accounts for the social nature of TV viewing. It is assumed that content users watch with different social groups would be noticeably different. In order to accommodate this situation, the EPG would give collaborated profiles where content watched with a specific group or person will be available in the profile associated with that group.

Scenario 4, Inclusion of Internet of Things : The introduction of Internet of Things would include devices in the home automatically profiling users to produce accurate media recommendations, adapted to who is watching, at what time, in which room of the house, during what season of the year, etc. The case is demonstrated with the help of a sample data profile that would be generated as part of the data collection process, designed to provoke user thought and dialogue around media experiences leveraging their personal data.

Thus, four settings were presented to the participants through the design of four contexts, represented through EPG interface mockups, where the EPG was improvised using personal data. The users here were the actors of the scenario who would engage with these future technologies in their everyday lives through the four contextual settings provided. The plot of the scenario, while embodied within

the very design of each mockup through the functionalities it presents, is built up explicitly by presenting each setting to the actors and then detailing the various interactions and capabilities presented by each setting, which would lead to the actors achieving their goal of finding a piece of media content that is relevant and of interest to them, thus helping them visualise their use of the presented future technology in their everyday lives.

These scenarios were developed with the view of helping the users experience and understand what the future holds for data driven media and helping ground them in this near-future reality for a brief span of time. **This is ensured by the use of designs that draw upon current, popular media services (like the iPlayer EPG), that are familiar to the average users and extending this familiarity to the newly designed EPG mockups, which helps users extrapolate their current, everyday media consumption behaviours to the near future, when these services shift to increasingly modern functionalities that are user data driven.** This extrapolation would help the users easily relate the forecasted future to their current everyday media consumption reality, (drawing upon similarities and differences) and thus, express their viewpoints associated with this shift that makes some very fundamental changes in the way media is served to them and consumed by them in the near future.

4.5. Participant Recruitment

Ethics for the study was cleared through the Ethics Committee based in the Computer Science Department at the University of Nottingham. Consent forms were given to all participants before the study, which were to be duly filled in and signed for anyone to be able to take part in the study.

Four focus groups were held spread across Manchester and Nottingham, with a total of twenty participants (five per group). The recruitment was done using recruitment emails sent out through mailing lists both through the BBC and the University of Nottingham, asking for users to involve in a session about a new EPG technology.

All participants were above the age of 18, from varying socio-economic backgrounds and all regular media consumers. Most of them were acquainted with the basic concepts of modern broadcast and media, with 12 of them using an EPG and 17 of them using Video on Demand either daily, weekly or monthly.

4.6. Data Gathering

The scenarios were presented to groups of five (following Kreuger's suggestion of an optimum number to promote effective dialogue in a focus group (Krueger, 2014)) by the researcher, where the role of the researcher was to only introduce the technology by presenting the

mockups and describing the functionalities explained in Section 4.3. The floor was then left open to the group for discussion about the technology, its use and implications in their lives, with researcher involvement happening only if requested by the participants (eg., a doubt regarding the functionality). Here, the design of the mockups (detailed in Section 4.3.) were carefully constructed such that they "act[ed] as design intervention[s] that elicit inspirational material while avoiding the understood social roles of researchers and researched" (Gaver, Dunne and Pacenti, 1999). Each mockup incorporated elements (like the various functionalities introduced, the two different methods of tailoring EPGs, the example of collaborative profiles that include children's data, use of the very granular data profile etc.) that served as talking points, probing, encouraging and sustaining conversations that explored the use of data driven media technologies and it's implications in various social contexts. Thus, the participants were moved closer to the design space, empowering them to be able to discuss, agree and disagree among themselves, thereby reasoning out and validating (Hemmings et al., 2002) their viewpoints about future data driven media experiences.

The scenarios and cases were presented to the participants using mockups of the proposed EPG versions followed by explanations of the various functionalities. After each case was described, the participants were encouraged to discuss among themselves what they thought about the proposed technology, what they liked and what they did not like and to express any questions that they had. The discussions took around 30 to 45 minutes. Questionnaires were provided before and after the focus group. The one at the beginning had questions regarding the current TV watching habits of the participants and the one after the discussions asked them to rate the functionalities described and to choose one of the EPG versions : current, recommender or calendar (Both attached in Appendix I).

Given that this study was a preliminary study that aimed to primarily check if users had concerns and passions when considering personal data use by their media experiences and if so, to investigate what factors contributed to these concerns, the goal here was promotion of maximum dialogue through involvement of a varying range of perspectives. With this view in mind, focus groups were chosen as a means for eliciting data. Focus groups, while enabling effective dialogue between a diverse range of participants would help bring up topics of interest while uncovering both the popular and conflicting areas, thus divulging the data and insights required by this study, within the constraints of time.

The data collected from the focus groups were analysed using thematic analysis and grounded theory to reveal the major themes in the conversations. As the research agenda leaned on a primarily qualitative analysis no statistical methods were applied on the data

from the questionnaires. Rather they were used to understand the demographics of the participant population and to provide context to the results, as necessary.

4.7. Findings

The results of this analysis show that the users appreciated the advantages afforded by personalisation of EPGs and suggested better support for easier interactions and social viewing. But they also expressed a genuine lack of trust associated with the exchange of personal data, often stemming from privacy concerns, fear of data collection in the context of the home and a current lack of understanding and control over the underlying data practices.

4.7.1. An improved media experience

The users appreciate the innovation provided by personal data in media experiences, particularly the personalization of media experiences. Here, they also voice their preference for simple and quick interactions and their want for the media to be reflective of their offline social contexts.

4.7.1.1. Appreciation of personalisation

The results show a definite appreciation of the personalised EPG. The users saw personalising the EPG to be "*an improvement on the conventional EPG*" [Zack].

In fact, users wanted the recommendations to go beyond being just a machine output, which was the current practice with systems like Amazon that pushed "*incredibly intrusive and phenomenally annoying*"[Tom] recommendations. Their current experiences with recommendations, were often flagged as unsatisfactory and users urged for the availability of more accurate recommendations. They requested for efficient and effective recommendations which would go beyond recommending just mainstream and popular content like, Top Gear and suggest specialised recommendations that reflect their particular media consumption habits, like "*70s car chase movies*"[Meg].

While users appreciated tailored recommendations they also expected not to be "*constrained*"[Kelly] by the system and its recommendations. They highlighted the possibility of losing the serendipity of "*discover[ing]*"[Paul] new shows when media is purely recommended to them, with little space for exploration. As a response to these concerns, users asked for a mix of automation and user control over their media. Thus, while the users were happy about the "*front-loaded*" [Amelia] recommender, they wanted enough space, freedom and opportunity to exercise some level of control (offered by the calendar version), as stated by Peter here, "*I don't think I will start it from scratch but maybe if you have an interface that you can*

alter, like then change it, I would like that". For the same reasons, the availability of the search option on the interface was often noticed and acknowledged by participants as a necessity to ensure user agency within the system.

4.7.1.2. Quick and simple interactions

Users wanted their interactions with the EPG to be fast and simple and so liked the idea of ease of navigation in the recommender EPG through the elimination of the need to "*scroll all the way to the bottom*"[Zack], searching for something relevant. Instead, "*there is always something to watch [rather than] empty slots*"[Kelly] or content they are not interested in.

Users mentioned that the calendar version "sounds like a lot of work"[Peter] and that they could not "imagine..[]..sitting down with a schedule and sort of saying, I can watch that at 8 p.m. and I can watch that at 9 p.m."[Tristan]. So the idea of planning TV viewing using something like a calendar was considered "a lot of effort […] in setting […] up with very little gain compared to the .[]. recommender system"[Tess]. Owing to these reasons, more than 10 out of the 20 respondents chose the recommender EPG while only 4 preferred the calendar version.

The quote by Peter discussed in the previous section, where they demand agency within the system might be considered a contradiction to this requirement of minimum effort from the users' side. But here, the request for opportunity to exercise control is explicitly prefaced by the requirement for the user to not be expected to "*start from scratch*", but be given options to make changes. This shows how the user would like to be assured of options for control over their media personalisation but simultaneously expects most of this control to be intelligently dealt with by the technology itself, while providing the user with ample space to exercise the expected level of agency through quick and easy control mechanisms, if required.

4.7.1.3. Support for social viewing

The relation between media consumption and the social life of the users was shown to be rich. Users showed genuine interest in aspects of the technology that enabled easier integration of social viewing both in co-located and remote scenarios.

In situations of co-located viewing, users considered the concept of collaborative profiles especially helpful. They thought it would help moderate tensions that arise while deciding what content to watch. Parents thought it would help while watching TV with children. For example, to "*find something that we all want to watch. I don't want to watch My Old Pony I'd like to watch Up or something like that*"[Paul].

Users were particularly interested in the possible overlaps that would configure in shared profiles with various social groups, which would in effect, result in content that everyone would like to watch and also in the filtering out of content that maybe one of them would have already watched, so that they don't have to repeat it. An example of this is the instance where P9 expressed her wish to have five lifestyle categories on her personal EPG while watching TV alone and having a balance of lifestyle and sports programmes when her father or brother (who would have only sports categories on their profiles) watched TV with her.

Co-located viewing using collaborative profiles was also expected to alleviate social tensions that would have arisen in previous scenarios. Indie, who lives in a shared house and does not interact much with her housemates except while watching TV in the common room expressed that it "*not only [provides] the opportunity of recommending the programme that fits everyones' taste but it also […] let[s] other people know this person dislikes these kind of programmes*"[Indie]. Also brought up was the possibility of avoiding uncomfortable social situations that arise from serving inappropriate content, particularly when children or guests were present in the group.

While they were not watching TV with someone, users expressed the need to share what they were watching with friends and family and so they liked the idea of forming private social groups where recommendations could be shared between members of the group. The idea of "*orchestrating shared viewing*"[Key] or as Basil puts it, "*it could be like a WhatsApp group like oh… we are watching this tonight and put it in the thing and we can see if they watched it or not*" was considered useful in enhancing the experience. Users like having the capability to share in private groups versus through established social media as they consider public shares to "*often [be] very irritating for other people.*" (Lull, 1980) and because they "trust my friends more than the system"[Meg] when it comes to the quality of media recommendations.

4.7.2. Mistrust associated with personal data exchange

While the users were positive about the innovation presented to media experiences by the use of personal data, they also identified concerns like privacy issues, lack of trust and issues around data legibility and control, that they saw as challenges to this shift.

4.7.2.1. Privacy issues and fear

Users' acknowledgement of the value presented by personalised media was often overshadowed by their fear associated with personal data exchange. This can be noticed in comments like "*I have got privacy issues, I'm not ready to give out data just to get better recommendations of programmes I watch*"[Tess].

This fear stemmed from the concern of loss of privacy that could result from mismanagement of their personal data. So, users wanted to know more about how the data is distributed and more importantly, to whom it is passed onto through questions like, *"[Data is] collected and then its sent or is it processed locally?…. its like obviously sold to third parties?*"[Paul].

The use of IoT instigated predictions of dystopian scenarios where audience privacy was heavily compromised to the point where everyday media consumption would be under constant surveillance, leading to severe loss of user trust, fear and aversion towards such data driven technologies. "*So its detecting who I am and say I'm watching a lot of new programmes about terrorism, what guarantee do I have that this information isn't going to put me on a watch list and suddenly the camera is watching you all the time, its detecting anger when I'm watching programmes about terrorism. I mean this sort of thing could easily be interpreted in ways that would have impact on greater scrutiny*" [Meg].

4.7.2.2. Lack of trust in the context of the home

The fear and reluctance to share data was often traced to users' lack of confidence in the corporations rather than the technology itself. Users thought "*[The companies] are gonna be selling where I am, what I'm doing, to all these other companies*"[Key].

This is particularly perturbing because media consumption is often associated with the home. And since data is relational (Crabtree and Mortier, 2015) and hence data about one person often involves data about other associated individuals in the home as well, the aforementioned concern of privacy and fear of infringement of the same is exacerbated. Thus, the idea of sharing mundane, yet personal daily rituals one does with others in intimate spaces was not welcomed with positivity. Here, the participant explains the same by stating that they don't want to "*broadcast where I am at every point in my life and who I'm with as well because its not just your room location, if you are with other people then you don't necessarily want that*"[Mark].

Also, media data in the home was repeatedly associated with children which led to a number of emotional comments and even requests as the legal, social and ethical guidelines associated with the use and management of children's personal data is yet to be constructed and communicated to the average users successfully. The uncertainty and lack of understanding in this scenario led to the users expressing concerns like "*my kids, are you recording my kids? [….] I'd be freaking out*"[Paul].

Generally, the idea of being able to record one's media habits, especially in the context of one's home, was received with a lack of trust and discomfort as shown by the following quotes. "*That creeps*

me out"[Westin], "*Thats super creepy.*"[Villa] and "*It gets slightly spooky*"[Westin].

4.7.2.3. Need for better legibility and control

Users attributed their genuine paucity of trust in the present day data economy to the current lack of transparency in data exchange mechanisms. They mostly requested for transparency through dashboards and visualisations so that they "*actually ...[...]... know what is happening*"[Kelly].

Users claimed that providing more transparency and control, particularly when data collection happens in their homes and involves family, friends, children, guests etc, helps improve trust in the organisation and foster a better data exchange scenario helping alleviate the current fear of privacy compromises.

Users pointed out to present day convoluted data practices where they feel a genuine lack of understanding and control and where they are forced to trust organisations rather than be encouraged to organically develop trust. Tristan explains this situation through an example where he believes the only way of exercising any form of control is to "*spend hours setting up the privacy controls in Facebook and in the end, even if you have the best privacy settings on Facebook, you still have to somehow trust Facebook to actually do it*"[Tristan], highlighting how the lowered levels of legibility and control puts out a call for better organisational accountability from the service provider in this scenario.

4.8. Study Conclusions

The results of the study showcase user response to a personal data driven Electronic Programming Guide. These responses form a number of themes that highlight user expectations and potential challenges in this domain that either speak specifically to the future design of data driven EPGs or more broadly towards the future design of data driven media experiences. The rest of this section unpacks these user viewpoints with the aim of informing future design endeavours in data-driven media.

4.8.1. EPG-specific findings

A number of responses to the study often involved user viewpoints on the specific media service provided by a personalised EPG. These findings that talk about the preferred kind of interactions and capabilities expected from the EPG help specifically inform research and design of future data driven EPGs.

4.8.1.1. An improved experience

The results show that users appreciate the concept of personalising the EPG. They feel like this makes the service more relevant as the personalised version presents them with content that is of interest to them, resulting in better user satisfaction and an improved experience when using the EPG.

Also noted was the capability afforded by personalisation to avoid the need to scroll and search the length of the EPG to find something of interest. Thus, making the EPG automatically populated with content of interest, reducing the user need to spend time and effort searching for content is preferred.

This capability of quicker interactions is closely tied to the users' predominant choice of the Recommender version, which serves the user with a system-generated, personalised EPG over the Calendar version, which demanded the users to 'build' their own EPG, further emphasising the need for simple, quick interactions with the EPG.

Thus, the study identifies the user preference of personalising the EPG while highlighting the need for an improved EPG experience through quicker finds and simple interactions that easily serve the users with content that they prefer.

4.8.1.2. EPG to support offline interactions

The results also identify the need for EPGs, especially when used in social settings, to be able to support and align with the audiences' offline social interactions. When data collection is involved, users expect the service to intelligently understand and adapt to the social circumstances, helping relieve social tensions in some instances and helping share valuable content in others.

Our proposal for collaborative profiles and content sharing through social groups for the EPG were acknowledged as an instantiation of such technology mediated support for offline interactions. Possible examples of the same could be through measures like EPG collaboration with the users' calendar, non-intrusive social media involvement, novel methods of sharing content through mobile devices etc.

Future research in both personalised EPGs and broader research in data driven media could explore this domain further, by looking into alternatives that help accommodate better integration of the media experience with the users' offline social lives.

4.8.2. Results for future data driven experiences

The conversations around personalised EPGs often opened up to questions around the use of personal data collection that enables this

kind of technology. This resulted in dialogue around the use of personal data in novel media experiences, particularly helping identify user concerns in this domain.

4.8.2.1. Building privacy and trust

From the results of this study, it can be seen that the primary concerns arising from the use of personal data in media experiences is the concern over user privacy and loss of user trust. Here, privacy concerns are associated with dystopian scenarios of what the collected data could be used for, the need to know more about the data flows and the question of whether the risk of giving away such intimate data is worth the return of getting media recommendations. Lack of trust was associated again to the current lack of legibility in data flows, which combined with the sensitivity of intimate contexts like the home that caters to vulnerable populations like children and the elderly, makes the concept of trusting a service provider more challenging for the users.

Having uncovered the dominant concern around trust and privacy in this scenario, the rest of this discussion unpacks the findings to highlight data legibility, control and accountability as potential practical responses to these seemingly high-level, abstract challenges posed by the use of personal data in media experiences.

4.8.2.2. Improved legibility

The lack of data transparency afforded by organisations that collect and use personal data is a major concern that the study surfaces. The lack of understanding of the data practises is in turn contributing to the lack of trust users show in data driven media services, encouraging them to question the need for data collection and the value of going through that process.

Here, the need for providing users with improved transparency becomes a priority. Thus, when designing data driven media services, in order to preserve user trust, there is a call for provision of mechanisms that provide users with a better understanding of the underlying data practices.

This shift could be through the use of appropriate data visualisations, dashboards that display monthly statements of data use and inferences made, or alerts that inform the user of unexpected data usage patterns. Ensuring choice of the most effective methods that efficiently communicate information to the users for various different use cases, through methods that are appropriate for the experience being served, is an immediate concern that requires response on both holistic and case-specific levels as media consumption becomes increasingly data driven.

4.8.2.3. Importance of data control in the home

The importance of the home being the context for data collection and the concerns regarding the same were discussed in Section 4.7.2.2. Consideration of data collection within the home brought forth fear associated with privacy compromises and a lack of trust in the ethos of the data exchange.

This sensitivity is owed to two reasons. Firstly, the involvement of the people associated with the context of the home. The home does not often consist of a single, uniform, homogenous cohort of users, it is a mix of user groups which might include parents, children, guests, lodgers etc. The use of personal data comes with a different set of implications for each of these cohorts which needs to be addressed.

Secondly, the sensitivity associated with the everyday mundane rituals and activities that happen in the home (both media consumption related or otherwise) that is considered a very personal and protected space. Collecting data in such a space is increasing the vulnerability of the people in the home through the possibility of exposing or exploiting knowledge about these personal habits by sharing it with external entities.

Here, in addition to increased levels of legibility, the users put forth the demand for opportunities to control their data. Possible practical solutions could include smartly designed control panels, settings pages or granular control interfaces etc. It could also be alterable, predefined modes of use for when children or guests are present, which could be initiated with a single click.

Just like in the case of transparency, practical and flexible data control is a space with much opportunity for research where response is required in unique and specific manners that align with the service being offered and the context it is being consumed in, so as to not distract or fatigue the users with the overload of interactions that might seem unnecessary to them.

4.8.2.4. Accountability

The low levels of data transparency and inefficient data control mechanisms has in turn resulted in a lack of accountability. Any party that would consume personal data is accountable to both regulatory bodies and to the user. Legislation often expects systems design to provide computational tools that enable accountability. The Consumer Privacy Bill of Rights (House, 2012) in the US and the GDPR (EU GDPR, 2016) in the EU, all highlight the need for building accountability into systems that collect and use personal data.

This lack of accountability poses service providers the challenge of overcoming it and helping users engage with future data driven media experiences in a sustainable manner. This requires much study,

research and effective response, in order to help alleviate a majority of concerns identified in this study.

A first step in responding to concerns of accountability provided by an organisation is to understand how the organisations and media service providers themselves view this shift towards personal data. This requires the study of how current practices reflect upon these challenges and how service providers expect to overcome these challenges in order to provide sufficient organisational accountability to the users.

With the view of understanding these service provider viewpoints, this thesis now moves onto it's second study which considers service provider viewpoints around the use of personal data by media experiences more closely.

Limitations of Methods Chosen : This study uses scenario based exploration which used projections of potential future scenarios to elicit user responses. While the method helps probe user thought and dialogue around the use of media experiences leveraging user data, there is the possibility of limitations to this projection when compared to response from users experiencing and interacting with a tangible media service that leverages user personal data.

5. Study II

The following chapter details the second study done as part of this PhD which studied **service provider** viewpoints on the use of personal data in media experiences.

This study undertakes this challenge through an interview study of 20 employees within the BBC who are actively involved with the creation and delivery of diverse personal data driven media services to millions of audience members on a daily basis. The study was held during a time when the organisation was shifting its business and technological frameworks to one that leveraged user personal data.

The results of the study identify a need for better interactions in the user-data-service ecosystem where trust and value are to be prioritised and balanced at all times. The major considerations and responses that help maintain this balance are identified as the need for being legally compliant, going beyond just the mandatory by ensuring social accountability and exercising ethical responsibility as an organisation. The study also presents how the technology is seen and used as a solution for overcoming some of these challenges and realising priorities to provide value while preserving trust within the personal data ecosystem.

5.1. Service Provider Viewpoints

Study I reported in this thesis piloted the idea of media experiences leveraging user personal data which identified the challenges confronting this scenario from the user perspective. Respecting the larger agenda of this research which aims to understand the challenges of using personal data in novel media, there is the call to further expand on these findings identified in order to gain a more holistic view of this landscape.

Inspired by the notion of viewpoints from Software Engineering, the research now turns towards looking at this emergent space of personal data use in media from the perspective of the service provider. Understanding the service provider viewpoints on the challenges associated with personal data use would help further explicate and maybe add to the already identified challenges from the point of view of the makers of the experience, thus resulting in a more comprehensive mapping of this space.

To enable this initiative, this study was designed to conduct 20 interviews with service providers from the BBC, a media service provider who was at the time shifting towards a model of personal data collection and use

The BBC was partnered with for this research study primarily because of the opportunity it provided to study data driven media of different

calibers. On one end of this spectrum were BBC teams like BBC iPlayer, BBC News and BBC Sports which competes with and provides services similar to that of mainstream media and data giants like Netflix. Hence, their perspectives, approach and processes would be more in line with these external parties as well. At the same time, the BBC also houses teams like R&D, Childrens and Learning whose work with more vulnerable populations and sensitivity to the wider social implications of media experiences make them approach their work with slightly different perspectives. While these teams are varied in certain aspects of their media service delivery, the BBC being a public service entity, the organisation as a whole is interested in developing and disseminating media experiences that act in the best interests of the public.

Thus, working with the BBC for this study not only enabled studying these mix of employees who served a varied range of media experiences to the British and international audiences but also opened up the possibility of working with teams that would be more representative of different kinds of organisational perspectives on the use of personal data use in media experiences, while being underpinned by the common theme of public service. Thus, the findings here would be applicable not just to the BBC, but with the growing user base for services like Netflix and Amazon Prime. The increasing societal implications of the data collection by these media services means that the results of the study would be a response to the wider practitioner need for a better understanding around data driven approaches that act more in line with the audience interests.

Also, to note was the timing of the study during which the BBC's personal data collection initiative was launched, which made the entire organisation more sensitive towards it. This encouraged many teams to engage in more dialogue around the shift, its challenges and its future, thus providing fertile ground for the study to probe and explore this space effectively.

Thus, this study reports media service provider conversations about the challenges of using personal data in media, guided by seeding topics around the kinds of data used in media experiences, the motivation and purpose for the same, the benefits and risks posed by the use of personal data and the near future expectations for personal data driven media.

5.2. Participant Recruitment

A consent form was communicated and served to the participants. The **ethics** for this was cleared through the School of Computer Science at the University of Nottingham.

The recruitment began with the help of an individual within BBC R&D (familiar with the work of various teams across the BBC who

were turning towards personal data use), who enabled access to certain key teams. An initial set of participants were chosen from different teams with careful consideration regarding their involvement with the use of personal data in their services. They were then contacted individually through recruitment emails that outlined the aims of the research and their involvement in it. The rest were recruited by snowballing references through participants who identified potential candidates who worked within the same team, but in a role starkly different from theirs. This selection helped the recruitment of more than one participant from most teams and ensured that the different interviewees from a single team served in different roles with respect to their use of data, to help maximise the representation of viewpoints.

A total of 20 interviews were conducted with stakeholders who worked in diverse teams, which included BBC Sports, BBC News, Children's content, Education, BBC iPlayer, BBC Radio, BBC Research and Development, Audience Platform, Marketing and Audiences and Data Management. These participants served varying roles within these teams which included engineers, technical architects, digital planners, analysts, UX practitioners, designers and producers. These participants also worked in diverse capacities and hierarchical levels within the organisation which included, directors and executives (4), department leads and heads (5), senior professionals (7) and relatively newer recruits (4). The group included 6 females and 14 males with ages ranging from 20s to 50s. This mix of teams and experience ensured representation of the varying priorities raised at different levels of work organisation while serving a wide range of media experiences. The common connection that served in the selection of participation was the individual's involvement with data driven technologies and personal data. The following is a list of participants and their role/ team within the BBC (the names have been changed for maintaining participant privacy) .

Participant Name	BBC Team
Lewis	UXD
Nate	BBC Audience Platform
Parker	New Age of Wonder
Jessica	Audience Planning
Walter	UXD
Schyler	Data Management
Angelica	Research and Development
Mary	Children's
Moses	Sports and Children's Digital Architecture
Alan	iPlayer/ Central Editorial

Megan	Audience Planning
Harry	Audience Analytics
William	BBC Taster
Charles	UXD
Diana	BBC News
Carlos	BBC News/ Digital
James	Voice Software Engineering/ Taster
Prince	myBBC
Ed	TV and iPlayer PC
Duke	IRFS Engineering Team

Table 1. List of Participant Teams.

5.3. Data Gathering

The study method consisted of interviews and was performed in collaboration with the BBC, a large publicly funded media organisation within the U.K.. After years of resistance, the organisation was in the process of shifting to mandatory user sign-ins and collection of personal data. The BBC serve a diverse variety of media experiences catering to audiences of varying demographics. This includes News, Sports, Children's entertainment, Knowledge and Learning, Radio, Video on Demand, 360 degree experiences, Virtual Reality experiences, Live Broadcast etc.

As detailed in Section 5.2., service providers who worked in diverse roles in different teams of the BBC, who serve very different kinds of media experiences were chosen. This method of participant selection was chosen to ensure representation of views of service providers who worked in different levels and played different roles with respect to their contribution to the service, thus ensuring inclusion of viewpoints of a diverse range of individuals who work with personal data in media. Whether their views were the same, had overlaps, were closely aligned, yet dissimilar or starkly different was something that the study would further help reveal.

The choice of interviews over other methods like focus groups or group workshops was due to the need to uncover the varied range of unique practical reasonings in detail. While an individual interview would help each of the participants go deep into the priorities and challenges faced by their particular role while serving their particular service, a group discussion might lead to a less detailed discussion of just the most overarching and popular concerns that apply to the majority. For the purpose of this study, where the focus is on uncovering the most practical reasonings and responses to the use of personal data in new media, the need for probing a diverse range of service providers working on different services, in different

capacities and uncovering each of their most significant viewpoints was of highest importance. As mentioned previously, whether these viewpoints overlapped, varied or was significantly contrasting in nature would be a finding that the study would uncover rather than the researchers designing the study to force the findings to just present the most common themes.

The interviews were either in person (3) or by audio call (telephone or Skype) (17), and lasted thirty to forty five minutes. All interviews were audio recorded and the recordings transcribed for data analysis purposes. The data collection and transcription was done simultaneously where the collection was discontinued once the data reached saturation, as the responses and discussions became repetitive.

5.4. Design of the Study

Since the participants hailed from teams that served vastly different kinds of experiences and hence had different priorities, in order to successfully achieve the goal of the study, the interviews had to accommodate for the diversity in the views of the participants. Simultaneously, the interviews also had to accommodate dialogue around the issues that the study aimed to learn about.

The interview design was unstructured, consisting of a set of topics that helped guide the conversations if they ever stalled (Robson, 1993). The three areas of interest that were used as seeds to drive the conversations were:

- the current collection and use of personal data,
- the benefits and risks of using personal data, and
- the future of personal data in new media.

These seeding topics were selected with the aim of encouraging the participants to actively engage in dialogue about their unique reasonings around using personal data. Hence, the topics were broad enough to initiate conversations for all the interviewees but also delivered the possibility to engage in deeper dialogue. For this purpose and with the view of using the results of the study in design endeavours that would enable future data driven media experiences, all the three topics were applied to the context of the service and role catered to by the specific interviewee, helping them respond through their very practical reasonings. These contexts also worked as probes, in the most mundane sense of the word as in, they encouraged conversation, often helping the participants express significant aspects of their very unique everyday priorities and challenges associated with using personal data and to visualise future scenarios that the participants envisaged would help overcome the current challenges they face. Here, the first topic helped answer the question

of ‘what’ the data was and ‘how’ it was being used, the second topic helped explore the ‘why’s and ‘why not’s of the use by exploring the benefits and risks of the same and the third explored the ‘what if’s of the future of this scenario.

The conversations began with the researcher introducing the research and the aims of the study, followed by the first seeding topic of the current collection and use of personal data by the participant/ team. This helped initiate the participants into the scheme of personal data driven media without overwhelming them. Here, the conversation around the uses of personal data probed the participants to think more deeply about their particular service and how they leveraged personal data for its enhancement, which provided an easy transition to the next seeding topic of benefits of personal data. The risks, if not explicitly pointed out by the interviewees by this point, was brought up by the researcher, as a follow-up to the benefits, in a manner of discussing ‘two sides of the same coin’. And once the participants completed their conversation around these topics or any other that the researcher or the participant deemed helpful to the investigation, the researcher wrapped up the interview with the final topic of the future of personal data and the media service providers’ projection of the potential future extensions in personal data use.

The conversations flowed where the respondents directed them making the method “informant” (Lazar, Feng and Hochheiser, 2017) in nature. In instances where topics that were aimed to be discussed were ignored, the researchers ensured that they seeded them into the conversations through questions, follow up questions, comments or queries. Given the highly complex nature of the topics discussed wherein the attitudes towards personal data could be wide and varying with the reasoning for the same being highly contextual, we believe the exploratory approach helped accommodate the complexity and multifaceted nature of the data in its full richness.

The collected data was audio recorded, transcribed and analysed using Endogenous Topic Analysis (Section 3.3.3.).

5.5. Findings

The findings of the study unpack responses to the following seeding topics employed by the study in a deep and rich manner.

- the different kinds of personal data used in media services and their purposes,
- the benefits and risks of using this data and
- the future of personal data in media enhancement.

These topics are often contextualised in the everyday practicalities encountered by media service providers.

Analysis of the data also highlighted that the conversations opened up to extend beyond just the seeding topics to expose more topics of interest that were discussed by the participants as part of their dialogue around personal data use. These included :

- the current response strategies to the discussed risks,
- business priorities supported by the use of personal data,
- personalisation,
- formation of user cohorts and
- the concept of user empowerment through personal data use.

They are all discussed in detail in the following sections, often using text and quotations from the raw data, staying true to the everyday practical reasonings and examples the interviewees grounded their views in. Participant's names are not their real ones in the edited extracts below.

5.5.1. Types and Uses of Personal Data

The organisation and its various teams collect and use user personal data of various kinds, through diverse sources. The responses to the interview helped uncover these groups of user data that is utilised to enhance a varied range of media services.

5.5.1.1. Behavioural Data

The most commonly talked about personal data type was behavioural data. This data describes user behaviour on BBC platforms and was often collected passively through tracking user interactions on these platforms.

The following quote is an instance of someone describing how behavioural data is tracked on BBC platforms and then subsequently used for personalising the BBC experience for the users.

> *Kim: We will track your activity on our website and then we will use that information in a way which allows us to then target um….target some of the BBC experiences based on the personal data a lot of which is based on behaviour data.*

As explained by the quote, this type of data is often used to personalise the experiences, both content and recommendations, with the aim of increasing their relevance in the users' lives. Different types of behavioural data like the viewing history, browsing history and interaction logs are leveraged by service providers to help build a rich picture of the users on the BBC platforms, helping them

understand and serve these users improved and tailored content that is relevant to them.

5.5.1.2. Demographic Data

The other main kind of personal data discussed was user demographic data. Demographic data is data that helps describe the user better and helps with the formation of user cohorts that could be used to spot and respond to trends in consumption.

> *Paul : We do quite a lot of... we collect stuff like the location and the age and the gender. So we are able to say, I don't know on the lines of.... X number of users visited iPlayer and y number of users visited Sports, this percentage is male, this percentage is female, and they are between certain ages and around which location of the U.K. they are based.*

Different types of demographic data are collected. They include user age, email addresses, location and gender.

> *Michael : Personal data comes into three buckets from my perspective. They are credentials and personally identifiable information, people call PII and the other data is for example your age, your date of birth, your post code of where you live, obviously your email address and username and a password.*

Age or date of birth is collected both for service enhancement that ensures dissemination of relevant content and also for legal and regulatory reasons to ensure the right kind of data collection, personalisation and service delivery is ensured for every user and also to ensure that the services are relevant to all ages of audiences.

Other demographic data like post code and gender are collected to fit into this regulatory model where younger audiences are exempted from providing data like their post code but required to provide their full date of birth to help calculate their transition day to a higher order age category.

Here, location was spotted as a particularly good indicator for studying and predicting the kind of content a user might be interested in, as explained by the following quotation.

> *Lolita : We have started rethinking, recording post code as part of the license fee changes. If we knew there was a regional thing, so we knew people in Manchester like Manchester United for example, we could use that knowledge if we knew where a person's based. So I think we could use it to create more tailored experiences and better understand how people interact differently.*

Demographic data is used for the following :

- understanding the users by uncovering the response of the audiences to the services provided,
- to respond to consumption patterns by highlighting trends and gaps in consumption,
- to provide services that are relevant to all users by spotting and removing any kind of demographic bias and
- staying in line with regulation by realising the law through formation of appropriate user cohorts (Eg., age based cohorts that manage the delivery of the legal grade of personalization to be served to audiences of different ages).

5.5.1.3. System Specific Data

The interviewees also saw other system specific data like IP addresses and cookies as material worth discussion when talking about personal data. This kind of data was used for *geoblocking*, *spotting frequent users* and for the *maintenance of incident logs* of user interactions with the services.

> *Alex : IP addresses do sometimes get used for things like geolocation and theres a couple of things that we have done that do geoblocking so you know that content can be only available in a certain area.*

> *Jess : We do have cookies, we do have quite a long lasting session cookie which actually lasts for a year, so we have some sort of arbitrary token that lets us identify repeating users.*

5.5.1.4. Data from External Sources

Along with all the personal data collected through the various BBC platforms is the data that is triangulated from external sources and third party platforms. These include data from Facebook and Telegram where users engage with BBC content.

> *Casper : The information that comes to us through Facebook, which is for our Messenger pilot which is mostly anonymised by Facebook before they pass it onto us….[….]…..we just collect whatever data is passed onto us by Facebook…..[….]….. So what we are using is we are piggy backing on Facebook's systems, so people's Facebook identity. I'm afraid I'm not sure how much of that data we get or whether we just get little bits of it passed onto us by Facebook…..[….]…..So, we are using Facebook, so we are putting our experiment into Telegram, into Facebook. So we are going onto social media platforms and we are offering*

experiments, we are offering BBC pilots, BBC services which people can use, using their existing credentials on that platform.

Also mentioned was the separate silo of user data from TV licensing. Currently this data is managed entirely separate from the data collected through online services but it poses the potential for future uses, yet to be explored.

5.5.2. Benefits of using personal data

A number of benefits were outlined for the current shift in media services towards leveraging user personal data. *Understanding the audiences* and their wants and needs was considered one of the most prominent advantages personal data offers. The second was to *provide an improved experience* to the user through tailoring, personalisation and providing richer and even novel genres of experiences. Thirdly, to *provide good value to all the users* by learning consumption trends, identifying gaps and looking into measures that would fill them by serving all audiences with what they are interested in.

5.5.2.1. Understand the Audiences

Personal data affords the service provider with the invaluable benefit of understanding their users and their preferences at a *very granular level*, through empirical methods rather than making philosophical assumptions about audience responses to delivered content, as expressed in the following quote.

Mike : You need to understand who your users are and then what do they need, and then how do they actually behave in order for you to start saying oh… they actually want more of that than what I thought they would like.

The need to understand the user is of significant interest as it enables the service providers *to overcome current flaws* in their production processes, further *enhance systems* that are deemed popular and ensure *delivery of content that is relevant* to the audiences contexts.

Eve : We can use personal data to improve our marketing, our user journeys and our product experiences. We can make those things more tailored, more appropriate to the user, we can also use the information flowing back into the system, to improve as we go along. So, for example, we can also use it to enhance our own internal skills, we could learn more about what users want from us, we could improve our curation skills in that way, we might spot a pattern or behaviour that means that we offer up a certain mixture of content.

Audience reaction to content disseminated is often seen as a metric for evaluation and also as feedback to be used for future production cycles. This data coupled with other audience data like demographics helps *understand systemic issues and biases*, particularly in news media, that the content might be embodying, which could then be worked upon.

> *Sam : To help us understand what kind of people are responding in what ways to our experiments. Because the work we do is very much about trying to work out new ways to deliver news to our audiences. So we need to understand who from those audiences are reacting in what ways to the experiments we are doing.*

Also discussed was the potential to *understand user intentions* to deliver experiences that are aligned with every user visit. Thus, understanding user intentions in every visit to a platform is expected to be a radical novelty that would help transform the current service delivery norms as it would help the service providers deliver the exact kind and length of content that the users anticipate, as detailed in the following quotation.

> *Kate : It will, what will we use it for, again it's about providing useful experiences to people. So, if you are just wasting time, so you just want to be entertained, then we might provide you with smaller, more upbeat, more light hearted news. But if you want to be informed about a topic, then we would offer you more deeper, more related content and deeper related content to the same content that you are interested in.*

5.5.2.2. Enhancing the Experience

Understanding the users through their personal data was expected to help deliver enhanced experiences to the users. This enhancement was seen by different interviewees in different forms.

The first was the ability to deliver a very *specific* experience that would be personalised and hence tailored to fit the users' context, life and interests, as illustrated by the following quotation, through an example.

> *Meg : You could provide your users a more specific experience you know ....specific to their specific individual needs, rather than a more generalised view of the user needs of a lot of people…[....]…. So for instance you know, in the case of the Bitesize app, somebody who is going to school in Wales would have a different set of content from somebody who is going to school in Scotland.*

Another possibility was that of creating *new genres of experiences* through personal data. One of the interviewees talks about an application that they had used, which was novel and based on their personal data. The interviewee then goes on to propose his expectations of personal data being able to unveil such new genres of experiences that would interest the user.

> *Archie : And then another thing thats interesting to me is that ...[....]..... Basically what is does is it allows you to sync up in an OAuth way to your online accounts like Twitter and Facebook and Instagram and what it does is that everyday it gives you sort of an on this day in the past view of your social media accounts in the past and what you posted sort of todays... whatever to day is... what is today? Today is the fifteenth of March so in TimeHop today if I go on to my TimeHop app it will, it will be telling me what I posted on March 15th in the past.... well whatever year. I think what I guess what Im trying to get at there is that thats a whole experience that based only on my personal data and its based on my personal data not yours. Or rather yours is based on your personal data and mine is based on mine. And its a kind of genre of experience that I think is very powerful and has been..... been popular as over the past five, six years you know.*

The third is to be able to make the experiences *richer and more engaging*. By knowing more about the user the service providers expect to add in more layers to the experience that would be relevant to the audiences and appeal more to them, resulting in more reasons for the user to engage and immerse in the experience. Examples of using personal data to enhance the richness of the media experience by allowing for user participation in the media experience through their data is outlined in the following quotation.

> *Rinni : You know one example is to customise the experience so, for instance, you know there might be one, we did one pilot where you.... was set in a world of a TV programme called Hinterland... its sort of like a ... comes out of BBC Wales and in this experience you play the part of someone solving a crime and there are a couple of things.... customising the experience, you could see a newspaper and it has a picture of yourself in the newspaper at the end, I think there is also a poster on the wall that is also done and those things were done by getting your Facebook profile. So that was, in that case it was customised to make it have more of an impact on yourself. Another example of a pilot using personal data is one where you could upload your face, and this was one.... we did one for Strictly and we also did one for the ^*^*^ Christmas as well, but the strictly one, had videos of people dancing from the Strictly come Dancing*

show. If you would upload your face and then it would upload your face instead of one of the dancers faces, and then the result was able to render a video of what looked like you dancing with a professional dancer in one their routines.

5.5.2.3. Providing Value for All

Also cited as a benefit of using personal data was to provide more *value to all* members of the audience. Here, the idea is to move from a generic one-size-fits-all model of media delivery that is expected to deliver one value for all but finding value for everybody by understanding the users through their personal data.

Betty : There is also very strong responsibility within the BBC that we deliver something for everyone. So everyone pays their license fee, we want to make sure that we are giving each of them some value for their license fee.

The following quote comprehensively expresses the potential of personal data to help *understand consumption patterns*, *unveil gaps in consumption*, *spot underserved audiences* and *adopt measures to overcome these flaws*, so that all the various groups and segments of audiences are served.

Taylor : So for me….the big win, which is going to take a while to get to is that we will be able to properly understand which audiences, or which cohorts of audiences or which sections or whatever it is, however you would describe it, which audiences are watching what types of programmes because I think that we have got gaps in… in what kind of programme we create and the way we tell stories, that once we know more about the audiences and what the audiences are doing and we can currently do, by identifying the gaps that we are not doing, to be able to try and encourage those younger more diverse audiences to try and make sure that we are creating the right sort of content experiences for them. And thats the issue for me, its….. its about the sort of underserved audiences that we have. I… I think we do loads of great stuff for certain sections of the community, but we don't do enough for other sections and I think, I'm hoping, that when this data becomes more mature and more people understand what it will…. what it means, is that…. that will highlight some of the holes that we have got in our content strategy.

If all members of the audience are not served well, the gap is identified and further worked upon by involving user data in the decision making process. This could be by responding to the gap at a macro level (based on age or location) or even at a micro level by leveraging the personal data to speak specifically to individual

audience members. Thus, a data driven approach is taken towards serving relevant content, ensuring that future endeavours reflect actual user priorities rather than service provider assumptions.

5.5.3. Risks

A number of risks associated with using personal data were discussed in the interviews. These included fear of reputation damage, intangible risks, constraining and biasing user attitudes, the risk of serving inappropriate content, creative challenges, data security risks and loss of user trust.

5.5.3.1. Reputation Damage

The potential of personal data, if not used in a diligent manner, to lead to possible maligning of the organisation's reputation was highlighted by many as a major risk.

> *Harper : The main risk is that you could really screw up as an organization….[….]….. And your reputation would be ruined as an organisation you know.*

This organisational risk of reputation damage meant that the service providers saw themselves as having more *responsibility* to their users, with regards to the personal data they collect and use and the content that they deliver, as cited in the following quotation.

> *Amith : So theres lots of risks…. I mean theres the…. theres reputational risks in terms of we are collecting lots more information and therefore we are…. you know, trust is, is such an important part of the relationship that we have got, with the BBC, so thats a lot of extra responsibility, a lot of extra things that can go wrong.*

These stakeholders recognise their role as a data controller and the importance of being both *legally* and *ethically responsible* for their use of personal data, so that the legacy and reputation of the organisation is preserved.

> *Smith : Ok legal risks, in the sense of, if we are thinking in terms of the audience then, in the data protection law, we are the data controller, I think that's the terminology. And I think we have certain legal obligations to look after, in our role as the data controller, that we do appropriate things with people's data. And then there's also just….. I think from our own point of view, what we want to do our own ethical things as far as the BBC's reputation is at risk.*

Also acknowledged was the added *social responsibility* of the organisation as a data controller to ensure maximum security and accountability over the data that is collected so that it is not hacked

or let out to anyone, which could lead to the organisation's reputation being damaged.

5.5.3.2. Intangible Risks

The risk of making out-of-context decisions while using user personal data was identified by some as a risk. These risks were termed 'intangible', 'judgemental risks' because of their nature which did not cause legal repercussions to the organisation but resulted in social discomfort among the audiences, leading to user dissatisfaction.

> *Piggie : Risks for me are again like the intangible risks. So, I care very much about people making decisions about me based on out of context behaviour.*

An example of this form of risk is detailed in the following quotation. Here, a member from Sports explains how they should be careful about revealing just the right amount of information about a user's preferences regarding Sports teams to other users so as to prevent social tension.

> *Angelica : Sport is a very passionate subject and I think thats exactly…. some audience members want their voice heard, they like to have other people see their opinions and things like that. Their sort of interests and sort of things that they can output but obviously the more we look at user personal data, the more we should be conscious, that not everyone wants to put themselves out there in those sort of forms so their…. not to safety degree but I think personal, that people don't want to have that information broadcast because their love of a sporting team might go against their sort of either social group or maybe where they live, all those sort of things. So that is quite important I think. Because Sport is a very passionate thing. Its also like, for us in Manchester when we talk about Man United fans and Man City fans, those two have very conflicting views in a very constructive way about Sport and you know thats really good but also thats not something we should just broadcast publicly to everyone….. everyone information.*

Thus, the organisation feels that a certain sense of social accountability is demanded when they use user personal data to enhance media services, especially owing to the close relationship between media and one's social life.

5.5.3.3. Constraining and Biasing User Attitudes

An interesting comparison of a user's browsing space in the internet to that of a neighbourhood, helped demonstrate the risk of personal data being used to bias and change user behaviour to match and benefit organisational expectations.

Elise : If you think about back in the days, you lived in a neighbourhood, this neighbourhood has a specific group of people associated with it and so socio-economic kind of status. This information has been then aggregated and sold to advertising when you are walking in a specific neighbourhood you have a different advertising. The same thing happens right now online, however, we can be much more specific. You know much more latest precision targeting specific people.

With access to this kind of enhanced customisation afforded by user personal data, media service providers recognise the importance of using it ethically. Within media services that leverage user personal data bias manifests in two ways. The first one, particularly prevalent among the news media providers was the need to prioritise being nonpartisan in their content delivery and serve the audiences with content that would not constrain or bias their views and attitudes, particularly when enabled with personal information on the audience's preferences.

The second is the fear of pushing irrelevant content to groups of users based on inaccurate decisions made due to their membership in a particular cohort. This type of bias was considered a risk that constrained user experiences, thus leading to user frustration (also identified in Study I). The following quotes express this concern by citing age as a potentially inaccurate indicator of user interest, if not used in a diligent manner.

Schuyler : Um… so, you might want to make some decisions based on age, but to be honest, I'm mostly looking at behaviour, as ways to kind of. Well, ways to show content to people. So, I find that I don't really want to discriminate by age, you know, somebody who is fifty may be interested in the same stuff as somebody who is twenty one. And behaviour is probably a better measure for whether somebody is interested in something.

5.5.3.4. Serving Inappropriate Content

The ethical need to serve users with appropriate content is further emphasised with the involvement of users below the age of eighteen. With the use of shared devices and tailored user profiles, the risks of children being exposed to content inappropriate for them is increased, which calls for attention.

Martha : I suppose the other thing is, this is probably less…. for the BBC, which is a trusted brand, meaning that people have lots of faith in what we do and all sorts of consideration behind the scenes to make sure our services are appropriate for the intended audiences, you know….more so in the case

of Children's content, where we go to great lengths to make sure that our online users who are children, doesn't necessarily have access to certain aspects of what we do. And then we think about parental approval etc. When you have signed in users or signed in devices that are potentially shared, there are risks there.

5.5.3.5. Creative Challenge

Another risk outlined was the creative challenge of utilizing data to its fullest potential. As a service provider, once you start collecting personal data, the responsibility of giving maximum value from that data to the users is one that is both socially and ethically driven. The following quote uncovers that call and proposes proper training of organisational employees as a potential response to this need.

Lynn : And then there's also the risk that people aren't trained well enough to understand how to use that data for the benefit of all audiences, because it's kind of a new thinking and it's new and so there's a lot of change needed in people's minds and the ability to understand how to utilise and exploit this data for good, rather than it just adding to the noise of information that's there.

5.5.3.6. Data Security

The importance of securing the personal data collected was highlighted continuously in discussions about risks by almost all the participants, irrespective of the teams they worked with or the jobs they catered to. The potential to become a target for hackers was perceived as a high-stake risk that should be effectively responded to.

Caspian : If its exposed for other purposes that wasn't designed to be collected, or if it is exposed for purposes outside of the BBC that is bad on our part both reputationally and legally that's wrong on our part. So, we have got to put lots of security around it.

5.5.3.7. Trust

The repercussions of all the previously discussed risks, from reputation damage to making out of context decisions for the user to serving inappropriate content, ultimately bleed into the overarching risk of loss of user trust.

Adding onto the loss of user trust from these various sources is the risk of being intrusive as a media service provider during the collection and use of personal data. The participants cited being

"*creepy[*P2]" as a risk that would easily undermine the trust audiences have on the organisation and its services and one that they have taken up as a 'mission' to minimise.

> *Jill : So, we, the BBC is a very well trusted brand and you know to maintain that trust we need to ensure that we are not doing anything weird with the data that people provide to us…[…]… And we also have the informal mission of don't be creepy.*

5.5.4. Risk Management

Along with the discussion of the various risks posed by personal data, risk mitigation strategies undertaken were also mentioned. These included responses like *spreading awareness about data practices*, *being transparent*, *adopting a public service approach* to data practices and *technological responses* like anonymisation, data minimisation, encryption and a shift to client-side data processing.

5.5.4.1. Spreading Awareness

When discussing risks and risk mitigation strategies within the organisation, a few respondents admitted wanting to know more about the current risk mitigation measures the organisation has undertaken. These emotions, when probed further, brought forth the need for organisations spreading awareness about data practices both within and outside of the organisation so that both the service providers and the users are aware of the underlying data practices.

It is important for service providers at all levels and roles in the organisation to know about the basic data processes and policies so that they are able to maximise the value they create for their service while using personal data by keeping in mind both the regulations and responsibilities that they are expected to cater to. Awareness among the users is a priority as it enables the user to make decisions and exercise agency within the data economy that they are involved in, thus empowering them. The following quote explains the same two initiatives from the perspective of the service provider.

> *Adrian : So, there is a whole lot of extra training that we have created and delivered in terms of the risk around the data security and all that kind of issues and making sure that thats all…. so theres you know, theres loads of… theres loads of processes in place and extra things that we have done to…. to make… to do as best as we can in terms of the trust aspect. In terms of the education and change piece for winning the hearts and minds of people to understand what they can then do with the data, that is an on going…. sort of process, that involves lots of change activities, lots of evangelising but the BBC is quite a…. big legacies, complex*

organisations like the BBC would take a bit longer to change.

5.5.4.2.Transparency

One of the key risk management methods discussed for spreading awareness, particularly among the users is through provision of transparency through design.

> *Andy : So, we have worked with the information commissions office to come up with a strategy that helps our users to understand how we use our data. So, our mitigation for that trust question is that we try to get really transparent in how we use the data.*

Transparency would be provided at different stages of the service and user experience, beginning right at the Terms and Conditions statement, extending to other aspects of the experience that leverage user personal data. The following quotation explains how the organisation has adopted the use of everyday language to communicate data practices to the users, in an effort to increase the transparency provided.

> *Philip : But what the advice has always been to me in doing this kind of thing is to be extremely explicit, using everyday language as to what is being collected and what happens to it.*

5.5.4.3. Public Service Approach to Data

Reflecting the BBC's funding model and their very nature of being a public entrusted body, some of the suggestions of risk mitigation bordered on taking a public service approach to the data. This would involve an increased focus on the user, where the user wants and needs are always the primary driver for data collection and use, and always exploring avenues for delivering maximum value to the user through diversified manners, as expressed in the following quotation.

> *Derek : And, where we are at, at the moment, in our journey of collecting and understanding the use of personal data, we are at a quite early stage where I think its really important that the BBC develops a public service approach to the collection and use of personal data. Everything we do should be in the interests of our users and audiences. So, the BBC should establish a strong public service ethos. Its really crucial that the BBC is doing this to the best of its ability."*

One of the methods of realizing this public service approach to data was to increase the value returned to the users by diversifying the services offered to them. The following quote describes an example of such sharing of data back to the users in a way that would create

more value for them. This is expected to give them more reasons to engage with the data economy through enhancing their interest.

> *Peter : One very kind of like open sharing of the data we capture…with….back to people. But then, also set design patterns, that help to kind of expose the processes happening behind your interaction with something. So, you know, if….so…. that sort of thinking at the moment around kind of like patterns that allow you to kind of continuously refer back to what is happening and you know that could be something like a tool tip, it could be like, you know, something that you can flip over a thing that you are interacting with and then kind of understand it more about what something is, I guess, and then to kind of unpack that first point a bit more is like, you know feeding back whats happening based on peoples feedback. So, you know like, hey you…. you provided…. thanks so much for… for feeding back on this project, heres what you said, heres what everyone else said and heres what we are going to do. This is why you are here. You know, so you…. you are, you are not hiding the fact that you are capturing data but you are giving some really good reasons why you are doing it. And also, what you are going to do with it.*

5.5.4.4.Technological Strategies

Utilising enhanced security measures to ensure privacy of the collected data at several stages of collection, processing and storage was mentioned by many interviewees. Here, anonymisation and encryption of data was often emphasised upon as a method to increase data security and privacy.

Data minimisation(Supervisor, 1995, 2017), was also seen as a strategy that would help ensure better accountability as collecting the minimum amount of data from the users is seen as a method of 'risk aversion' from the organisation's side.

> *Cox : Obviously quite tempting to want to get more so that we can kind of segment the answers more…[…]… the BBC is very careful about collecting only the data that is most essential for their services …[…]… We don't… we go to great lengths not to ask for anything that we don't absolutely tend to use. To minimise the risks of privacy implications.*

The possibility of minimising data exchange with the organisational servers through client side data processing was another technological strategy that was discussed. This kind of shift means that the organisation is not effectively collecting or storing user personal data but simultaneously using it to derive value for the users, thus eliminating many of the risks discussed here. The following quote

explains an example of one such service where user data was processed on the client side without transfer to BBC servers, thus eliminating the need for data collection.

> *Moses : That Digital Me, we didn't actually collect it. If you see what Im saying right? So, it wasn't actually collected or stored on BBC servers and this that and the other. It was just brought into the browser and the browser crunched it and in theory, in throes when you turned the browser off, the data disappears….[……]……the Digital Me project is that we went out of our way as much as we possibly could so that Facebook and Twitter data that we were using to power that experience dint go and sit on any servers anywhere much less than on your herd drives anyway. That was something that we put into the technical architecture of that project.*

While this kind of processing seems like a straightforward solution to many of the concerns highlighted in this study, it was also acknowledged that such alternative measures that provided alternative data practices often constrained the technological innovation to a good extend, as stated in the next quotation by the same interviewee.

> *Joel : But of course it meant that certain things were limited as well. Like say for instance, it was harder for us to have to store states and then you could come back to the application another day and pick up where you left off and those kind of things.*

5.5.5. Projected Future Extension of Personal Data Use in Media

When talking about the future of personal data in media experiences, the responses were varied. Certain respondents felt that they already had enough data in hand and wanted to focus on maximising the potential of this data by looking into alternatives that would deliver more value to the users.

> *Seb : Honestly, I don't think theres any data that we should be trying to get that would directly affect the quality of the experience that we could give our audience. I think we sort of pride ourselves on not being intrusive, not prying, to allow the users to have complete control over their data that they give us and ensure they get value from providing it to us. Whether its sort of passively or actively.*

Others deliberated over collecting more data from the users, while carefully considering the ethics and justification behind the same. In the following quote, the interviewee expresses their need to collect data on the users' ethnicity, following it up with the reasoning that supports this decision. But, this is quickly followed up by thoughts

on whether such a move would be 'comfortable' for the users, checking if the rationale for the data collection was enough to make this extension in the future.

> *Quentin : To some extent, it would be, it would be interesting to see if we are catering across to basically the whole spectrum of the British society. I would like to know around the…. yea, cater to the needs of sort of like your average age, gender and location, but perhaps sort of the ethnic background of our audiences. I'd be interested to know but just like my first point, just to make sure that we are doing that but that we are delivering value to everyone. But I'm not sure at the same time that these are information that I would like to ask our audiences for at the moment we are asking for your age or your date of birth, gender and location. And I am not sure how comfortable they would be in answering what is your ethnical background or your religious beliefs and things like that…. I would like to…. I would like for us to deliver value for everyone but I'm not so sure that I would like to ask for that information. So, its a rather tricky situation.*

The rest of the participants were quick to think about more data that they could use to help them enhance their technologies in different ways. An example of the same is :

> *Callum : Well you could get, yes you could get much more granular, say for example what football team do you support? You could get much more granular around interests. What specific groups do you like? You are listening to…. lets say to music stations, what types of music? Yea, so what genres of music? So you could go at…. at the moment audiences can go and choose, kind of collate their own playlists. It might get to a stage where BBC creates those playlists depending on the data that people have provided. And, similarly with football teams and things like that you can create your own…. your own filters but then again, the BBC might capture that data much more centrally yea and serve information based on teams or based of types of news that you might want to look at.*

Another way of extending the current use of personal data in the future was by accessing competitor data which is expected to help understand user interests better by learning what they were consuming from different service providers and to help diversify the type of experiences currently served to the audiences. Meta-data enrichment was another possible future strategy, which was cited in different instances. This included pulling together different data sets like behavioural and demographic data, data between different Sports experiences (like using tennis data for football experiences) and exchange of data between teams. The current inability to access

children's data and future aspirations for being able to collect and use data from them to help with easier interactions in Children's experiences was another example of future extension of personal data use discussed.

These data are currently not collected or used in the proposed manner primarily due to economic barriers like lack of funding, technological barriers like infrastructure issues, legal or regulatory barriers or ethical concerns that challenge the need for such extended collection and use of data.

5.5.6. Business Priorities

The results show that the use of personal data helps achieve and regulate a range of organisational business priorities as follows.

5.5.6.1. Fulfilling the Business Mission

Ensuring *delivery of valuable service to every single user* is one of the BBC's premier missions especially because it is part of the mandate that supports their very right to exist and serve the society. Personal data is expected to become a key tool in this process.

> *Heron : It has to do especially with the BBC's mission of delivering value to everyone. So, by having an idea of the location and age and gender, we can make sure that we are delivering something for everyone.*

Extending further from just the regulatory mandate, the organisation also recognises an *ethical responsibility* to serve all members of the audience. Here, they explain how their unique funding model enables their shift to collecting and using personal data to be driven by this ethical need to provide value to all users as opposed to being imposed for exploiting user personal data to generate income through advertisements etc.

> *Mik : Our unique funding model… we don't need to like sell advertisements and target people based on I don't know, their socio economic background and things like that which other companies do. The BBC only uses personal data and information for delivering value to everyone. So, the benefit of collecting this information is precisely as I said before, to have an understanding of who our audiences are. And how our outputs reflect maybe…. what is expected of us, in terms of giving content to all of our audiences.*

Use of personal data is also expected to *shift the relationship* between the service provider and the user from the traditional one to many broadcast model to one to one relationships that cater to each users' particular needs and interests.

Jack : Optimally the world is changing where rather than one to many, its one to one relationships that organisations create with their audiences and customers. And so, we have to move into that world.

Personal data provides opportunities to understand consumption trends across various demographic segments, based on age, location, etc. This helps as an evaluative tool in *learning how different content is consumed* across the spectrum of audiences, identifying gaps in consumption and making amends to future content creation to ensure consumption across all cohorts.

5.5.6.2. Provide Symmetrical Exchange of Value

Providing users with a good exchange of value is vital to the success of any business. Provision for user accounts and profiles are expected to give users opportunities to express themselves better online and to be more active participants in the data economy. This enhanced contribution from the users could then be reflected to them in the experiences that they receive in return. These reflections could go beyond traditional media services to educational initiatives that help with personal growth, self-reflection programmes that help the user make sense of his/ her personal data with regards to contexts frequented by the user, interactive experiences that show a comparison of user data against the rest of the audiences etc.

Nate : So, we are going around public service remit for data so where Amazon and commercial organisations it is very much about how they can market it back to you, so our question was what was a public service organisation, that's not making money do? So, we are looking at what's the most meaningful feedback? If you give us your data, how can we give you something back? That is enhancing for you, as opposed to just for us.

5.5.6.3. Help with Product Development

Currently, product development in media is based on input from many different sources within the organisation with the user having very little participation in the process. With the use of personal data, users can play a more central role in this process with the service providers being able to understand user needs down to the granularity of the specific devices used.

Serena : So, I guess it informs the development of the product. So, it basically helps to inform staff, so for example, this is the kind of stuff that can help tell us, you know, TV platforms for example, are becoming more and more, are becoming a bigger and bigger part of [Video on Demand] usage. That's then really important for us to say, ok, should

we then be putting in more refills on TV, should we be trialling more stuff on TV, that kind of thing.

5.5.6.4. Staying Relevant

With the high level of entropy in the number of products and services that are available to users, it is an existential need for every business to stay relevant in the users' everyday lives in order to ensure their sustained success and growth. Personal data, is currently considered to be a tool that would help understand the users' changing wants and needs and hence help the business and its products stay relevant to the user in the current day.

Blaire : I also think that there's a risk though in a sense to avoiding collecting personal data .....[....]…… I think what I'm saying is that consumers demand personalised experiences and that requires personal data. And if you cannot supply them what they expect, then they will consider you to be irrelevant.

To stay relevant in the users' everyday lives, the organisation should also ensure that the services they offer are on par with those offered by their competition. The interviews also uncovered that one of the motivations for shifting to the use of personal data would be to serve the audiences with their expectations, by keeping up with competition, which demands a shift to personal data leverage.

Polly : The BBC doesn't realise that lots of its competitors and lots of people who have digital experiences, have much more personalised experiences, through their digital activity. So its really sort of falling onto line with sort of what audience expectations are ...[...]....I also think that theres a risk though in a sense to avoiding collecting personal data because especially some sectors of the economy, its just…. its irrelevant. You know, you are just irrelevant if you don't. So, its a tricky place to be in.....[....]…… So I think what I'm saying is that consumers demand personalised experiences and that requires personal data. And if you cannot supply them what they expect, then they will consider you to be irrelevant."

5.5.6.5. Importance of Trust

Trust was stressed as one of the key values of the organisation and as an international brand, it represents a public service entity that has always worked in ways that upheld user trust as a priority. This trust, that has been built over the years, is part of the organisation's legacy and is very important to the organisation. As a business, they recognise the possible erosion of user trust the mismanagement of

user personal data could lead to and hence are very careful that their experiences reflect this value at all levels and dimensions.

> *Thompson : There's another risk which is…. and this is back to our relationship with our audience. So, we, the BBC is a very well trusted brand and you know to maintain that trust we need to ensure that we are not doing anything weird with the data that people provide to us.*

5.5.6.6. Legalities of Using Personal Data

The General Data Protection Regulation was referred to as a guide for designing future media experiences. It was acknowledged that future regulation supports provision of higher levels of data transparency to the user which is aimed to empower the data subjects with an understanding of data practises around services offered by the organisation. This puts more pressure on service providers to deliver valuable services where the need, use, consent and control of personal data is easily justifiable not just to the user but legally as well.

> *Christian : I think what the GDPR is bringing is a bigger literacy about us, of users giving away our data to get better benefit, but we need to be conscious about to whom we give that data and what is the actual value we get from that. Because if we are much more literate, people will have to treat our data with care, with much more care. And if we end up in a situation where this information would be lost, that would be a bad thing for the organisation and we might just lose completely, the trust.*

5.5.7. Personalisation

Personal data affords service providers with the benefit of personalisation, which was discussed in varying contexts and from diverse perspectives in the interviews. Personalisation helps the organisation provide tailored services to users, building a one to one relationship model, where the user individuality is respected, thus helping the service providers meet audience expectations and keep up with competition.

> *Gideon : Gone are the days where… the BBC…. well, so the BBC still does broadcast to the many but its trying to, as well as trying to broadcasting to the many, its also trying to tailor to the individuals.*

The benefit of using personal data to provide users with a more specific, richer and diverse selection of relevant content was discussed previously. This section extends that discussion, focusing upon personalisation and how it affords *enhancement of experiences* and *user journeys* in novel data driven media. The creative

contribution of personalisation in *crafting current* and future experiences and the *legal implications* of the same is also recognized and discussed here.

5.5.7.1. Personalising the Media Experience

Personalisation is used to enhance new media experiences in various ways including both personalisation of recommendations and notifications, as well as the content served, e.g., video, audio, news, sports etc.

> *Max : It's not about recommendations like, because we know about you, you might like to buy this book. It's more like we know that you live in wherever it is you say that you live in off Facebook and we know that the main football team there is and that kind of thing. So the football team in the story, it's the team that's closest to you….we think you like that.*

Personalisation of different types were talked about. An example mentioned in the following account is the possibility of having 'explicit' and 'implicit' personalisation based on how the data is collected from the experience.

> *Molly : Another thing that we are going to use ID for is we are going to customise the experience for the user. So, that would be based partly on things that they explicitly choose so they could sort of say what types of pilots they like so we make some changes to the Taster site based on that. But also implicitly with their history so, for instance we are always suggesting, well we are suggesting what other pilots you might like to try, so the persons already tried one and rated it, so we don't need to recommend that anymore. So, customising based on what they have already done, is an example.*

Personalisation is also extended to scenarios where the user collaboratively engages with the experience to drive it in a unique and personalised fashion. A novel media experience that outlines a practical example of this style of data driven media, is a "cook along" kitchen experience where IoT data from smart utensils and/ or user input are used to continuously tailor the content to match user expertise and pace.

5.5.7.2. Personalisation for Enhancing User Journeys

Personalisation is expected to help in the process of exposing the audiences to the breadth of the content available to them as explained by the following quotation.

> *Welma : So, you know we have many many hours of TV and radio every single day. Both international and regional*

radio stations and TV networks. And we create about one and a half thousand new webpages every single day. Most of all that content…… it would be impossible for any individual to be across it all…[..]… Its just packages of information or content, that may well be of interest to an individual but they are not going to know that its even available… I mean in the digital era just because something goes up in a regional radio station, it might be about a subject that somebody who doesn't live in that region is interested in.

With diverse content being produced and served at such a rampant rate, personalisation offers the possibility of not just exposing the users to this diversity but also pointing the users to the exact content that might be relevant to them in this vast expanse, thus optimising the user journey.

5.5.7.3. Personalisation and the Law

In the U.K. it is expected that not all individuals are exposed to the same degree and type of personalisation. Service providers are sensitive to this gradation in personalisation services and are considering design alternatives that accommodate such legal constraints.

Zachaary : In the UK there is a, we have this idea of not every single user will get the personalisation on the same level as others. In line with the future regulation GDPR, we would not be able to serve personalised recommendations, personalised experiences, without the consent of a parent or a guardian, for anyone under 13. So, currently, for under 13 years old have their personalisation disabled by default which means that when they register and they are signed in, they are not receiving any personalised experiences as we are not able to use any of their data for personalisation.

5.5.8. Forming Cohorts

Forming user cohorts based on demographic data or user behaviour was often discussed as a method of leveraging personal data that afforded a range of possibilities, like *abiding by legislation*, *delivering relevant content* and evaluating *the reach and response* to disseminated content.

5.5.8.1. To Abide by Legislation

The formation and use of user cohorts based on age was adopted to ensure alignment with legislative constraints regarding the different types of personal data that was collected from signed up users. Cohorts also help manage users' age transitions, which is

accompanied by changes in the jurisdiction with respect to their personal data use.

> *Justin : For children under 13 we do not collect full post code, we collect town, the town where they live. And, we also, for practical purposes today, if you are under 18 we collect your year of birth, sorry your date of birth so that we can work out when you transition from being under 13 to over 13 and when you transition from being under 18 to over 18, so that we can track those transitions.*

Cohorts also ensure that every user has, by default, available to him/ her the legally allowed grade and form of services for their age. These include mechanisms like the access to comment on platforms and ensuring parental regulation for under eighteens.

5.5.8.2. To Deliver Relevant Content

Cohorts help target content that is relevant to users. Targeting is enabled across diverse types of content, including television shows, online programmes, radio shows and even marketing emails that are served to audiences. Forming cohorts helps with the personalisation of these varied types of content to make more targeted decisions about the material to be served to a user based on his/ her membership in various cohorts.

> *Miles : So, personalisation comes in many types and forms. So, it's creating um, improved websites, it improves programming, it's giving recommendations to people what they might want to watch, depending on which cohort they sit in and even sending newsletters out to people. So, if I, for example, if I'm middle aged and if I really like nature programming and I like Radio 4, but I don't like Sports, why would you be sending me emails about sports or why would you be sending me emails about Radio 1 extra when you know that I'm a Radio 4 listener.*

Personal data also affords the service providers with the ability to form age based cohorts which could be used to ensure delivery of age appropriate content to the audiences. This helps the organisation to not just exercise appropriate delivery of content as per regulatory norms (as discussed in previous sections that discuss the legalities of personal data use) but extends its responsibility to cover social and ethical grounds by ensuring audiences are served with content that is appropriate for them.

5.5.8.3. To Measure Response and Reach

Cohorts also perform the function of evaluation of disseminated content. It is used "*to measure our conversion rate and measure how we succeeded with the making of better experiences[P5]*". It also

helps to highlight gaps in content consumption, especially helping uncover flaws that might be preventing delivery of service to specific sets of audiences that belong to particular demographic groups.

> *Mila : In terms of developing new services, understanding which bits of those new services are reaching audiences that we are not currently reaching. Or we are just building things that are super serving our existing audiences. Helping us understand if there are gender imbalances, if we are serving people in different parts of the world in different ways. You know, we are doing particularly well with women in the South of England but we are doing very poorly with women in India.*

5.5.9. User Empowerment

Dialogue around personal data renders it as a tool that enables improved interactions through user empowerment by helping *understand the user* (discussed previously in the Benefits section), *delivering content* that is relevant to them, *including the user in the creative processes* and exercising *responsibility for care of users*.

5.5.9.1. Deliver to the User

The opportunity to use the understanding about the user to help editorial teams to make decisions and deliver content that is relevant to the user, became a broader topic of discussion within the interviews.

> *Nate : I work very closely with the editorial team and they are interested in making sure you can tell what the audiences want, what sort of information they need and how we can change the content based on that information that we gather.*

Personal data contributes to various techniques that make content delivery more efficient and appropriate to the users. Understanding the users with the help of personal data, by combining content from differing temporal, geographic and contextual sources helps collate the most appropriate experiences for every user.

> *Simba : So, if we take full post code, in our digital products, we resolve that to things like what TV region are you in, so that on iPlayer we can show you the appropriate version of your TV. We also use it to determine your local radio station, we also use it to determine the weather forecast for your location. And, when we have any local news or local news alerts we can update you based on your location.*

Also the capability of being able to understand user context through personal data and deliver content that aligns with the same so that the user maximises their value and engagement with the service. This

would involve looking into the time of consumption, platform of use and other factors like previous viewing history and preferences to triangulate the most appropriate length and form of media content to be served to a particular user at a particular time. The following quote explains how service providers are hoping personal data to help inform them about the user context to make accurate decisions about the optimal kind of programmes (refills, trials) to be delivered on the user's platform of choice (TV versus online media platforms).

> *Minnie : So, if we know…. its basically helps to inform staff, so for example, this is the kind of stuff that can help tell us, you know, TV platforms for example, are becoming more and more, are becoming a bigger and bigger part of iPlayer usage. Thats then really important for us to say, ok, should we then be putting in more refills on TV, should we be trialling more stuff on TV, that kind of thing. The question about the infrequent users is really important for us because as a product we want to increase our reach overall, we want people to use the BBC more, so anything we can understand about whats making, some people stay longer and other people not come as often, is really useful, for informing those decisions, yea.*

With the increase in popularity of internet television like Netflix etc, the popularity of broadcast television has seen a drop (Bacon, 2016). One of the biggest challenges this shift has introduced is the need to stay relevant to younger audiences. The learnings gained from the use of personal data is expected to help maintain relevance over all demographic bands by ensuring delivery of relevant content to everyone.

5.5.9.2. Include the User in Crafting Future Media Content

Using personal data could help involve user priorities when making decisions regarding crafting future media experiences. By using personal data to shape media experiences, the scope of decision making essentially broadens from that of just editors and content commissioners to the users themselves.

> *Racheal : I'm actually quite interested in moving beyond that involving users with the algorithms, instead of us trying to guess what people are interested in. I'm more interested in them telling us. And that's what so much personal data and indicating preference and storing people's preferences and being able to go back and edit those and update them as they change.*

5.5.9.3. User centred consent models

Thus, personal data is considered as a "*research tool for the creative process[P5]*". It could help rethink the crafting of future experience design and development to help enhance experiences by involving the user in the media creation process itself, so that the content is more in line with the user wants and needs.

User consent is one of many legal bases for any technology to collect, store or use personal data (GDPR, 2016). Consent has been translated into design by allowing the users with the choice of opting in or out of personalisation. Also, to ensure user consent happens through user empowerment the need for the user to be knowledgeable about data practises was agreed upon. The provision of alternative models of layout for the organisation's privacy policy was considered a response to this scenario.

5.5.9.4. Responsibility for care of the User

The service providers' view of potential reputation damage as a result of collecting and using personal data and the need to have an increased sense of responsibility to help mitigate this situation was discussed previously (in the Risks Section : 5.5.3.). The interviewees further unpacked this notion by outlining scenarios where they have to be responsible to the users and looking into solutions for exercising this responsibility through practical accountability measures.

One such scenario outlined was the possibility of using personal data to understand user preferences and using that information to control and manipulate their behaviour. The need to take such scenarios into consideration and taking precautions to avoid them was discussed.

> *Patty : I think the longer term war is if all of this modelling, all of this data, I'm not talking about BBC I'm talking about general, will be used for reasons which are more associated with the government. You know….prediction and changing of the behaviour with communicating misfacts and putting people into specific mindsets. We have a, we could have potentially a self induced… you know…. Orwellian state run by one of the corporations. Thats obviously a dark scenario but it is possible in a way.*

The importance of protecting the users and their privacy through careful consideration of their data and its uses was also outlined. Particularly with respect to media experiences, the nature and type of data involved might be sensitive enough to cause social discomfort to the users, if not managed properly. This calls for accountability from the service providers' side.

Wyatt : And theres also, in the case of certain types of user generated content, like face like I said before, we do want to make sure that or we want to reduce the risks of harm to that person, so perhaps their face being misused in a certain way or you know, something like….something, somebody trying to do something defamatory or offensive involving that person's likeness.

One of the means for building accountability discussed in the interviews was the effective communication of data practices with the users. The importance of transparency as a value has been discussed previously. But for transparency to be effective, the communication has to be conducted through methods that are engaging to the users. The following is an explanation about the organisation pre-emptively designing the UX of a website that shares information about user data practises to be easily understandable to the users.

Charles : Yea thats right, we have…. we publish information at three different levels. In the UX, in the user experience, as you are entering your fields of why kind of question that you can drop down to a short summary. That makes through to a more detailed description, still written in English rather than legalese that describes what is the…. how do we use the data, why are we collecting it and that sort of thing. And then that works through to our privacy policy which is a kind of legal document. So, the three layered approach, one, in the UX, very simple, second, still aimed at consumers and… but slightly more detailed and then the third is the thing that very few will actually end up reading but it is the sort of legal document.

The idea of affording the users with choice over their data sharing was also discussed as a method of providing accountability to the users. Here, the concept of the entire data processing being driven by user consent alone is brought up by an interviewee as a method of provision of choice and consent to the user.

Parker : Everything is driven by user consent, so when you sign up for the BBC ID, you are signing up for what is mandatory as the policy. The policy clearly says we will be collecting this data, when there is an option for them to say I don't want you to collect certain types of data, so personalisation, as I said is a…is a feature that they have to… they can switch off, on their BBC ID.

5.6. Study Conclusions

The findings of the study present and elaborate a number of dimensions of service provider viewpoints with respect to data driven media experiences, ranging from the purpose of collection of personal

data, the benefits and risks involved, the future of this process and other topics that evolved from the conversations like business priorities, personalisation, forming cohorts and user empowerment.

The following section reflects upon these findings to present the conclusions of this study, which is the overarching need for the value created by personal data in novel media experiences to be constantly balanced with the challenge of loss of user trust. It then collates the findings to catalogue a list of challenges that contribute to this scenario, as elicited by the service providers, rooted in their everyday practicalities of serving data driven media experiences.

5.6.1. Value

The notion of value is one that occurs repeatedly in the conversations around personal data. Personal data provides value to both the media service providers and the users in multifarious ways which makes its use within future media services an obvious choice.

To the service providers, value from personal data manifests in the following ways :

- *Better Business Performance* : Personal data helps the business mandate of media service providers in two ways. The first is to identify consumption gaps and trends, spot underserved audiences and adopt measures to respond to these flaws to provide content of value to all audiences equally. The second is by helping with product development by improving upon existing products and enabling creation of future products by serving the creative and editorial teams with actual preferences, interests and expectations of the audiences.

- *Better Audience Satisfaction* : Personal data helps the service providers in the delivery of content that is more aligned with audience expectations, thus increasing user engagement and satisfaction. Firstly, personal data helps form user cohorts based on variables like age, location etc. that helps in understanding the user response better. Such response is uncovered through trends and patterns in consumption and this data is used to check on and respond to audience satisfaction. Secondly, personal data helps the possibility of triangulation of various variables (eg., location, preferences etc.) to possibly understand user intention behind each visit. Using variables like which platform (eg., TV, online, radio), what time of the day, location and previous preferences, the service provider expects to be able to make reasonably accurate predictions about the kind of content that is most suited to that visit (Eg., short form, long form, refills, trials, VOD, live, content that was paused midway in a previous viewing, the new episode of a series that is being followed).

- *Remain Salient* : In the rapidly changing landscape of media creation and delivery, personal data gives media service providers the significant advantage of staying relevant in the audiences everyday lives. It helps understand the audiences at a much more granular level, helping create future experiences that reflect these learnings and thus match audience expectations. With the rest of the world turning towards personal data at a rampant rate, user expectations around services are much more centered in the affordances provided by personal data use. Here, to cater to user expectations and to stay on par with competition, personal data becomes a key resource. Personal data, along with helping match audience expectations also helps by constantly improving the flaws in current services, making them more relevant to the audiences and also enabling enhancement of future media to reflect these learnings thus making them a salient part of the audiences' lives.

- *Reach and Impact Feedback* : Personal data presents a number of benefits and opportunities because of its potential to feedback the reach and impact of the media services delivered. While abstracted statistics around media consumption was available even through traditional television sets, with the introduction of smart TV's, online media platforms, smart phone applications and digital radio, personal data extends this opportunity to include more granular feedback metrics that make responding to issues, overcoming flaws and enhancement of services more efficient and effective.

- *Relationship with the Audiences* : Personal data also shifts the very relationship of the media service provider with the audiences. The traditional one-to-many broadcast model is replaced by a personalised one-to-one relationship where every user's priorities are considered and wants catered to. The inclusion of personal data in the processes of understanding the users, matching their expectations and responding to flaws to create future experiences that are relevant to all audience members shifts the focus of media production to the users. This makes the process more user centered and alters the very nature of the relationship between the media service providers and audiences by giving the users increased input, involvement and a more central role in the very creation of future media experiences that they consume.

To the audiences, value from personal data within media experiences manifests in the following ways :

- *New Forms of Content* : Personal data introduces the possibility of enhancing media experiences either by building on top of current experiences or by creating new genres of experiences.

Firstly, personal data allows the introduction of specific experiences that are tailored to reflect a user's context and preference, like an educational service reflecting the curriculum of a particular nation and remembering the progress of the user. Secondly, personal data helps build richer experiences by adding in layers that are relatable to the users, thus encouraging more engagement and user participation, like a show involving photographs or other personal details about the user in its' content. Thirdly, personal data allows for the creation of entirely new genres of experiences like timelines, memories, stories and maps that collate audience history to present a new media experience, unlike any witnessed before.

- *Tailored Experiences* : Personal data presents significant opportunity when it comes to personalisation of media. Here personalisation of a media experience could begin right from the formative phases of content discovery with recommendations and playlists being curated to present the users with relevant content. Such tailoring then extends to user journeys being personalised to present the most relevant content and engagement strategies out of the vast breadth of media that is produced every single day. Personalisation is also applied on the content of the media recently, wherein the content itself speaks to the audiences by being more relatable and relevant to them.

- *Appropriate Content* : Personal data also enables the possibility of delivering appropriate media content to audiences. This is particularly relevant in households with children, where shared devices are used for media consumption. It is also helpful in negotiating social situations when more than one person is present, consuming content and the relevance and appropriateness of the content to all parties involved should be considered. Here, using personal data to determine the right kind of content to be delivered, thus alleviating potentially socially awkward situations or avoiding situations of delivery of inappropriate content to children is a significant possibility of value.

- *More Engagement in the Future of Media* : Personal data, as discussed previously, alters the relationship between the service provider and the audience members. The audience members are no longer just passive receivers of content disseminated to them, but, through their data, they are able to shape the very experiences that they are offered, thus playing a key role in the future of media creation. The audiences will be able to express their preferences and interests through their data and expect enhancements and responses to previously ineffective aspects of their media experience, once they start marrying their personal data with their media.

5.6.2. Trust

While personal data helps create value to both the service provider and the user (as listed in the previous Section), it also simultaneously increases the risk of loss of user trust, emphasising the need for service providers to constantly ensure preservation of trusted relationships with their users during this process.

Firstly, for the BBC, trust is one of the core organisational values. Thus, prioritising user trust and ensuring placement of proper safeguards for the same throughout the process of leveraging user personal data was discussed often as a concern in the shift towards personal data use.

Secondly, owing to the key placement of trust as a value driving the institution, the possible risk of reputation damage introduced by the use of personal data was often associated with this loss of user trust. Here, the importance of mitigating such circumstances through diligent management of personal data to preserve user trust in the organization was leaned upon.

Thirdly, one of the key methods of response to such issues around trust was to ensure user privacy and not to seem 'creepy' through this shift. Instead, to focus upon transparency and effective communication of the data practices to the users through novel and efficient methods.

Finally, the risk of loss of trust, while mentioned explicitly within the risks section, also forms an umbrella that encompasses most of the other risks discussed (ranging from 'intangible' risks of making out-of-context decisions to security compromises). Any of these other risks, if managed well could contribute positively to user trust in the organisation and if not, could easily lead to loss of user trust.

While the instances where trust was explicitly addressed has just been discussed, there were also other concerns mentioned by the interviewees, often interleaved within other topics, yet exposing the capacity for the concern to lead to potential loss of user trust. For example, the findings reveal access to personal data has led to the demand for more data and an increase in the resolution of potential data in the future. When data collection becomes increasingly granular and detailed, the dangers could start outweighing the benefits if the process and its rationale are not appropriately justified to the users. The ability to cohort users while advantageous in many ways, also introduces the risk of bad stereotyping leading to lowered user satisfaction and trust in the organisation. Also, the need to stay relevant by keeping up with competition demands constant recalibration of service provider priorities to maintain focus on user needs to ensure that interactions which compromise user trust are not adopted at any stage of the process.

5.6.3. Need for balance

The study uncovers an increasing demand for personal data to provide value to modern media experiences (through supporting business priorities, providing increased service affordances through personalisation, understanding the users better through cohort formation etc.,), which is challenged by the potential for loss of user trust (which further unpacks to explicate the other discussed risks like reputation damage, making out-of-context decisions about the users, constraining and biasing user attitudes etc.). Thus, there arises the need for balancing of two conflicting phenomena that originate from the same source : personal data use, in order to ensure sustainable use and growth of the same.

While the need for balance is uniform across the various teams interviewed in this study, the position of the pivot of this balance is often variant, depending upon the nature of the team and the particular media service they provide. On one side there are the data maximalists who ask for maximum amount of data to ensure that their products are on par with competition, this need is particularly stressed upon when their competition includes international players like Netflix and Amazon, who leverage massive amounts of personal data to produce media experience enhancements. In the middle resides those who appreciate the need for personal data when considering substantial innovation in future media but are reluctant to make this move owing to the potential loss of user trust and the other risks that add up to it. And on the other end of the spectrum are the data minimalists who, while acknowledging the key role of personal data use in future media enhancement, challenge this very shift owing to its inherent risks. If personal data is to be leveraged in this instance, they ask for alternate approaches to the data processing (eg., client side processing, public service approach to data) that minimize the data collection and the risk of losing user trust.

Thus, whilst the shift towards use of personal data in novel media services is seen as a necessity and an enabler to provide audiences with increased value, this need demands to be balanced with the potential risk of loss of user trust introduced by the same. The rest of this discussion presents considerations shared by service provider interviewees, which contribute to the maintenance of this balance, while using personal data to drive media experiences.

5.6.4. Contributors to the Balance

The results of this study presents a number of contributing considerations that could help maintain or disrupt the aforementioned balance when media experiences turn towards user personal data. These considerations could be catalogued in the way they approach the issue in hand and hence be classified as either legal accountability, social accountability or ethical considerations. Irrespective of which of the three approaches are chosen, it was

agreed upon that technological responses often formed a viable method for responding to the situation.

5.6.4.1. Legal Accountability

Legal requirements are a prominent sculpting force while designing and delivering personal data driven technologies as they enforce accountability into the scheme in a mandatory manner. The then impending GDPR combined with the affordance it provided service providers to help ensure improved legal accountability to the users, made the provision of legal requirements seem a primary response measure to preserving user trust.

> *"With the younger audiences, we are allowed to keep the data in order to be able to say that so many people do this, that or the other, but we are not allowed to keep that data attributed to one person, that says because you watched this, you can then watch that, because that becomes marketing, we are not supposed to be using that data for in that sense. So, there are all sorts of legal frameworks that we are trying to sort of... to work within.[P18]"*

Ensuring the Right Grade of Personalisation

In the U.K., legislation dictates stark distinction between the grade of personal data collection and personalisation served to users under the age of eighteen and over eighteen. As stated in the previous quotation, abiding by these 'legal frameworks' helps the service providers respect and cater to users from different age groups appropriately, through provision of legal accountability. But, this provision becomes complicated with the introduction of shared entities like media devices, data sources, social spaces and settings which demand more attention to ensuring delivery of appropriate content to all the users involved. Responding to such socially inspired legal challenges posed by personal data collection would help the organisation build up it's legal accountability and through it, user trust in their services.

Leveraging Age Based User Cohorts

The use of cohorts and segmentation of users were seen as a potential method of approaching legal accountability concerns around personal data use. This becomes particularly effective in the case of determining the right grade of personalization for a user. Here, age based cohorts could be used for regulating such scenarios by automatically managing the collection of the right type and amount of personal data, delivering the right grade of content and also managing age transitions between cohorts easily and with least effort for the users.

Realising the GDPR by Exercising Transparency

Reflected in the findings were also the measures required for the realisation of the GDPR. Providing transparency about data practises was identified as a method for translating legislation into design. Currently, transparency is enabled through introduction of new websites that present information to users in simple and engaging ways, a shift in the format of the terms and conditions statement and the general use of simple language when discussing data practises. Looking into alternative methods of transparency provision and other methods that help practically realise the GDPR will contribute towards the establishment and maintenance of trusted data flows between media service providers and audiences.

5.6.4.2. Social Accountability

The turn towards personal data in new media experiences highlights several challenges that places the call for better social accountability. Social accountability, while not regulated like in the legal scenario, is considered a requirement to help build user trust in personal data environments. It refers to being accountable as a societal entity, respecting and mitigating the social implications of using personal data that could lead to loss of user trust.

> *"And I am not sure how comfortable they would be in answering what is your ethnic background or your religious beliefs and things like that. I would like for us to deliver value for everyone but I'm not so sure that I would like to ask for that information. So, it's a rather tricky situation…[…]….to put it bluntly, it would be a bit creepy to ask for that kind of information.[P12]"*

The quotation clearly represents the dilemma of whether to ask the users for data that might make them uncomfortable in their offline social lives versus the need to know the same information to ensure improved delivery of services. It clearly represents the kind of predicaments faced by media service providers when dealing with issues of social accountability, that might lead to loss of user trust.

Effective Management of User Cohorts

One such challenge identified within the findings is the formation of user cohorts and deciding how to make decisions based on a user's cohort membership. When considering user cohorts, personal data can act as a tool that enables understanding the audiences in many ways that help accurate delivery of appropriate content to users. However, knowing the balance of when to utilise information about user membership in a particular service-created cohort versus when to respect user individuality that sometimes precedes cohort patterns, is a social challenge which requires further attention.

Responding to 'Intangible Risks'

Also to be considered here are the 'intangible risks' of making out-of-context decisions that could lead to user discomfort in social scenarios. The findings discusses this issue with the example of a sports fan needing to keep their preference of teams private, if it goes against the regular social choice. When media shifts to knowing such information about the users, it is considered vital for the service providers to be responsible with the information so as not to create social scenarios of tension for the users, instead using it to alleviate said social tensions.

Being Socially Accountable as an Organisation

Another risk that is closely associated to social accountability is the possibility of reputation damage, which is closely associated with the potential for loss of user trust from the use of personal data. Here, the organisation as a social entity believes the need to preserve their reputation within the society helps them to be seen as a more accountable entity socially, contributing positively to the trust users place on them. This requires the service provider taking active responsibility for the data practices and responding to the challenges presented here in a timely and effective manner. The study exposed a number of ways of responding to this concern. These include active measures to spread awareness both within and outside of the organisation and the various proposals for being responsible to the user through proper communication of data practices, provision of choice and control and the shift to client side processing which involves minimum collection of personal data by the service provider.

5.6.4.3. Ethical Considerations

Ethical considerations are responses to both legal and social accountability challenges. While not dictated, or regulated by any specific entity in the media domain, ethical data practices have been recognised by service providers as a key necessity in building user trust in personal data ecosystems.

> *"I think that you can do it the right way and the wrong way. For me, I think that if you are asking something, if you are asking people for more information about who they are, or what they think about stuff, then you need to give something back, or you need to justify why you are asking for it.[P13]"*

As the quote states, the necessity to justify the data collection and use to the audiences is an ethical measure that helps both the service provider and the users. For the users, it helps them understand why they are asked to share their data and the reasoning behind the exchange. For the service providers, it helps them justify the underlying data collection to both their customers and to themselves

and through the transparency they provide, they are able to work on the trust users place on them as an organisation.

Ensuring Bias-Free Experiences

Delivery of bias free experiences is a key ethical challenge that requires attention within media experiences. Within the scope of personal data leverage, two dimensions of bias was talked about. The first is the importance of non-partisan reporting when it comes to News media, which is highlighted when the service provider is empowered with personal information about every single user. Here, the importance of not exploiting such information to bias user attitudes but rather utilizing it only for ethical purposes such as enhancing the media experience is stressed upon. The same emotion is also extended to the second manifestation of constrain and bias that result from the use of personal data. These include echo chambers and filter bubbles, that could limit the users' reach and diversity in consumption, making them question the value of the underlying data use.

Delivering Appropriate Content

While age appropriateness of broadcast is often legally regulated, with the expansion of media moving beyond just broadcast coupled with the previously discussed involvement of shared media consumption settings, the ethical need to regulate the appropriateness of the content served in these social scenarios also becomes ethical considerations that demand further thought.

Maximising the Potential of the Data Collected

Personal data is currently considered by some as an asset class (Schwab *et al.*, 2011) on the same lines as oil and gold, which calls for a 'symmetric' exchange of value for this asset that is collected from the users. And also, in contrast to these fungible assets, personal data is highly contextual and thus raises the creative challenge of ethically utilising it to its fullest potential for these diverse contexts. The ethical challenge then is to innovate as a service provider and provide services that are on par with the competition and that maximise the value returned to *all* the users while not compromising on user trust. This maximization of value begins internally through training of the employees who work with personal data and are responsible for converting it into value through innovation. It also extends to possibilities of diversification of services provided to include feedback mechanisms like the Quantified Self initiative (Lupton, 2016) that would help the audiences gain more insight and value from their data.

Spreading Awareness

A viable ethical consideration that could help with the risk of loss of user trust discussed in the study was to actively initiate spreading awareness both within and outside the organisation to ensure proper knowledge of the data practices that enable innovation and the risks they pose to all the stakeholders involved. Also mentioned was the potential adoption of a public service approach to data that focuses on 'back to the people'. This would include looking into mechanisms that make the data sharing more 'open' to the users by providing extra value to them through alternative services and feedback that helps the users understand the benefits of sharing their data. This also included recommendations of exposing the underlying design patterns to the users to help them better understand the data interactions that support their experience.

5.6.4.4. Technological Reasonings

Along with the legal, social and ethical considerations the service providers also presented technological reasonings about the use of personal data in media experiences. Here, technology becomes a medium that embodies solutions which help maintain the 'need for balance', either by directly contributing to user trust or value provided, or indirectly, by helping respond to the various legal, social and ethical considerations discussed.

Building User Trust through Technological Affordances

One primary set of technological reasonings often focused upon was *data security* where techniques such as anonymisation, data minimisation, encryption etc was suggested as methods to ensure protection of user data, preservation of the organisation's reputation damage and through it, user trust. Provision of enhanced data security and effective communication of the same to the audiences is expected to result in improved data interactions that are backed by user trust.

Using *user personal data to understand audiences* to help craft future content was discussed previously. Thus, the user is now a participant in the creative, editorial and commissioning processes, making future media experience design and dissemination procedures more user centric. Focus on initiatives that shift the onus to actual user preferences help address the concern of bias by involving user behavioural and consumption data in the crafting process. This shift is expected to increase the relevance of services to all the users involved, ensuring service of value to all users and thereby building user trust.

Using Technology to Respond to Legal Considerations

Transparency was highlighted as one of the legal requirements of the GDPR that is hoped to result in increased user trust. But exercising

said transparency effectively so as to empower the users is a process that demands much attention and focus. Starting from the terms and conditions statements, extending to all parts of the media experience that is influenced by user personal data, there is a need to effectively expose underlying data processes to the users to help them understand their role in the service being provided. This puts forth the call to encourage the design and inclusion of data interactions that lead to increased awareness and understanding of the underlying data practises while creating and serving future data driven media.

Using Technology to Respond to Social Considerations

Use of cohorts is a technological response to concerns like the need to manage legal requirements, evaluate content consumption, deliver relevant content and identify and cater to underserved audiences. But here, the use of cohorts also presents the challenge of recognising the difference between contexts that require cohort use and those where cohorts are dysfunctional. This calls for careful study of the socio technical implications of cohorts and identifying scenarios where they disempower versus empower the users. Thus, future design of media experiences that leverage user cohort data have to be sensitive to these context specificities by developing interaction strategies that overcome erroneous stereotyping that could result from extensive user segmentation.

Using Technology to Respond to Ethical Considerations of 'Symmetrical' Value Exchange

Providing meaningful feedback to the users through alternative strategies for value creation like self-reflection summaries, quantified-self initiatives or visualisations that demonstrate the user's role in the digital economy have been identified as alternatives for responding to the concerns of provision of symmetrical value to the users. Diversifying the value delivered to the users by maximising the potential of the collected data through creative initiatives that benefit the users are technological responses to ethical dilemmas that demand further justification for the value returned for the data collected.

Also, switching to *user centered models of data processing* that looks into user empowerment through increased user awareness and engagement in the digital economy is expected to shift the dynamics of the balance to engender a more trusted data use environment for the users while simultaneously allowing for personal data leverage. Other methods of user empowerment within the digital economy included the possibility of adopting *alternative* technological solutions where the user is treated with more ethical consideration and capacity for involvement with the data practices. These include more ethical ways of managing personal data like a public service approach to data and the use of client side processing without transfer of data to organizational servers.

5.7. Study II Conclusions Summary

The following account summarises the challenges of personal data use within media experiences from the viewpoint of the service providers, as revealed by this study.

The primary motivation for media experiences to turn towards user personal data was the multiple ways in which it manifested **value** to both the service providers and the users. But while personal data presents so many avenues for value creation, it also simultaneously introduces the challenge of **loss of user trust** which puts forth the call **for balancing the value created by personal data with the loss of user trust** it might result in.

This balance is unpacked to present **legal**, **social** and **ethical** considerations put forth by the service providers, which are all expected to be responded to through technological interventions that realise them.

Legal accountability provision which often forms a regulated mandate in these contexts, were highlighted as measures to contribute to maintaining this balance by alleviating the challenge of loss of user trust. Such legal accountability challenges include :

- Ensuring the right grade of personalization.
- Leveraging age based cohorts to deliver the right grade of personalisation for each user.
- Realising GDPR requirements like transparency.

Social accountability provision extends beyond what is mandatory to include extra measures that help build user trust in media experiences that use personal data. Examples of such responses are :

- Effective management of user cohorts that do not stereotype users based on age, gender or location.
- Responding to ‘intangible risks’ like the data resulting in out-of-context-decisions that lead to user discomfort in social situations.
- Being socially accountable as an organisation by spreading awareness about the data practices.
- Looking into solutions that give more choice and control of the personal data to the users.

Ethical considerations respond to both legal and social accountability concerns. They involve :

- Ensuring bias free media experiences by both provision of non-partisan reporting and through elimination of filter bubbles and echo chambers.

- Ensuring delivery of appropriate content particularly when children are involved.

- Maximising the potential of the personal data collected by looking into diversifying the value returned to the users.

Technology was seen as a medium that was expected to help respond to all of these considerations.

- It could provide *legal accountability* through provision of transparency right from the design of the Terms and Conditions statements to the content of the experience itself.

- It could help respond to *social accountability* challenges through diligent management of user cohorts that could result in accurate decisions and personalisation that abstracts away the possibility of situations of social discomfort.

- It also helps realise *ethical considerations* by allowing for alternative means of value delivery through self-reflection summaries, quantified-self initiatives or visualisations that demonstrate the user's role in the digital economy, to maximise the value of the personal data collected from the users.

Limitations of Methods Chosen : This study was conducted solely in association with the BBC. The study included a range of teams from within the BBC (News, Sports, iPlayer, Childrens', R&D, Marketing and Audiences, Radio etc.) and all of them serve very different formats and kinds of media content, work and manage projects in very varied manners and prioritises different aspects of the service provision and data usage accordingly. Thus, while they could be considered almost as a number of 'different' media organisations, they are all associated with a single organization, which could be considered a limitation of the study. This could be overcome through future work that extends this research by looking into other external media organisations who would be undergoing similar shifts to leveraging user personal data and study how the priorities uncovered here overlap/ vary across them.

6. Study III

Having explored both the user viewpoints and service provider viewpoints concerning the shift of novel media services to user personal data, the next step was to study user and service provider response to an actual media experience that leverages user personal data. This intervention was designed to include a '**data dialogue**' that **embodied responses to the challenges previously uncovered.** This 'data dialogue' was used as a mechanism to **provoke the respondents to elicit their response to their data being collected by media experiences and the inclusion of such measures as the 'data dialogue' by engaging in further conversations around personal data leverage by media experiences.**

The following chapter details how these priorities are achieved by bridging the learnings of the previous studies. The result is the co-design, deployment and evaluation of a novel, physically immersive, data driven media experience (the Living Room of the Future, LRoTF) that is tailored to the audiences while simultaneously being sensitive to the use of user data within the experience. Reflecting the nature of the previous studies, the chapter presents the viewpoints of both the users and the service providers derived from conversations that focused particularly around the 'data dialogue' the experience offered.

The results of the deployment show that audiences found the LRoTF highly immersive, engaging and 'magical' and responded positively to the inclusion of the 'data dialogue within the LRoTF which contributed towards alleviation of certain challenges previously identified. However, audience members were clear about the demarcation between enjoying a short-lived novel media experience within a museum installation versus the challenges associated with adopting such technologies into their everyday lives. This feedback laid out a range of subtleties that added further nuance to the understanding of the challenges in this scenario. These results are voiced as a set of specific user reasonings around immersive media experiences that leverage user personal data and broader concerns associated with media experiences using personal data. The responses of the service providers in this study focused on the importance of data legibility and control within data driven media experiences and the balance of provision of value and loss of user trust. These topics were unpacked within the scope of the LRoTF and media experiences that leverage personal data specifically. These findings present the service provider practicalities associated with such challenges and their future expectations around the inclusion of responses sensitive to personal data leverage like the 'data dialogue'.

6.1. The Living Room of the Future

This chapter presents the co-design and evaluation of the Living Room of the Future (LRoTF), a provocative media experience that leverages user personal data while simultaneously exposing the data practices that shape the experience customization through a 'data dialogue'.

This design fiction experience bridges the learnings of the previous studies to design a 'data dialogue' with the users, that seeks to respond to the challenges previously identified and provoke the users to elicit the use of personal data by media experiences and thier response to the inclusion of such measures as the 'data dialogue' to engage in deeper dialogue around the challenges of using personal data in media experiences. This experience is also used to probe service provider response to the LRoTF as a media experience that leverages user personal data, the inclusion of the 'data dialogue' in it and their take on the future of such responses within media experiences that leverage user personal data.

The two previous studies discussed media experiences that leverage user personal data as scenarios of the future to identify challenges in this shift from the user and service provider perspectives. This study bridges these findings to produce a media experience that leverages user personal data and exposes the underlying data practices through the 'data dialogue' in a provocative manner. Thus, the results of this study elicit **responses to the actual experience of media leveraging personal data** which includes feedback on the 'data dialogue' interventions through reasonings around the challenges of using personal data in media experiences that are more nuanced than (or rather help further unpack) the findings of the first two studies.

The Living Room of the Future (LRoTF) is a physically immersive media installation, in which media content adapts in response to user data collected from IoT devices situated in the living room. The adaptation, in turn, actuates connected devices (e.g., speakers, lights, window blinds, heating, etc.) in order to enable the media experience to reach beyond the screen into the living room to create a physically immersive experience.

6.2 A Multi-Partner Project

The LRoTF was a multi-partner collaborative project (funded by AHRC project Objects of Immersion [AH/R008728/1]) which partnered the University of Nottingham, the University of Lancaster, the University of York, BBC, FACT and the British Council. Here, while the design of the holistic project was supported by all parties, every partner pioneered certain aspects of the experience and its deployment, which is listed below.

The BBC along with the University of York was interested in and hence worked on the Object Based Media (Shotton *et al.*, 2014; Evans *et al.*, 2016) aspect of the technology that helped deliver a movie experience that could be adapted in real-time, with respect to the audiences interactions.

The University of Lancaster worked predominantly on the design and deployment of the LRoTF space itself where they focused on the sensors, actuators and their responses within the movie and the room.

The British Council enabled the possibility of running a workshop as part of the Sarajevo Unlimited Conference (*Sarajevo Unlimited*, 2019) in Sarajevo, Bosnia and Herzegovina in 2017. This brought together media directors, producers, journalists and technologists to think about possibilities for future media experiences that marries IoT technologies to the media itself. The participants of the workshop were encouraged to apply to a two week placement in the U.K. where they produce the assets for the OBM to be used for the LRoTF movie. The winners of the competition, a technologist and a film maker from the Western Balkans, worked with the rest of the team to help shape the media aspect of the experience.

FACT focused on the media content that formed the movie experience within the LRoTF. They worked in close relationship with the competition winners who created the media content for the film, orienting them more deeply to the concept behind OBM and helping them shape content that is suitable for an adaptable experience. In addition, FACT also provided space in their theatre facility in Liverpool, United Kingdom to host the LRoTF deployment.

This thesis focuses on the input from the University of Nottingham, who led the data management aspect of the project. This includes the design and inclusion of the 'data dialogue' that exposes the underlying data practices of the media experience and the design and execution of the evaluation study of the research reported in the following account.

The LRoTF was chosen was as an instrument for investigation in Study III for the following reasons. Firstly, it provided the opportunity to design responses to the challenges uncovered in the previous study through the potential inclusion of the DataBox (Amar, Haddadi and Mortier, 2016) and a 'data dialogue' within a near-future yet tangible media experience. Secondly, the method of Design Fiction used within this experiment allowed to immerse the audiences in a world where media experiences actively leveraged user personal data while pronouncing this use of personal data through the data dialogue employed, which were conducive to the motives of this research. Thirdly, this multi-partner initiative, which also included the BBC allowed for the study of the co-design and response to this data driven media experience to not just be conducted with the users but also the service providers, thus providing more cohesiveness,

completeness and comprehensiveness with respect to the multi-stakeholder viewpoints approach adopted by this thesis.

6.3. Design

The LRoTF forms the confluence of Object Based Media research (Shotton *et al.*, 2014), Internet of Things in the home and the application of Human Data Interaction (Mortier *et al.*, 2014) to create a novel immersive data driven media experience which is sensitive to the socio-technical challenges of the use of personal data (identified in the previous studies reported in this thesis). It does this by creating an 'embodied' Design Fiction (Coulton, Lindley and Cooper, 2018) that incorporates real and fictional technologies in a familiar but futuristic environment enabling new physically immersive media experiences.

6.3.1. Key Design Elements

Key design dimensions of the LRoTF include the **physical space**, the **media component**, **connected objects** and the '**data dialogue**'. These aspects were worked upon by the various stakeholders of the project (the partners and their contributions are detailed in Section 6.2) in a continuous, collaborative and iterative fashion. Here, while all key design decisions were discussed and decided upon collaboratively, the design itself of the various aspects were done individually by the various persons involved, culminating in a final stage of combining these various aspects to produce the **holistic LRoTF experience**. Details of each of these design dimensions are provided below (these details provided are limited to support the scope and argument of this thesis).

6.3.1.1. The LRoTF experience

The LRoTF experience was designed to last between 5 and 10 minutes, depending on the audience's interactions with connected objects. When an audience member (or a group) enters the room, the experience begins as soon as one of them sits on the sofas provided. The pressure sensors in the cushions trigger the movie to begin, the lights to adjust and the window blinds to go down automatically. This marks the beginning of the first part of the experience where the audiences are not explicitly told about the connected objects and their data connections beforehand, thus allowing for their exploration within the experience. The movie here, is continuously adapted to work consistently with actuators in the fan, lights and clock radio throughout, to reflect the unique interactions showcased by the group of people present in the room. The 'voice of the living room', which comes from the clock radio, introduces audience members to the LRoTF. During the film, the lights adjust colour and brightness to ambiently match what is on the screen. When the wind audibly blows through trees or a girl's hair in the film, the fan within the room blows warm air on the audience. The data monitor on the coffee table would

also be simultaneously displaying the data collected as a result of these interactions within the experience. Thus, every time an audience member engages with any smart object, the data collected is displayed on the monitor. According to the 'serendipitous' set of interactions displayed by the people in the room, the movie on the television is also simultaneously adapted (being edited in real-time) to create a resulting 'cut' that is unique to the interactions engaged in by the audiences. For eg., if an audience member picks up a book about the particular building in Sarajevo in which the film was partially shot from the magazine rack, extra footage is added to the experience; picking up a travel card adds more film of an underground 'tubeway' train journey; picking up a drinks bottle adds more shots involving rain; and picking up a TV remote control triggers the voice of the living room to say 'Do you think the living room of the future can be controlled with a remote?' Once the first part of the movie is run, there is the UV reveal which exposes the data connections and gives the audiences a second, shorter run of the movie wherein they can again engage with the connected objects, only this time they are informed of the data connections from the UV reveal and the information and their experience from the previous turn. Upon culmination of the second session, the data receipt is printed from under the coffee table, the lights come back on and the blinds go back up, cueing the audiences on the end of the experience.

Figure 8. The Living Room of The Future.

6.3.1.2. The LRoTF Space

The living room represents a readily identifiable space in which media is consumed in many homes. It is of course not the only space in which media is consumed in the home, but it lends itself well to 'world building' (Coulton, Lindley and Cooper, 2018) or the crafting of an embodied experience that may be situated in public spaces to facilitate broad public engagement. To support deployment, the

LRoTF is intentionally small (approx. 3.5m2). Nonetheless it houses components one might find in a typical living room: a sofa, coffee table, side tables, lamps, a clock and a TV.

6.3.1.3. The Media Component

Media for the LRoTF was developed by two successful candidates selected from a workshop held as part of the British Council's PlayUK programme at the regional innovation forum Sarajevo Unlimited (*Sarajevo Unlimited*, 2019). Details of the workshop and subsequent media development are beyond the scope of this thesis, except to say that the final output consisted of a set of 'object based media' (Shotton *et al.*, 2014; Evans *et al.*, 2016). OBM allows media content to be 'tailored' to the viewer's circumstances, as it uses 'objects', which are essentially media assets that can be arranged and re-arranged according to user context data to deliver every audience member a unique media consumption experience (Churnside, 2013; Shotton *et al.*, 2014; Evans *et al.*, 2016). The LRoTF marries OBM to the IoT via the Databox platform to dynamically contextualise media objects using data generated by viewers' interactions with connected objects located in the living room, while all data processing is run on the DataBox. (Section 6.3.1.5 Page 128)

6.3.1.4. Connected Objects

The LRoTF exploited a mix of off-the-shelf and bespoke connected objects which include connected lamps, a fan, a window blind controller, and a bespoke clock radio that acts as an audio speaker to play ambient sounds. Alongside these objects, which can be considered as 'outputs' in that they are controlled by the media objects, there were also a number of data 'inputs' that directly affected the media content. These consisted of pressure sensors and NFC readers. The pressure sensors were embedded within the sofa to detect when the audience sat down and NFC readers were also integrated into the tables to detect when drinks and magazines were interacted with.

6.3.1.5. Data Dialogue

A 'data dialogue' was designed as part of the LRoTF experience, to make it sensitive to the personal data collection underlying the media experience customisation. The aim of this data dialogue was to bridge the learnings of the previous findings to include responses that would help alleviate challenges uncovered in those studies and to use it to provoke the users to elicit more in-depth responses to media experiences leveraging user personal data, through a first hand experience that exposed the data practices.

Human Data Interaction

The design of the 'data dialogue' within the LRoTF was inspired by Human Data Interaction, a novel line of thought that proposes responses to the challenges of personal data use which resides in the highly complex intersection of multiple disciples like 'computer science, statistics, sociology, psychology and behavioural economics', through three dominant principles.

The first of these principles is legibility which states that the user should, at all times be in complete knowledge of what data is being shared, with who and for what purpose. Legibility deals with responses like increasing awareness of data collection (e.g., European legislation requiring the use of website cookies to be made visible) and presents further possibilities of increasing user awareness around the data processes like sharing the implications of the data to the users. When taking a human-centred approach towards data it could be said that different individuals have different priorities and insecurities associated with the data which could vary with time, context etc. It is necessary to show the potential uses and inferences that data could result in in these instances, so that the user is always empowered with awareness.

The second principle of HDI is agency, which states that the user should have control over their personal data at all times and in all contexts. This suggests a step up from the current trend of giving informed consent (though sometimes even that is not made possible (Ioannidis, 2013)) to move towards user centered controls. And this control is not just over consent but extends to the data itself, thus even enabling revocation of collected personal data (Whitley et al., 2009). Enhanced control will not only improve trust and privacy but will also help eliminate biases in the data collection process which could be because of semantic misunderstandings, sampling or temporal biases or contextual dependencies.

The third core principle of HDI is negotiability, which states that the user should be able to choose what data to share in exchange for services, at any level of granularity and in any context. This is the key bridge which leads the user from having good legibility to exercising agency thereby bringing in symmetry of power into the system. It is claimed that self-management of personal data by individual users might be too far fetched a dream due to the enormous amounts of data being collected (Solove, 2013). Hence, negotiability should be made practical through a combination of high level supportive legal and regulatory frameworks which would help abstract the need for managing large amounts of data and technological solutions which would help give granular individual negotiation techniques for data that might be considered sensitive in particular contextual or cultural settings.

By providing users with legibility, agency and negotiability over their personal data, data driven technology design is expected to take a turn in the direction of being ethical, thereby providing more support to the users of the digital economy, in turn alleviating their declining trust in the system.

DataBox : An instantiation of HDI

The DataBox is a personal networked 'edge' device (and associated services) that collates and mediates access to personal data (Amar, Haddadi and Mortier, 2016). The DataBox would exist as a physical object in the home which enables the users with the possibilities of physical interactions with it. It is a solution which could help data to be collected, combined and processed in ways that would let the users explore workflows and manage their own data and data involving other stake-holders (through ad hoc social interaction).

DataBox enables client-side processing of user personal data, thereby ensuring that all data stays on the box, within user homes, where it would be processed by one or many service provider applications that run on the box (much like a smart phone with its multiple applications). These DataBox applications would provide the users with SLAs that allow users to see and understand what types of data they would be exchanging, provide space for negotiation of exchange of this data and allow for explicit control over this data as well. By providing legibility and negotiability of data through the SLAs and continued control over the personal data that is managed client-side, the DataBox becomes a technological instantiation of the principles of HDI.

Within the LRoTF, the DataBox was used to enable client-side processing of all personal data collected. This data, after being used for experience customisation was deleted from the system in a bid to obviate the privacy and trust concerns that attach to the collection and use of personal data in the home. The inherent nature of the DataBox use within the LRoTF also abstracts away the need for explicit user control over the exchange of personal data as the data is not sent to any external servers or used post the experience. While the technical intricacies of making the LRoTF run on the DataBox is not part of the work of this report, the inclusion of the DataBox and the implications of application of its underlying principles within the LRoTF were leveraged to contribute to the 'data dialogue', particularly within the exit interviews.

Data Monitor

The data dialogue was enabled through many methods. The premier one among these was the data monitor, a small tablet that was situated on the coffee table that displayed a scrolling list of textual 'data announcements' to make it visible that data was being collected, used to trigger filmic events and personalise the experience. For example,

if an audience member picked up a book, a travel card or water. Not all uses of the data were real, e.g., 'emotional analysis' was a fictional element of the experience, used to present the possibility of and probe user reaction to such data collection by a media experience. This design response was inspired by the conversations around transparency and legibility engaged with in the first two studies and particularly the response to the IoT data collection scenario in Study I. The granular data collection that was exposed in this scenario in Study I provoked the respondents to think beyond the media aspect of the EPG to elicit the challenges of the underlying data practices. Hence, the data monitor here was employed to reveal the same kind of data collection, exposed at a very granular level but also simultaneously interleaved with the rest of the media experience.

Major Components : The data dialogue on the data monitor, enabled by the script on it primarily has three components.

- Data announcements that reflect actuator triggers within the room that the users witness within the space in a passive manner

- Data announcements that surface data collected from the users' active and direct interactions within the space

- Fictional data announcements that were included to provoke the users' thought and dialogue around granular personal data use by media experiences.

Design Process : The data dialogue presented through this monitor was designed and coded into the LRoTF in advance, through an iterative design process. The first stage of this data dialogue design was initiated after the storyboard and narrative was finalised and the inclusion of the appropriate sensors and actuators were decided upon. A first basic sketch of the data dialogue was made based on the actions of these technologies. This version of the data dialogue included a list of technological triggers that would happen within the LRoTF (both active and passive) in a chronological order, from start to finish.

The second iteration of the data dialogue design filtered out this basic list to produce a list of triggers that would involve a mix of passive and active triggers within the experience. The passive triggers are those actuator responses that happen in the room and are explicitly witnessed and noticed by the users. Such passive interactions would be "Blinds going down" and "Lights changing" where the audiences experience the effect but does not trigger it directly. Active interactions include user interactions with sensors within the LRoTF. These are executed by the users and the data monitor would surface these actions and the data collected from them. Examples of active interactions would be "Oyster card picked up" and "Water Bottle picked up" which are triggers that are actively initiated by the users

within the LRoTF by picking up an Oyster card or water bottle. These mix of triggers, being both witnessed and initiated by the users are thus 'announced' simultaneously by the data monitor. These announcements were designed to form a dialogue that, rather than present the user with a list of data logs, eased the audiences into the experience and the data flows through informal conversation styles like "Enjoy the book… you'll find more info about the building in the film". Thus, the data dialogue on the monitor was designed to open a conversation with the users wherein it would help the users understand the data practices and through it, the rest of the media experience, in a holistic manner. By exposing the data flows through everyday conversational language and careful interweaving of the dialogue with the rest of the media experience, the data monitor was designed to realise the principle of data legibility, while being sensitive to the context of the media experience being served, thus communicating the data practices to the users through the media experience itself.

Since the data monitor, by continuously revealing the underlying data flows within the room through the conversation it was having with the users, almost builds a relationship with the users, this relationship was used in the third iteration of the data dialogue to include extra data announcements that were meant to provoke the users. These bits of data dialogue were not reflections of actual data collection but possibilities within such a space that was included to probe user thought and action around experiencing such personal data collection within media experiences. These included "xxx dust mites detected in the room…. you have xx% of developing respiratory illnesses if continued at this rate" and "Low hydration level detected… Please stay hydrated". These data was not collected within the LRoTF but the inclusion of this dialogue on the data monitor along with other actual data collection and response dialogues that they witnessed within the LRoTF was intended to lead the users to believe in this fiction and provoke their response to such data collection.

The following is the final list of all 'data announcements' that were coded into the data monitor, to be displayed alongside the LRoTF experience, from start to finish, according to different triggers initiated within the room.

Announcement Type	**Announcement Scope**	**'Data Dialogue' on the monitor**
Response to sensor data	Personal	Audience detected... "Hi! Welcome to the Living Room of the Future. Sit down, relax and enjoy your experience! "
Response to sensor data	Personal	Audience seated.... Great, now you're sat down get comfy... lets dim the lights and draw the blinds
Passive action	Local	Starting Film...
Passive action	Local	Starting Fan...

Passive action	Local	Turning fan off
Response to sensor data	Local	Enjoy the book... you'll find more info about the building in the film
Response to sensor data	Personal	Oyster Card Data Analysed....The last vistor left their oyster card...
Passive action	Local	Change in brightness of film... lights changed
Biometric	Personal	Low Hydration level detected....Please stay hydrated :)
Audience state	Personal	Audience Attention Level Checked
Audience state	Personal	Audience Excitement Level Checked
Response to sensor data	Personal	Do you think that remotes will really be used in the living room of the future?
Passive action	Local	Film audio switching to radio...
Passive action	Local	Radio off... sound back to main TV
Passive action	Personal	Look around for the connected things...
Predictive Biometric	Personal	xxx dust mites detected in the room... you have xx% chance of developing respiratory illnesses if continued at this rate
Passive action	Local	Room temperature high ... cooling down
Passive action	Local	Lights turning back on...
Passive action	Personal	Deleting your personal data ... printing your receipt
Passive action	Personal	Hope you enjoyed your experience... please take your data receipt! :)

Table 2. Data Monitor Display Script.

Data Receipt

The second part of the data dialogue was a data receipt. This was a physical receipt detailing the data generated by the audience to drive their experience that is printed out by a thermal printer built into the coffee table. The receipt presents a list of the data associated with the experience the users just had, items that had been interacted with and the effect of the interactions on the unique ending of the film. It also highlights that all the data had been processed securely within the Databox and not passed to external parties. The aim of the data receipt was to provide a physical souvenir of the data practices within the LRoTF by reflecting a very specific set of user interactions which was also later on used as a resource to enable conversations during the exit interviews.

The following is the format used for the data receipt with the details of the users' data being filled in as per the interactions within each experience session.

~ Living Room of the Future ~

********** Data receipt **********

Experiment number: xxx (number of experiments)

Personal data processing started at:
Personal data processing stopped at:
Audience seated at:
Time in experience:

********** Room activity *********

Network speed:
Room luminance: (Luminance)

********** Your activity **********

This data includes:
Water drank: yes/no
Number of times the lights were changed:
Oyster card picked up: yes/no
Text book picked up: yes/no
Remote control picked up: yes/no
Data from these were used to choose unique ending: A/B/C

***** Beyond your experience today ******

LRoTF Databox has processed your personal data locally and ethically within this room.
We have not shared your data with any third parties and we have deleted your personal data upon your exit!

Thank you for participating!
This experience was developed by the objects of immersion team
Twitter: @immersiveobject
Tweet us about your experience

Figure 9. The Data Receipt.

Data Conversations

The data dialogue within the LRoTF also involved conversations during the exit interviews where the researcher explicated the data

processes further using the diagram displayed on the television during the UV reveal. This diagram (designed by the University of Lancaster) showed the various connected elements within the LRoTF. The researcher used this diagram to reiterate and reinforce to the users how the data flows were managed within the LRoTF, particularly discussing the role of the DataBox in containing the data within the room and deleting it at the end of their session. Just like the data receipt, these data dialogue conversations were used to help with the exit interviews, by encouraging user thought and dialogue around the data collection and responses included within the LRoTF.

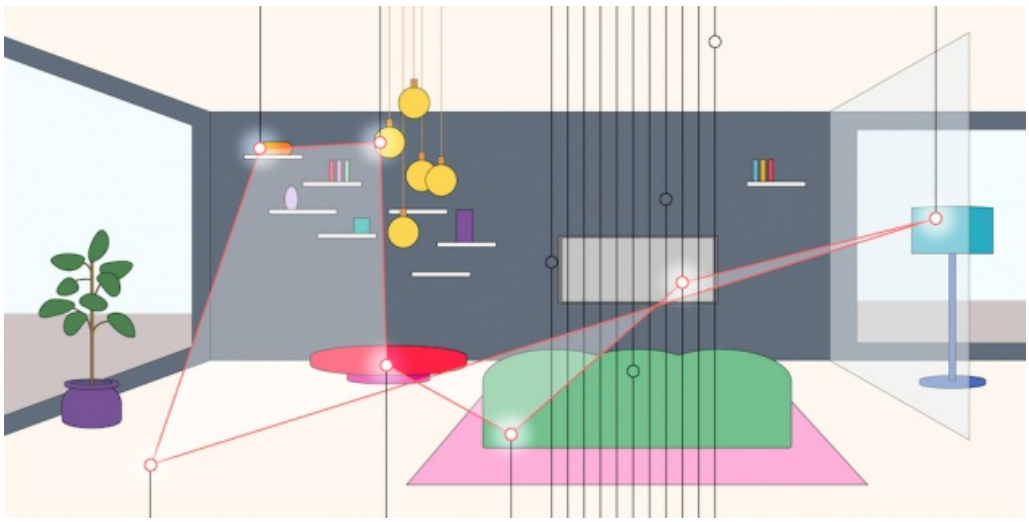


Figure 10. The UV Reveal Diagram.

The inclusion of the DataBox within the LRoTF in itself is an instantiation of the principles of Human Data Interaction. Further, multi-modal measures like the dialogue about the DataBox and the UV reveal diagram, along with the content of the data receipt and the data announcements on the data monitor during the experience consumption which formed the different aspects of the data dialogue, was woven into the appropriate parts of the media experience to make the data practices more accountable thereby realising better data legibility and agency (core principles of HDI (Mortier *et al.*, 2014)) within the LRoTF.

6.4. User Viewpoints

Participants in the study were served with ethics documents, approved by the University of Nottingham's Computer Science Ethics Committee.

The LRoTF exploited the affordances of being a design fiction experience by helping the audiences experience a near-future possibility in a tangible manner while simultaneously forming a research tool that helped provoke and elicit user response to such data-driven shifts and innovation in media. Through its world building approach it rendered itself to designing 'things' that produced the impression of a future world, where everyday mundane objects from a person's living room would collect personal data to provide customized media experiences.

Following the spirit of Design Fiction, the focus here was on encouraging people to engage critically with the potential issues that the forecasted future design embodies, thereby utilizing the design to probe and ask questions.

The LRoTF was conscientiously designed to do so by presenting the user with possibilities of data driven media innovation, while simultaneously affording them with mechanisms that provide more transparency, control and accountability over their personal data. **Here, the primary aim was to employ the 'data dialogue' as a mechanism to provoke the users' thought and dialogue around the collection and use of their data by media technologies in the near future.**

The evaluation of this innovation, does not only look into user response to the data management design decisions made to create this tangible data driven media experience, but by going beyond scenario projections of the future to situating the users within the reality of such an experience helps effectively probe and surface user reasonings about the very mundane use of such experiences in the users' near-future everyday lives, helping further inform and expand on the challenges uncovered in previous studies through more contextualised and practically situated reasonings.

6.4.1. Participants

For participant recruitment, an online reservation system was provided for interested users to book slots during the five days of research. Wherever there were free slots, walk-ins were also allowed. The event was publicised on social media and the websites of the various partners involved which resulted in most of the slots being pre-booked.

Each slot allowed a maximum of three participants to sign up. They could be a social group who knew each other, a couple of people who

knew each other and a stranger or a group of three strangers. Also, while the maximum capacity of the experience was three, there were also instances of groups of two and individual users using the space. This mix of numbers and social relationships was allowed to ensure representation of different possible scenarios within real Living Rooms where the numbers and connections are diverse, depending upon the users' lifestyle and context. Thus, the groups were a mixed bag of people who had previously planned and signed up to experience the LRoTF or walked in off the streets while they were simply spending a day or afternoon 'in town' or 'taking a look' around FACT. No reward was offered for participation.

6.4.2. Data Gathering

The evaluation of this experience was conducted 'in the wild'. This method seeks to elicit naturally occurring responses to novel technologies in contrast to the carefully crafted responses that result from highly controlled lab-based experiments (Crabtree et al., 2013; Rogers and Marshall, 2017). The LRoTF was thus deployed, and made available to the public at large, at the Foundation for Art and Creative Technologies (FACT) in Liverpool for an initial planned period from May 4 to May 8, 2018. However, response to the LRoTF was such that the deployment was extended for another 15 days, during which time more than 2000 people visited the LRoTF. The experience has since been deployed at the British Council's Play Festival in Skopje, Macedonia, and at the V&A as part of its Digital Design Weekend. Nonetheless, the evaluation here focuses on the initial planned period of deployment at FACT, during which time 59 people in 30 groups of up to 3 agreed to take part in the study.

6.4.3. Pre-Experience

Before entering the LRoTF, participants were served a short demographic questionnaire. The questionnaire began with a number of basic demographic questions about age, gender, ethnicity and occupation. This then continued on with more open-ended questions which inquired about the participants' motivations for visiting the LRoTF, their involvement with and use of technology, their take on data privacy and media consumption patterns (Questionnaire attached in Appendix I).

The data from the questionnaires were analysed to understand the mix of populations that attended the experience and to uncover and accommodate for possible biases in the population considered in terms of their demographics, use of technology, interest in the digital economy and media consumption habits.

All participants were above the age of 18. The cohort consisted of 26 females and 33 males from diverse ethnicities (White British, White Irish, Asian Indian, Asian Chinese, Middle Eastern, Mixed) and nationalities (English, Irish, Chinese, Indian, Romanian, Cypriots,

Macedonian). While we make no claims to the representativeness of the cohort, participants had a similarly diverse array of occupations, ranging from barmen, nurses, students, film makers, producers, software developers and designers to landscape architects, account executives, animators, marketing specialists, gardeners, journalists, researchers, corporate managers, company directors and civil servants. 36 participants reported that they had heard about the Internet of Things and 51 said they owned at least one connected device. 43 participants reported that they were 'concerned' about data privacy, 12 responded 'maybe', and 3 said 'no'.

6.4.4. Post-Experience

After the experience, the participants took part in an exit interview, which was recorded on audio and subsequently transcribed. The data receipt (Figure 9.) and a schematic diagram of the connected objects and data flows (Figure 10.) within the LRoTF were drawn on to seed conversation and probe participants' experience. The interviews were informal and unstructured, driven by what participants had to say about the LRoTF rather than a pre-planned script and series of questions. The goal was to root discussion of the LRoTF and the innovation it represents both in the participants' immediate personal experience and their mundane reasoning about the projected role and place of such services in their everyday lives. These interviews focused on 4 broad areas of interest:

- participants' immediate response to the LRoTF experience;
- their response to the collection and use of personal data to drive such experiences
- their response to the 'data dialogue' the LRoTF engaged in
- their reasoning about future physically immersive, data-driven media experiences in their own everyday lives.

6.4.5. Data Analysis

Subsequent analysis of the interview transcripts attended to the 'endogenous topics' (Pollner, 2010) that were accountably (Garfinkel, 1996) attended to and elaborated by participants and the interviewer in their talk together. This analytic approach differs from thematic analysis (Boyatzis, 1998; Braun and Clarke, 2006b) in the attention it pays to the objects of mundane reasoning manifest in conversation, rather than the identification of thematic patterns across a dataset. Rather than formulate thematic categories then, the analytic task is to attend carefully to the talk's objects – the topics that inhabit and animate discussion – and to carefully explicate these and make them available for broader consideration. It is an ethnomethodological approach that takes inspiration from Conversation Analysis as originally construed by Harvey Sacks

(Sacks, 1992). However, its focus is not on 'the machinery' of talk but rather, following Lynch (Lynch, 1997), on what people do through the mundanely competent use of that machinery.

Using this method of analysis helped uncover topics in the data that in addition to elaborating the experience of the LRoTF, illuminated further the adoption challenges of new media services that exploit connected objects and personal data in everyday life. The following sections describe these results. Further exemplifying the data driven approach of the analysis, the results are presented through the words of the audiences which form little vignettes of talk that help represent the topics, wherever possible.

6.4.6. Experiencing the LRoTF

The analysis reveals that participants found the LRoTF to be an engaging but not unproblematic experience. Their talk surfaced a range of topical concerns with privacy, agency and control, value, the potential social impact of data driven media experiences, personalisation and filter bubbles, and trust, which is unpacked below. We start first with participants' reaction to the LRoTF experience, before moving on in subsequent sections to treat their reasoning about collecting and using personal data to drive such experiences, and adopting data driven media experiences in everyday life. Participant's names are not their real ones in the edited extracts below.

6.4.6.1. A 'mad', 'magical' experience

While the LRoTF surfaces a range of adoption challenges, the initial response to the LRoTF as a novel media experience was overwhelmingly positive. These positive participant responses to the LRoTF ranged from 'calm' and 'relaxing' to 'mad' and 'magical', making the room feel almost 'alive'.

Peter: It was a nice, calm, relaxing atmosphere, it was. I liked that the blind went up on its own. I'd have that.

***

Sarah: It was really good. It was a different experience. A bit mad but in a good way. I wasn't expecting it to be like that.

***

Clayton: It was really good. It was a different experience. Like, it was a bit mad, but in a good way. I don't know, I wasn't expecting it to be like that. It wasn't half a bit intense. Like, I don't know, I want my living room to be like that.

Interviewer: Oh, do you?

Clayton: Yes. Like, that was really cool. Like, the lights changing and everything according to what was going on on the TV, and then the fan coming on and off and stuff. It was 'wow'! Like, I don't know, it just messed with my mind … really cool. This has made my day a really good day now.

***

Alex: To my friends I'd use the word 'magic'.

Interviewer: Magic?

Alex: I would say it's magic. It's just, it's in sync. It just happens with what you're doing, and whoever's in the room controls it.

***

Sarah: It was alive. The lights changing, anything I moved was reacting, it was a heartbeat, that room.

Clayton, Alex and Sarah were and are in no way special cases; all of our respondents found the LRoTF thought-provoking and engaging. Participants were particularly receptive to the synergy between physical objects and media content and reported that the interplay between the two made for a richer, more 'sensory' experience:

Jo: It was good. I didn't expect it to be as sensory. I thought there would be an interconnectivity of all the objects in some manner, but say if there was a yellow hue on the screen, that would then reflect in the lighting, which I thought was quite nice.

Lee: The way the room was reflecting what you were seeing colourwise – if there were shots where there was a lot of greenery on it then the room was turning green. Then there was a lemon, it was turning lemon. It just seemed to reflect what you were looking at really.

Jon: Also, like, the breeze. Like on the train, the underground.

Lee: I liked that moment. Her hair was blowing and then suddenly the fan went on. It was nice.

Jo: And the temperature, yes. Was that just coincidence that it linked in with the train scene?

Inclusion of everyday physical objects into the experience often elicited strong responses from audience members, who felt the experience was more immersive, entertaining, and novel. These higher levels of immersion was owed to the audiences feeling like the room itself was 'reacting' to them, making them feel like they were

much more a part of the experience, whether it be by having a 'conversation' with it, by empathising with the characters on-screen or even by feeling like a 'star' of the film.

Janet : I got a response from the TV, from the small screen on the table, I'm sorry, yes.

Interviewer : Did that help you in any way?

Janet: Yes, of course, that was my response, that I interacted with the room and that the room reacted back to me.

***

Bill : So, I love the idea that it immerses you so much. The lights changing, for me, was so good, because you felt what she felt, and the fan was so nice.

***

Alice: It was very immersive.

Dave: It gives you a better feeling of watching a movie.

Alice: Yes.

John: It just makes the whole experience more enjoyable. It's not just looking at a 2D screen in front of you. It's creating the whole mood and everything.

Dave: It makes it more exciting.

Alice: Yes, it could benefit a lot of films, giving you that experience of being the star of the film, in the film, kind of thing.

The overall response to the LRoTF was largely positive: mad, intense, alive, magical, reacting to the viewer and immersing them in a rich sensory media experience. However, that is not to say that the LRoTF is without its problems. Respondents had a number of reservations regarding the collection and use of personal data and the prospect of living with data-driven media experiences in their everyday lives, both of which impact the potential for their adoption into the audiences everyday lives. The rest of the findings, in turn, focus on unpacking these user sensitivities in detail.

6.4.5. Collection and Use of Personal Data In LRoTF

This 'mad', 'magical' immersive experience traded on the collection and use of personal data, which LRoTF explicitly surfaced through the data monitor, UV reveal and data receipt. These design decisions allowed us to probe audience member's reactions to the collection and use of personal data and to engage them in detailed discussions

about their concerns and expectations regarding data-driven media services.

Here, audiences voiced their views on sharing their personal data for a media experience, their reservations regarding the same, concern over sharing some types of data versus others for media experiences, and their current discomfort due to lack of transparency and control. When talking about data collection in the LRoTF, audiences very often explicitly voiced their appreciation and feedback on the inclusion of design solutions that provide better data transparency like the data monitor within the room.

6.4.5.1. User attitude about data collection by media experiences

Users often acknowledged the potential benefits of sharing their data and were happy to do so if they saw value in the process. Within a media experience, value was dependant upon the individual's priorities and defined as something that would 'benefit' audiences. This value ranged from simple auto-fill and auto-correct measures to automatically ensuring service of age appropriate content to enhancement of the very media experience itself, as described in the following account.

> *Sim: It's difficult to quantify what I'd find beneficial at the moment. It's just that if it's going to make your life easier then it's a benefit, I guess. As to exactly what you'd define as being making your life easier, I'm not quite sure at the moment, if I'm honest, but yes, if it's giving you-, such as, you know, detecting the fact that, 'It's very hot inside a room so we'll cool it down for you. It's very cold so we'll heat it up for you.' Maybe, or even watching something in the wrong light that's bad for your eyes then it may adjust the lighting to make sure that it's healthier for you. That's fine.*

But, while they were mostly accepting and considerate of the data collection that was required to enable media experiences like the LRoTF where there was a 'benefit' involved, this willingness to share their data was often further conditioned on a number of factors, like the type of data and it's context of use, transparency afforded by the data practices, the option for user choice and control, (all unpacked in detail in the following sections), which were expected to ensure 'ethical' data practices that were deemed necessary if the user is expected to live in 'data-collecting' environments that are 'self-thinking' and 'intelligent'.

> *Kim: I would say, to answer this in two parts, on the one hand of course, I can understand that in terms of future technologies and future audience and an increasing flexibility in terms of the next generation's viewers, that we need more flexible systems. Yes, to fit into this transitional lifestyle, we are all going to proceed over the next couple of*

years, I completely understand this but this can only be based and should be based on this ethical framework you've showed us today….[…]…. To be honest, as I said, we definitely need this ethical framework in terms of intelligent data-collecting systems. Self-thinking systems, and this can be the only solution to be able to live in those surroundings, in my view.

6.4.5.2. The Data Monitor

The LRoTF was consciously designed with solutions that was intended to provide greater transparency of the data collection to the audiences. These included the data monitor, the UV reveal and the data receipt that was printed and handed to all the participants at the end of the experience.

While all three of these were acknowledged by the participants, the data monitor in particular was flagged by most of the groups as being the primary source of understanding the underlying data collection and its involvement with the experience. The following account demonstrates how a group understood the data collection within the LRoTF through the data monitor, how it impacted their interaction with the experience and how it probed their thoughts on the data practices.

Interviewer: Did you understand that the experience was driven by your interactions and data?

Martin: The small screen in front of us was showing what was going on, even down to saying what elements are in the room in terms of dust and things like that and how it could affect you.

Lesley: As soon as I realised I was more focused on the fact that everything was actually data-driven and the whole idea that data is subliminally being taken from us all the time became kind of more apparent I suppose.

Interviewer: At what point did it become apparent?

Lesley: I think immediately when you took the book out and the screen said 'enjoy the book', and I was like, 'right, OK' and I picked up the travel card and it was like, 'OK yes.' It's like a stripped down way of how in everyday life those bits of data are always collected from you. Like I say, the travel card's a really good example. You use it every day and you never think about it, when really it's collecting a lot of data about where you are, how much money you spend, where you're going, what times you're going. You're forever, kind of, getting data collected.

Interviewer: If this could become your living room in the near future, would you like to know about the data that flows in the background?

Arianna: I'd definitely want to know.

Interviewer: Why?

Arianna: I think partly curiosity but also it's a good way to keep track of it, make sure it's not going too far the wrong way.

Thus, respondents generally saw the monitor as a key part of the LRoTF, a source of interaction, communication and feedback between the experience and the audience, which they often looked at in order to try and work out what was happening within the room and in doing so came to see and subsequently voice retrospectively concerns about 'background flows of data' and the need to 'keep track' of them.

The audiences expressed the need for such alternatives that provide data legibility in similar media experiences by explaining how the data monitor helped them understand the involvement of their data within the experience. A few accounts that reflect the same are as follows:

Paul : Yes, if I didn't have the monitor, I'm not sure I would have understood that things-, because if you were to ask me what I felt was driven by my data, I wouldn't say the length of the interaction, or the lights, or the fan, although it does say the room's heating up.

***

Mackie : It was interesting the small screen in front of us as well was showing what was going on even down to saying what elements are in the room in terms of dust and things like that and how it could it affect you.

Pim: Yes, that was really cool, I thought.

Yu: Yes, it was useful information.

Pim: Yes, I thought it was very interesting.

Mackie : Exactly what was going on and why it was going on.

In the course of the experience the data monitor even encouraged participants to actively engage with the connected objects driving the experience:

Jack: I just grabbed everything.

Interviewer: You grabbed everything. Did you know that something was happening when you grabbed them?

Jack: I got a response from the small screen on the table, yes.

Interviewer: Did that help you in any way?

Jack: Yes, of course, that was my response, that I interacted with the room and that the room reacted back to me.

***

Louise: It helped me understand things that I would otherwise not know was happening, like the emotional analysis. I feel if you're doing stuff like that it has to be clear that that's happening and the screen was doing that. If the screen wasn't there then I would have no idea that that sort of data would be collected on me.

In reflection, as Louise's comment suggests, the data monitor surfaced the need people have to be 'clear about what's happening' when data is being collected, particularly what use it is being put to, e.g., emotional analysis.

The data monitor clearly had distinct affordances and invited participants to probe and engage further with the experience as it unfolded, but it was not without its troubles. Participants felt that there was something important missing from the data announcements:

Paula: The monitor kind of showed what data was being collected, but it didn't show how it was being used. So, actually, I think because it was like a novel experience, I kind of spent a lot of my time trying to figure out, 'I did this, so now what's happening', rather than focusing on the changes and content.

***

Stuart: I wasn't quite sure where exactly the connection was. I was aware because the black type tablet was telling me like, 'You're doing this and you're doing this.' That was the thing where I could connect. So, some things are changing in here because of me, but I didn't really know how the rest was changing because of me.

In reflection, it became clear that participants need to understand 'what's happening' with their data extends to include the 'connections' between data and effect so that they can understand how or on what basis particular effects are produced (a point which also surfaced in respect to the data receipt and the choice of film endings).

Iona: I wasn't quite sure where exactly the connection was. It was more, I was aware because of the black, type tablet was telling me, like, 'You're doing this and you're doing this.' That was the thing where I could connect. So, things are changing in here because of me, but I didn't really know how the rest was changing because of me.

***

Betty : It was cool, yes. It wasn't very clear how the narrative is actually affected by the choices ...[...]...I think it'd be quite interesting to see a bit more about exactly what these choices would then trigger and why the film went a certain way because the water was drank or whatever, for example.

Thus, the audiences asked "So, like, what did we do that it [the LRoTF] started to make a change?" because of the lack of connection between the data and its impact on the experience. The audiences often questioned what was happening within the experience, both when the changes were disruptive enough for them to notice and when it was too passive to be explicitly noticed by them. In both cases, due to the data monitor, they knew that their data was being collected but this knowledge without the understanding of the impact of the data, only probed further questions on how data collection was impacting their experience and what they should, could or not do about it, during the experience, which led to users feeling 'stress'ed and 'confused'.

Gretta : I think so but I didn't understand the purpose of it, yes. I understand how to interact with all the objects and how to interact with them, but I just didn't understand the purpose of it, maybe....[...]... Maybe, I was thinking either should I-, I don't know how to say it, it's like, am I supposed to watch this way or am I supposed to touch this now or should I-, I don't know, should I be more comfortable, should I be more serious, something like that, yes.

***

Snow : 'Have I just changed the movie because I've picked up this?' or, 'Have I just changed the movie because I've sat back?' To me, that was stressful because I thought, 'I just want to watch the movie, and enjoy it, and be immersed in it.' So, the idea of it changing to my actions was something I felt was a bit-, it didn't make sense to me.

***

Simon: I think, when the book was picked up, some lights overhead lit up so you could read it, but it didn't replicate so I'm not really sure.

It was not just during the experience that the audiences had questions and asked for more information. But, even after the experience, they wanted to know how their interactions and their data led to the experience they were catered. Here, audiences particularly asked for knowing the choices that were made based on their data, to know how the choices were made and see how it affected the unique ending of the movie they were presented with, so that there was more meaning to their interactions, their data and thus, their holistic experience of the LRoTF itself.

Olly: Yes, we also couldn't see the different alternate endings, so we didn't really know if anything influenced what we did, because, for all we saw, it could be just scripted. So, it's not like, for example, if the person to the right didn't pick up his water and the person to the left did pick up his water, we couldn't see any difference between that.

***

Molly: I was a little bit confused as to what the end result was, what you were trying to achieve from it. When we talked about personal information and stuff like that, how it was collected, I didn't understand how you were collecting that by what we were doing. Does that make sense? I'm not sure what the end result is.

The data receipt, which was served to the audiences at the end of their experience and which documented the particular ending that was chosen for each group based on their data, further exemplified these audience questions and doubts regarding the choices made and the unique ending chosen for them.

Kate: So, yes (reading from the data receipt) 'chose unique ending C'. That doesn't mean anything to me.

Kate, like many participants did not know why unique ending C (or A or B) had been 'chosen' and nor could they consult the data monitor or data receipt to find out, just as voiced in the next quotation.

Thomas : I didn't realise that the lights changed that often. 'Data from these were used to choose unique ending B.' See, I didn't see that. I didn't notice that, obviously, it was changing things on the screen, so I didn't see that there was a unique ending.

Our respondents' reflections on the collection and use of personal data as part and parcel of the LRoTF experience thus highlights several concerns that coalesce around the *legibility* of data collection and use – keeping track of background data flows, being clear about what's happening with data and the purpose it is being put to, making

the connection between data and effect perspicuous – and being able to exercise *control* over the experience. We do not suggest that an improved data monitor, UV reveal and data receipt will address these concerns in the future, these are provocational devices not solutions, but rather that these concerns should be taken into explicit account when designing data-driven media experiences.

6.4.6. Adopting New Media Services in Everyday Life

While it is clear that there are aspects of the LRoTF experience that can be improved, particularly with respect to data legibility and control, it is clear that the 'user' response was very positive. At least, it was very positive in the context of an experience situated in what is effectively a public arts gallery. But the environment of the experience being situated in a Living Room closely modelled after any average living space, resulted in the audiences considering the experience in the context of their everyday lives, in their homes.

> *Cindy : It seemed like a normal living room. You didn't feel that somebody's watching you, that something was going to happen. Only when the things started to turn on, it was like, 'Maybe they're doing it.' Yes, it was quite interesting. I felt really excited at the beginning because obviously I didn't know in what way the room would react or respond, but it was just very casually switching the colour of the lights or, you know, turning the blinds down.*
>
> ***
>
> *Barb : It still felt like a living room, which is a good thing I think.*
>
> *Ken : Yes. It felt, kind of, natural, gave us little advice on how dusty it was and to keep hydrated and if you want to have more information to read through the book.*

This naturality painted by the setting, coupled with the subtle design fiction elements used for thought around near-future media innovation, probed the users to think beyond just a museum experience or an experiment, to consider the implications of their adoption of similar media technologies into their everyday 'private spaces'. Such extrapolation of scenarios from the experiment setting to the audiences homes resulted in a shift from the overwhelmingly positive response about the experience to voicing of more subtle reservations and reticence, which form barriers to the adoption of data driven media experiences into the users lives.

> *Savinna : In this situation, it's in installation. It's a replica of a living room. I could see the benefit of going into something immersive in a cinema experience, but the idea of inviting that into the home environment, that makes me think about the relationship.*

***

Megan : I feel like it would be interesting maybe in like a cinema experience, but like, in the home, I feel, I don't know, I don't need to have that level of immersion. So, yes, I'm not sure if I would be willing to like, make that trade-off.

Like Savinna and Megan, the broader cohort of participants drew a strong demarcation between appreciating the novelty of the 'installation' versus the implications of situating the experience it provides in their own living rooms. While they enjoyed the innovation provided by the LRoTF, participants were sensitive to how their attitudes and interactions with such immersive, data driven media experiences would change if they were embedded in their own everyday lives. These user reasonings were either specific and focused towards physically immersive data driven media experiences or owing to the provocative nature of the study and the experience itself, probed the users to express concerns related to the adoption of data driven technologies that speak to broader domains. It is towards unpacking these sensitivities that we now turn.

6.4.6.1. User Reasonings about Physically Immersive Data-Driven Experiences

The users expressed a number of challenges arising from the use of personal data particularly concerning physically immersive data driven experiences. These include :

- Privacy concerns and potential dystopias
- Agency concerns both within and outside of the media experience
- Value trade-off within a data driven media experience
- User concerns around accountability.

The following sections discuss these challenges in detail.

Privacy Concerns and Potential Dystopias

Unsurprisingly the potential use of personal data from connected objects situated within the home provoked privacy concerns. Participants viewed the home as a particularly sensitive context and were wary that data-driven media services would make it possible for external parties to 'snoop' on their intimate habits:

Jackie : I think that until it comes to this very private set or this private room, your safe space where you feel safe and secure. It's comfortable. ...[...]...There must be a system invented which is protected enough and secure enough to take care of your most private spaces.

***

Jenny: I feel very shocked because this is my room and I don't want my data to be collected and analysed by another person – there's no privacy! People need some space to live, they need to not be disrupted or something else.

The concern over privacy, particularly when considering the intimate context of the home is social in nature. The following excerpts demonstrate how audiences are worried about how adoption of LRoTF type media experiences could lead to their intimate behaviours being broadcasted to the rest of the society, thus compromising their privacy, forming opinions about them and forcing them to alter their behaviours within their own homes.

Uriel : I think, because you know how your data's collected beforehand, you are okay with it, but if it's, like, real, it depends. If it's printed just for yourself I would be okay with it, but if you know it's streamed somewhere then it would be more, like-,

Gabby : You might change your behaviour.

Uriel : Yes.

Gabby : For example, if someone tracked how many times you opened the bottle of wine or a pint of beer and it's broadcast, people would probably change their behaviour.

Uriel : If it's like just personal, for example, somebody who wants to know if they're healthy or something and it tracks how much they drink, I would want to know that, to know if you're doing something good or not, but I wouldn't want it to be streamed somewhere else. Then somebody else has access to it.

Agency Concern within the Media Experience

The source of the 'confusion' faced by the participants, talked about in the Data Monitor section (Section 6.4.5.2.) was in part due to the absence of any clear means of controlling what was happening within the experience. Here, they felt like they had no control over the experience as they did not know how the experience was being tailored for them or how to make changes in that process.

Serena: I didn't get a sense that I was affecting it, or I didn't have any sense of agency. I understood, or I thought that what we were doing was changing what you were saying, but I didn't feel any benefit from that…[…]…I think I'd want to see a subtle change, based on my, sort of, behaviours, and that would be quite impactful, but for me, I didn't feel the impact.

Within the LRoTF, this design choice was consciously made so that it would not behave like a giant remote control for the TV wherein picking up a book or water bottle again will not 'replicate' an effect as the 'effect' has already been triggered, thus making it an example of perceptive media rather than interactive media (Section 2.1.4.2. for details).

But this lack of opportunity to exercise direct control over the experience was noticed and discussed often in the interviews. Respondents asked for opportunities to control and interact more directly with the experience in easy, quick and 'frictionless' manners as represented by the following quotation :

George: I was really surprised there wasn't like a Google Home or something in there.

***

Ahmed: Yes, voice. Frictionless interactions, more humanistic, humanistic to a point where it doesn't feel too futuristic.

***

Amida: For example, when you pick up the remote, we get a little message: 'Do you really think the living room of the future will be controlled with a remote?' How will it be controlled then, because there wasn't really any way of control? So, I think voice would be a good option.

They were sensitive about even the very social circumstances of their consumption and wanted to be able to exercise end-user control to customise data-driven media experiences around it because they thought that it might be 'scary' to actually live with such technologies in their homes with their data and behaviours being collected and processed continually, with them having no control over it.

Sara: So it knows I drank water because I picked up the water bottle. No, it's a bit scary in a way that there's technology that can track what you're doing.

***

Lily: I felt that there were sensors in the cushions, kind of round – tell me if it's wrong – I had this feeling, and I thought about taking a seat on the ground because the couch suddenly became very suspicious.

This need for control is exacerbated in social scenarios involving multiple audience members where the need to customise data-driven media experiences around the social circumstances of the consumption and the user expectation to be able to make changes in this scenario is highlighted.

Kelly: Also, if it's a group of people. So, if for example you were living on your own and the living room collects all your data and it's set by your own standards and then a group of people come. You know, how would the data react then? Would it be confused or would it afterwards adjust to the number of people?…[…]…What if there's one person who's extreme one way and one that's extreme the other? How does it try and be in the middle?…[…]…What is it? A compromise, no one wins, or something? Like, just someone go and get another jumper or someone take their jumper off, you know what I mean? Not, everyone half.

Whether it is a matter of turning sensor data off to avoid 'scaring' people or making participants feel uncomfortable or manipulating some other data sources to exclude 'bystanders' from the experience, our respondents felt it important that they be able to tailor the physical context of data-driven media experiences around the social context. This also included optimising media experiences for different members of the cohort:

Terry: Like, if there are three of us, you could at some point know that, 'This is person A, this is person B, this is person C, and I know that person A has got this problem, and person B may have this problem.' It may adjust things accordingly, but then again, it may have to make a collective decision if the three of you are in at the same time, 'Well, this person's got this problem,' or 'This person's got this problem, this person doesn't.' How do you make it so that it is a good experience for all three at the same time?"

To overcome this sense of lack of agency, they asked for better communication with the experience. While right now they felt like the LRoTF nudged and directed them in certain ways, they want their future experiences to be more open to audience suggestions, in other words, provide the audiences with options for improved, informed and enhanced control over the experience.

Terence: Never mind, take it. Of course, when you know what's going on in the background, you have the feeling that this room is alive, because it goes forward in a relationship with you or it's starting to have a relationship with you. Actually, it's, in my view, a very one-directed communication. So, like, the room is only communicating with me but I am not communicating with the room, means it's very one-directional, which is quite new because normally I am the one who is communicating in the room.

Thus, here, control is about configuring particular data sources as part of an experience, whether to alleviate the concerns of participants or to enhance the experience for others. Here, a one size fits all model

becomes a barrier to the adoption of data-driven media experiences in everyday life.

Agency Concerns outside of the Experience

Respondents expressed significant concerns regarding the impact of data-driven media experiences on their everyday agency within their homes if they were to find their way there. In addition to the lack of direct control over the media experience itself, as highlighted in the previous section, respondents voiced concerns over their ability to control the use of personal data and the impact of autonomous services on their everyday lives. Living in an environment that automatically responds to its inhabitants interactions was seen as something that could potentially undermine one's personal autonomy:

> *Jake: I liked the idea of it, but as I said it takes away your autonomy. When you get bored, you change the channel at the moment, whereas that system might change it for you, and open the blinds, close the blinds, etc. It feels like if I'm not following exactly what's on the screen, then am I doing something wrong. It's a bit like it's telling you how to live your life.*
>
> ***
>
> *Karl: It felt a bit robotic, because whatever we were doing, the screen was telling us, like, 'Now you're feeling thirsty, have a drink' ... It felt like we were not in control of the environment.*
>
> ***
>
> *Lynn: My living room, to me, is where I want to go to get away from it all. I don't want to necessarily be prompted to do things.*

This sense of loss of autonomy, was not appreciated, particularly in the context of the home. As Lynn says, "*My living room, to me, is where I want to go to get away from all*-", and hence the audiences "*don't want to necessarily be prompted to do things*". Rather, as Fedora states in the next quotation, they expect a sense of freedom and control over their everyday lives, which they feel might be challenged if such media experiences are let into it.

> *Fedora : Again, always goes back to if it provides what I want. Do I want my water bottle telling me, do I want my coaster telling me, how hydrated I am? Probably not. Do I want to know how many dust mites are in my house at any one time so that I can breathe better? Maybe, but I'd want a system that automatically cleans it out at the same time. So, I suppose instead of just telling you, that's something I'd want.*

In the home, users emphasised the want to be able to do stuff themselves. They develop a sense of control and agency from doing that. There are days when they would like to hand over control to automation but even then, they are very keen on being assured of their control over this shift. The next account details this user want to be able to feel that sense of agency within one's home by having the option to manually execute tasks and to be able to shift to automation, when needed, all while emphasising the need for user control across all possible scenarios in the home.

Dean : I like automation of things, definitely, but there's also an element of doing it yourself, instead of-, if you had something like that, you could just easily sit on the couch all day and let the computer do it all for you. There's an element to getting up, moving around, going here and there.

Henry : It's the same reason for me as well. I can very easily become a very lazy person. So, to have the opportunity to be that-, I don't know how healthy it would be. At the same time, forgetting about that, it's quite cool, just like, you know, the fan turning on, the shutters coming on. You switch on the TV, 'Oh, there's a movie happening. Draw the blinds.'

Dean : Yes. In certain circumstances-,

Ronnie : Yes, but it's only going to happen in the evening sometimes, when you're home. You don't want to get up and put the blinds down. At the weekends, when you're not working-,

Dean : Yes, it would be good for those kinds of things. Other times, you have the option to just turn it off or just put it on standby and do it yourself. Then, maybe in the evening, it gets cold, and stuff, so you just let that do it for you. A good mix.

One potential solution, which runs counter to current enterprise models of personal data use, was seen by respondents to lie in the ability to personally control not only the experience but the data that drives it:

Daz: I want to be able to own that data and negotiate that with them; if you like, be able to pick and choose.

***

Yule: Certainly, having control of [the data] myself, not to say that I wouldn't be happy for people to use the data, just as long as I consent to it. Having that choice is key, and is very important.

Like Daz , many respondents were sensitive to the ownership of their data and having the ability to control who can access it, to be able to 'pick and choose' as it were. This process of being in control of the data use was often brought up in the interviews through three different, but highly interdependent user wants.

Firstly, they were keen on understanding the data exchange process and the rationale behind the decisions made for them from the data collected. Here, they ask for improved legibility.

> *July : Maybe if I can visualise it in some sort of easy and understandable way of saying, 'You talked about drinking beer. We're going to now sell that drink of beer that you drink to the company to try and sell you more ads, are you okay with this?' Just be a bit more transparent with how that is then sold on.*

Secondly, just like within the experience, they wanted space for negotiation regarding their data. They asked for the capability to make choices rather than be served with what the system deems a best fit.

> *Mai: It does seem very much like it's an all or nothing thing with data sharing, is that as much as they say you can pick what you want to select, you can't. Your data's always being collected, and even if you opt out, you still have data collected.*

Thirdly, they want to be able to know that they have control over these choices and that they are in a position to exercise informed control over their data, an instance of which is explained in the next quotation.

> *Fibi : I like to say, 'This is not relevant to me, I'm not interested in football,' or whatever I'm not interested in. They might profile, male, whatever, whatever age, and a lot of it I'm just not interested in. I'm interested in my own things. So the ability to say 'not relevant to me' or something like that, and then that maybe going back into the data to be reused.*

While voicing these elaborate expectations with regards to informed data control, the respondents simultaneously recognised that <u>actively controlling their data could be time-consuming</u>. The issue of data control became a double-edged sword where the need for control was and is perceived as necessary, but the time and effort required becomes a barrier to its proper execution.

> *Michelle : I can see that that would help. I know this sounds a bit annoying, but I don't have time. In our normal life, physical life-, and then to think that, on my mobile*

phone, I've got to spend maybe twenty minutes sorting out what I want to see. I try to minimise how much I use my phone anyway. So, it's a bit like, 'Oh, now we have to sit there, and filter through things, and choose what we want to see.' It would help, and I could make myself do it, just because I have to, because of the world we live in anyway, but it would be a frustrating thing anyway.

Little surprise then that the majority of participants were enthusiastic about the Databox, which they felt would help them control what was done with their data and lower the chances of it being stored, shared with external entities or recycled for purposes they were unaware of, making the entire process more 'ethical' from the audiences standpoint.

William : I don't mind because it deletes it. I like the ethical data thing because something I do worry about is the way our data is used generally. So I think if it was in my own home and I had the security that it would only be used there or deleted, then I would quite like it.

***

Nani : I think the fact that it's a box that stays in my house and I do decide-, like, for instance, it would ask me, like, 'Oh, so we've noticed that you watch this show. Would you like the BBC to know it?' I'd be like, 'Okay.' I would feel like I have a bit of control over it. I can be like, 'Okay, I don't want to share this, I don't want to share that, but I'm happy with them knowing that I've travelled, whatever, so they can show me a documentary on this.

***

Marty : Well, I think it's fine if it's used in this case where it's just client-side. So, it's not the broadcaster having data on you to push an experience. They push a different possibility of experience and then it's rebuilt, you know, at your own side. So, I think that's completely fine, yes.

Not being able to feel a sense of agency and autonomy by not being able to exercise control over any underlying data collection and through it, one's environment appears to be a major barrier to the adoption of data-driven media services in everyday life. Mechanisms enabling consumers with control over their data and the access to it are needed to foster their uptake at scale. Here, solutions that only use data locally and delete the data from the service after use hold particular promise.

Value trade-off in a data driven media experience

> *Donnie : I'm slightly annoyed-, I was listening to something about this yesterday on the radio as well, that I provide a lot of information which can then be collected by private companies and monetised by them. Now, when I provide information, even if it's just, like, if I'm buying something-, like, you go to Topman now, you buy something, they ask for your email address, so they can email you the receipt. That's really great because it saves paper, but it also means they're getting then a database of information about the sort of things you buy. Well, why can't I charge them if you want that information, take 5p off what I'm buying. I can then personally monetise it. There's no way of doing that, it's very difficult to do that and they hold the power in that sense. So, that sort of issue concerns me.*

This account demonstrates the 'annoyance' of an audience member speaking about their data being used and monetised by organisations in ways where they are led to feel powerless. While this account speaks of a more generic scenario of user data exchange for services, this study shows that as media services turn towards user personal data, the audiences see this novel scenario also as a form of exchange between them and the service provider, thereby surfacing and unpacking the flaws, concerns and opportunities in this particular exchange.

> *Catherine: I'm a, 'here's my data, show me what you can do with it,' kind of person. I'll happily give up personal data if you can give me something in return.*

The question of course is what constitutes a 'return'? Something that is worth the data they are giving away, something that is of value to them, something that provides them with "*A more convenient life, a healthier life, a happier life.*"

But, this current exchange is not often seen as a fair one as many participants questioned whether the value of the service they get in return for their data was symmetric to the potential of the data collected from them.

> *Martha : Yes, I suppose the extra information you get. So, I picked up the Oyster card. Found out whose it was, or it gave me some information, so you do get information. Is it worth all that data you give away? Not sure….[…]… It depends on how-, I think, for me, it's-, I'm not sure you get back what you put in. As long as it's kept in that isolated environment, that's okay, but if it goes out, which I think it inevitably will do, then do you really get back all the services that you want versus all the information you give out? I'm not sure.*

Here, some respondents thought 'experience enhancement' enabled through more intelligent media services that go beyond current data-driven media recommendation systems 'valuable'.

> *Stewart : I don't know. Like, it told me I'd watched an action movie. It had the lights on red. Other people say having the lights on red on an action movie increases more immersion. Give me this setting. So, you know, share my data with others to give me a better experience next time round.*

Others felt that "*An immersive experience in its own right is not a strong enough service, it would have to provide something of real worth."* When the data is collected for enhancement of a media service, the participants questioned whether enhancement of the media service was a justifiable reason enough to undertake such a shift, particularly considering the risks of data privacy and security, introduced by the same, which further diminishes the perceived value of the service.

> *Fille : I wouldn't feel comfortable to-, I mean, I don't think it would be a fair exchange from my side to get this kind of augmented experience in exchange of giving away all my (? 04.45) or my emotional state or my maybe health state, if I'm drinking water, maybe my posture or other things.*
>
> *Interviewer : Yes, why do you think that the exchange is not equal?*
>
> *Fille : It's not worth giving away so much of your private data for just entertainment.*

Others wanted more 'utility' and 'convenience' and envisaged scenarios where the LRoTF shifted to the kitchen and became a 'digital maître d' enabling an immersive cooking experience that helps with the cook by automatically turning the stove on/off and up/down, and notifying the user of ingredients that they might be allergic to, etc.

Thus, participants often questioned if the 'worth' of the service they received in return for their personal data was a fair exchange. Users expressed the need for a 'rationale' that is strong enough for their personal data to be used. This rationale should be able to balance the perceived value of the service provided in exchange for the data collected by maximising the benefits offered by the service while minimising the potential risks posed by the same. The implication here is for data driven media experiences to be compelling enough for the user to be comfortable sharing their personal data, to ensure sustainable adoption of the same.

Sean: So I think it's an interesting alternative approach, and there's potentially some benefits in there, but I think in the broadcast situation, I'm just not sure whether I see enough value in that to warrant having it, because my data is still being collected and processed in some way and I just don't know if there's a strong enough rationale in that situation for me personally.

***

Jack: I'm not sure you get back what you put in. As long as it's kept in that isolated environment [the LRoTF], that's OK, but if it goes out, which I think it inevitably will do, then do you really get back all the services that you want versus all the information you give out? I'm not sure.

Like Jack, many respondents questioned whether the value they get in return for their data is symmetric to the value of the data collected from them to drive innovative media experiences.

This symmetry also often depended upon the type of data exchanged. Audiences highlighted "*Yes, but I think some data is more important than others.*" Some considered giving their identifying details like names, addresses or even camera footage from smart TVs, that might have direct implications on their lives, as too intrusive to be given for a media experience, while being comfortable sharing their behavioural data.

Lilly : If they find out my passport number, my insurance number, whatever, then that might be a little bit more annoying. So, I think it's if you really limit the type of data that these things can collect, then personally I'm fine with that….[…]… I think so, yes. Like, if it's just, 'How did you react to a specific thing?' or, 'Did you pick up this item?' or even if you listen to music, if you like that music choice or that music choice, I don't think that's bad if you collect that.

Others were concerned about their behavioural data that reflects their everyday life, collected within the LRoTF being used out of the context of the experience, without their knowledge or consent, to target content, provide recommendations, push products and introduce customisation that might limit their access to the diversity of content available as a potential concern.

Thus, depending upon the type of data collected and the service provided in return for it, there is a call for symmetry of value in this exchange of data for media services so that the data driven media services provided are considered worthy enough for the data given by the users. Audiences are willing to give away their data for media experiences if the 'benefits' outweigh their 'concerns'.

April : Again, I don't mind. If there's a benefit to me or makes my life easier then, as I say, until you've ever felt that point where you are security conscious then I don't think you'll ever be too worried.

Irrespective of what form value takes for particular people or what type of data it collected to produce that value it would seem that enabling a more symmetrical relationship that demonstrably benefits those whose data is exploited is an important challenge to the widespread adoption of data-driven media experiences in everyday life.

Accountability for data in the home

Accountability for the collected data was often discussed as a concern when it came to adopting data driven media experiences into the audiences' homes. Audiences pointed out a current lack of accountability they felt due to the lack of transparency and control wherein they did not know "*What people are doing with your data and so I think that is a real issue for me.*"

Boxer : The data is being accessed by someone else, and even if they don't get to keep it, I don't have to have faith in their use of the data, rather than, maybe they're secretly storing it. You can easily, right now, you can still sign in to a bad third-party service and it will misuse your data even if it says it's not going to. So, they're part of my concerns.

As seen in the previous quotation, audiences are concerned about who has access to their data, what kind of processing is applied on it and for what efforts this data and its inferences are used. Since they are not offered a clear chain of accountability here, they express lowered levels of trust and confidence in the technology and believe that the possibility of their data being compromised is high.

Thor : It's like on Facebook-, well, I'm not on it anymore, but when I was on it you could go onto the privacy settings, and how information was shared. As somebody that knows very little about data, and what have you, I found that I was like, 'Yes, I don't want to share that. I do want to share that. Those people can see this. Those people see that.' I felt like, even though I'd customised it, my information was still not shared in the way I wanted, and there were certain caveats or ways that they could loophole around what I'd said.

The situation becomes increasingly sensitive and exacerbated when the context shifts to the audiences' homes and data about their everyday mundane activities are used in the scheme. Without a reliable system of accountability, audiences are reluctant to consider adopting such technologies, especially with the risk of compromise of their very private lives being put on stake.

Steve : It's just that, you know, as with any new technology, we all think, 'This is great,' but we end up finding out that this information is being used for this purpose, etc. So, I would be a bit suspicious about installing it in my home, just in case that data eventually did get somewhere. People know exactly what I do in my home. They'll know how often I use the toilet. They'll know exactly what time I get up of a day, what time I go to bed, and things like that. So, it does concern me a little bit that you're inviting something into your home that records data that doesn't need to be recorded.

Due to the sensitive nature of ubiquitous data from the home, even in instances where they are told that their data is safe, audiences showed lowered confidence and expressed that they are forced to blindly depend on the service provider to ensure the safety of their data as there is currently no regulated means of accountability to assure them of the same.

Rogers : I guess I was told, 'Yes, your data is being deleted at the end.' I guess I don't have a way of being 100% sure that's true.

***

Banner : You say it's deleted, it could be undeleted, I don't know how you're deleting it. I've got no reason to believe that what you say isn't the truth but I'm still a little bit cautious.

Here, there is a call to the service providers to intervene and provide better accountability as an organisation, so that the users are encouraged to place more trust in data driven technologies in their homes. In the following excerpt a respondent details how they expect a clear chain of accountability for the data that is collected from them so that they have confidence in the service providers' use of their data.

Wayne : So, like, I'd just like to say who do I trust, it depends, and I'd like to, kind of, say, 'Here's the minimum, if I trust you a bit more, I'll give it to you but, in any situation, do not pass it on,' like, it should be completely against the law to pass that on. If it does, there should be some, kind of, blockchain-type thing which goes, like, I can see where you've got that data from and track it, but the thing is, it's too complex.

Here, to increase accountability the inclusion of possible legal frameworks was discussed as an option. At a moment in time when audiences have very low knowledge about the flow and direction of their data, they feel enabling legal accountability makes the process

seem more regulated, helping them build more confidence and trust in the system.

> *Roony : I would probably feel safe using someone like the BBC because I feel like they're more publicly controlled and they're more governmentally controlled. I don't particularly like companies that aren't based in the UK, for example, who are able to get away with things because the parliament here can't control what their actions are.*

The solution provided by the DataBox wherein the data stayed within the box and was deleted at the end of the experience was considered a promising solution to the current accountability debate. The idea of the data not being sent to any external servers gave more comfort to the audiences as they felt the box was being accountable for the processing and sharing of their data. Owing to the novelty of the concept, they questioned the reliability of the box itself but was positive about adopting similar technologies if the accountability presented by the Databox was reliable.

6.4.6.2. User Reasonings about Broader Concerns

Audience reservations to adopt LRoTF-like data driven media experiences was not solely because of the previously discussed issues spurred by the experience itself. There were also discussions about broader concerns which included :

- Privacy Concerns beyond the Experience to Projected Dystopias
- Personalisation and the Filter Bubble
- Trust
- Lack of data Transparency and Awareness

All of which contribute considerably to the audience reluctance to adopt data driven media experiences.

Privacy Concern beyond the Experience to Projected Dystopias

The concern over privacy spurred by LRoTF type data driven media experiences has been discussed previously. But these privacy discussions often extended beyond just those social repercussions to the fear of how private knowledge about habits and lifestyles could be used for purposes that could have negative impact on the audiences lives, for eg., a health insurance firm denying premium to a customer who lives in a dusty home, a recruiter using the data during a job interview process, bosses accessing such data to keep checks on their employees or a compromised smart home that could easily let in a burglar when the family is on vacation.

Kevin : It's like a health insurance firm looking at it and going, 'Right OK, well if you come for a premium with us we're going to hike the premium up because we know that you are living in a dusty home.' It's one of those.

***

Yorky : If it knows what time I get home and what time to turn the light on, fine, but if it knows my name, my national insurance number, my medical records, when I go on holiday, so then it can let somebody know to come and burgle me.

In particular to media experiences, was the fear of being biased and manipulated by the media once they know more about the audiences through their data. Popular incidents like the Cambridge Analytica scandal were often cited as an example of the same and was used by the participants as a reason why they are reluctant to give away data about their everyday lives to such media services.

Love : I'm a massive believer in the media's, like, brainwashing us and stuff like that. So, I see it as, like, when you're on Facebook and you're inboxing your friends and you're, like, 'Oh, I'm just going to go to the café,' and then two days later you see adverts on Facebook for cafés. Like, that's mind-controlling and I don't like stuff like that.

These fears and concerns often led to the respondents positing various dystopian scenarios that mitigated against the uptake of data-driven media experiences in everyday life due to what was seen as their fundamental potential to breach consumer privacy and through it, their very lives:

Ramsey : You feel like you're being watched or tracked, like you're not necessarily alone in this room. It's like, 'Are they going to use this for other things, not just for my own home'?

***

Light : You feel like you're constantly watched and that's not a good thing for your personal freedom. You don't want to feel like you're constantly watched. You don't want to feel like you're constantly being processed in this big database. You want to feel like you're at liberty to do anything you want without judgement.

The potentially dystopian effects of streaming personal data and sharing or even brokering it with other parties is an issue to be addressed in the adoption of data-driven media services in everyday life. Concrete assurances are asked for from service providers to encourage their future uptake at scale.

Personalisation and the Filter Bubble

While they could see benefits in personalisation, participants also felt that allowing media experiences to always be tailored for them could lead to loss of freedom. Their experiences of being recommended similar content on currently popular media platforms has developed the concern of being trapped in 'filter bubbles' with limited opportunities for exploration.

> *Rose : I'd be frustrated, because I want to experience new things. The fact that I am trapped in my own likes is very – Google does it already, and I get really frustrated. You open up your Google page and it tailors what you see according to what you search. I think that's horrible. I end up having to use someone else's mobile phone to get new music, because all the music recommended is things I've already heard. I'm just like, 'Can I put that in the search engine and not get the same things, simply because it's tailored to me?' So, I don't like it. I think it stops you from experiencing new things, just because the Internet thinks it knows what you like. It doesn't.*
>
> ***
>
> *Sammie : So, generally, in terms of adaptive experiences and personalisation I think they are a good thing and that also comes because I work in marketing, so obviously I want to sell people stuff through those personalised experiences, but fine. What I have an issue with is always personalising experiencse because then you end up in an echo chamber of only ever hearing the same thing that you want to hear, over and over again.*
>
> *Suny : No diversity in that model, is there?*
>
> *Sammie : Exactly, like the old days where you used to read a newspaper and you'd read a story that you'd never search for on Google, but you'd be like, 'Oh, that's surprising, that's interesting, that's, kind of, changed my perception of something in some way.'*
>
> *Suny : Because it's next to something else, it's within the gaze.*
>
> *Sammie : Exactly, so, it's that accidental discovery whereas if it's all personalised you never have that external discovery, so, I think there has to be some kind of balance between the two.*

As Rose makes perspicuous, the tailoring of experience in the LRoTF sparked broader concerns with data-driven technologies trapping and manipulating 'users' by reducing opportunities for exploration. This concern is further pronounced in the case of News media. While

during the consumption of drama the concern of being trapped in filter bubbles is restricted to the possibility of being served lesser diversity in the content, in News media, the concern is exacerbated to being served biased content which questions the very ethics of the underlying strategy.

> *Pratt : I think there definitely needs to be an opt in and out of the shows. Without that you wouldn't understand what's real or not and that's important because if you're watching the news, for example, the news could be tailored to fit your likes, dislikes.*
>
> *Interviewer : Location.*
>
> *Pratt : Yes, location, if it's a sport one, in Liverpool, obviously, the Champion's League would be the first story, but say if it's a political one in London then a more Labour-driven Sadiq Khan interview might be first. There did need to be a certain level of communication, otherwise I'd be seeing something and going am I being sold the whole story? Or am I just being told something that is half the truth and biasing my opinion on something? I think that's important, the opacity of what is actually going on.*

Not only was this viewed as frustrating, but such tailoring of new immersive media content 'just for me' was also seen as potentially problematic in more serious ways.

Trust

The value trade-off and uncertainty that accompanies the use of personal data inevitably bleeds into matters of trust. Respondents' reasoning here, as elsewhere in the interviews, was informed by a general concern with the ways in which personal data is currently being exploited in society at large and the scandals that accompany it. This 'data zeitgeist' was further exacerbated by the LRoTF and the hyper-personalisation of media experiences:

> *Tanya : My trust in broadcast comes from the fact that everyone gets the same as me. That's where my trust fundamentally lies, knowing that everyone else has the same content. So, there's something about this idea that there's some clever processing of data that you're not quite privy to and don't quite know why decisions are being made that might change or alter that experience in some way. I think, for me, there is a risk that that could undermine the trust in the content that I get.*

While they could see benefits in personalisation, our respondents also felt that allowing media experiences to be hyper-personalised could have negative consequences that undermine their trust. Of particular

concern was the potential impact on consumer behaviour in 'changing' or 'altering' media experiences in order to deliver highly personalised content:

Adrian : If it would be to influence you in some way or other it would be, like, unethical.

The ethics of what is done with data and the potential to manipulate or 'influence' persons attitudes, behaviour and choices is a core challenge to the adoption of data-driven media services in everyday life. As Tanya hints at, enabling trust very much turns upon transparency in data processing and automated decision-making.

It is also the case that the respondents located trust in the service provider organisation, but 'track record' will not guarantee the acceptability of data-driven media services alone.

Phil : I suppose it would depend for me what company was producing the things, and the track record of the company.

***

Lenny : Personally, that worries me. I mean people like the BBC, as one example, I trust more than let's say, Fox News or whatever, but it's a future that worries me.

Interviewer : Why does it worry you?

Lenny: I think because it's such new territory. I worry that people aren't thinking enough about the ethics, it's like people moving too fast to really think about the ethics behind the stuff.

Trust is also moderated by the service provider organisation, but reputation alone does not guarantee the acceptability of LRoTF type services in everyday life as sometimes the audiences are left with no choice but to share their data and place their trust on the service provider, if they are to have access to the technology, rather than be afforded with opportunities that help them build trust organically and make informed choices as a result of it.

Suzanne : We all put our trust in them, we're kind of like, 'Oh yes, it'll be fine, they won't do anything with my data, everything will be fine.'

***

Lewis : But it's like what I was saying before about Facebook and that. I don't trust it because someone else has got, like, control of all your data and there are so many times where someone can get hacked and then, your data's out there for everyone to see. There's always someone out there that can hack something.

Meg : It's a line that you can't really draw, do you know what I mean? It's a difficult line to draw, so what is acceptable, what isn't? I think I put my trust in data collection more than I trust it. I think we all do. We all put our trust in it, we're kind of like, 'Oh yes, it'll be fine, they won't do anything with my data, everything will be fine.' Whether I actually trust it is another scenario.

Evidently there is a strong and compelling need to build trust into the collection and use of personal data if LRoTF type services if they are to be adopted in everyday life. For many of our participants trust turns, as it does for Lenny, on ethical considerations, which they found manifest in the Databox. Participants' responses to the local processing of data and its deletion after the experience saw this as a good example of ethical data practice, particularly when considering the sensitive nature of the data in their homes. It made the participants feel more comfortable at the prospect of adopting immersive data driven services in their everyday lives.

Mohammed : I don't mind the Databox because it deletes it. I like the ethical data thing because that is something I do worry about, the way our data is used generally. So, I think if it was in my own home and I had the security that it would only be used there or deleted then I would quite like it.

***

Tristan : I prefer the idea of the Databox to outside influence or control. I think it's a level of trust. If you're handing over, almost your safety, in some senses, to an outside company or person, then (TC: 00:10:00) you need to be able to trust that they will keep that safety in mind.

Here, if consumers are to trust data-driven media experiences then service provider organisations need to go beyond just issuing privacy policies and put solid mechanisms in place that enable consumers to have confidence that they have considered the ethics of leveraging personal data and that their data will not be unduly exploited.

Lack of Data Transparency and Awareness

One of the biggest contributors to the accountability concern is transparency and the current lack thereof, which was discussed often as an inherent problem faced by the audiences when engaging with media technologies that are powered by user data collection.

Lachy: The fact that, in most scenarios now, you are required to give your data over if you want to use the service, but again, I don't think that's made-, my Nan will go on YouTube because she wants to see something. She doesn't realise what's being taken for her to do it, most

people don't. I think that's where the issue is. There's no transparency in what's actually being taken and why it's being taken.

Users expressed a want for better transparency in this scenario. They want to know what data about them is being collected and how it is used, at a very granular level. They stated that having a better understanding of the data flows would help their involvement with the experience on two different levels.

Firstly, understanding the involvement of data, the data flows and the decisions made based on the audience data is expected to help the audiences understand the experience better, helping them involve in more informed interactions, making the entire experience more meaningful to them.

Fanny : To be honest, it might have been really good to know that [the data connections] in advance because we didn't know when to pick up objects and then how many to-, you know, if we'd known that if we picked up certain objects-, it might have improved our experience of the room.

Secondly, they want to know about the data collection to know what happens to their data and who has access to it so that they can "*keep track of it to make sure it's not going too far the wrong way.*"

Lack of transparency spurs the start of the user perception of lack of control, as users expressed that they felt more in control when they knew more about the data collection driving their experience.

Rupert : There's a point where, I'm fine with the use of data, it's more so that people aren't really informed about how and when data's collected. That's more my issue with it…[…]….. If you're aware of it, you're still in control, but a lot of people aren't aware of it. A lot of people won't really read the Facebook app terms and conditions, it's a really good one for that and no one ever reads that, and they've got complete control over your device and over your life at any point. I don't know, I think there isn't really a loss of control if you're made away, but most people aren't made aware or choose not to do the research and be aware. I don't know, I think in some cases there is a huge loss of control without anyone realising it, so yes.

Here, audiences feel like the service providers are not investing enough to ensure proper delivery of effective transparency as currently the onus is on the user to know, learn and understand about the data aspect of experiences. The extra effort required here, discourages many users from actively pursuing knowledge about their data, leading to more unaware interactions and experience

engagement and in turn towards lack of user trust in similar technologies.

6.5. Service Provider Viewpoints

As explained previously, the LRoTF was a multi-partner project that included different partners with slightly different agendas and expectations from it. Here, the BBC R&D were a major contributor with their interest being primarily the use of Object Based Media for real-time, data driven, content customization and delivering a novel media experience that pushed the boundaries of what is currently possible in media experiences.

In this capacity, the BBC were active participants (almost co-designers) in the making of the LRoTF experience. Thus, they were aware of and familiar with the 'data dialogue' initiatives included within the LRoTF as responses to previously uncovered challenges.

Hence, the research of the LRoTF study extends beyond just collection of feedback from the users to include the service providers again, to involve the viewpoints from both perspectives.

6.5.1. Study Design

The approach here was to reach out to BBC stakeholders who were involved with or aware of the LRoTF to explicate their feedback on the LRoTF as a media experience that leveraged user personal data, the 'data dialogue' initiatives designed into the LRoTF, how they see such initiatives responding to challenges and the future of such responses within media experiences that leverage user personal data.

Nine interviews were conducted as part of this agenda. The participants included employees from within BBC R&D, researchers, producers, project managers, team leads and a few high level BBC stakeholders with an invested interest in the data dialogue responses the LRoTF embodied. One of these participants were also interviewed as part of Study II reported here. They were engaged in unstructured, informal interviews that lasted 30 minutes to 1 hour. Here again the dialogue was not constrained by a script but rather let to develop around the following seeding topics :

- The data dialogue initiatives included within the LRoTF
- How the service providers see such initiatives responding to challenges
- The future of such responses within media experiences that leverage user personal data.

The interview data was audio recorded and transcribed. Endogenous Topic Analysis was employed for the data analysis. The results of this analysis surfaced legibility, control and the balance of delivering

value versus preserving user trust as major topics. These topics were discussed situated within the LRoTF, where the service providers, while highlighting the importance of these values used examples from the LRoTF to support its practical possibility within a media experience. They often cited the use of the DataBox within the LRoTF as an example of a response that alleviated a number of their concerns by helping them deliver enhanced value through personal data leverage while managing user trust in this process. The very underlying principles of the DataBox that enabled better data legibility and control was seen as a best practice response, a vehicle for further research and exploration and a possible solution to some of the significant challenges faced currently.

6.5.1.1. Legibility

Legibility was one of the core topics of discussion among the service providers. When moving the organisational structure and its technologies to one that leveraged user personal data, the importance of communicating the need and implications of this shift to the users was of paramount importance. '*So that the users know that their 'data is being used for good purposes or a better experience and not for exterior motives then you can relax and enjoy the experience*.'

As a service provider, providing mechanisms for legibility was seen almost on the lines of a duty or responsibility towards the users by empowering them through awareness of motivations, justifications, processes and implications of this shift, thus building more user trust in the providers' data leverage.

> *Letty : Either they just don't realise what data they are using or if they are kind of super suspicious and thats something that theres such a lack of awareness as to what true data is being used. So, I think thats a huge challenge and actually just the importance of the data on the audience especially, the importance of why they should care about their data and why its part of who they are and the identity is. I think people don't really necessarily understand why they should care so I think as a media organisation, I think especially being the BBC we are in a position where we really need to set ourself up and kind of communicate all of this to people. I think that would be the main, the need to make it clear like legible, that data we are using, why we are using that data and really clearly articulate that.*

Increasing data legibility and making the technology more 'intelligible' to the users is also expected to build in more accountability into the system, which, again, in turn gives the users more assurance, confidence and trust in the underlying data practices.

Data Legibility within Media Experiences

When employing methods of data legibility, the importance of focusing on its effectiveness was further highlighted. For this, legibility practices were expected to be more 'human' and less like '*a load of logs, a lot of text*'. Here, the need for contextually appropriate design interventions like the ones used within the LRoTF was emphasised upon.

> *Lars : I definitely see that kind of design being a way in. Yea, I definitely think that the answer is in… within the design of the communication.*

The 'data dialogue' employed through multiple modes within the LRoTF was seen as a response to this need for effective legibility. The monitor was to capture user attention through it's continuous updates and realtime responses, thus triggering in them the awareness that data was being collected, in a continuous fashion. While it also posed the opportunity for exploring what data was collected and why, this kind of deeper understanding was to be exposed and then reinforced by the UV reveal and then by the data receipt.

These measures were considered 'visual representations' of the data collection that made abstract concepts tangible to the users, thereby capturing their attention, probing their thoughts and articulating their thoughts around data better, thus leading to more engagement in the personal data ecosystem.

The importance of such intelligent methods of scaffolding in 'tangible' legibility measures into different layers of a media experience to ensure effective communication of the data flows through multiple modes was commented upon. Here, particularly spoken of was the role of data monitor within the LRoTF, wherein legibility almost became a part of the media experience itself, thereby reducing the concern of distracting the users from the experience and opening an avenue for exploration of solutions that could help balance user attention between the experience and data awareness.

> *Illya : You could have just sat there and gone through the first part and gone ah ok I'm watching something, I see some lights changing, its very similar to what I have seen elsewhere. That would have been easy. The monitor makes it very clear that theres a lot going on. And I think that is really important. I think that answers your question. It definitely introduces people to the notion that there is more going on underneath.*

For the media service providers, effective data legibility also serves as a method of alleviating one of the major concerns associated with personalisation powered by user personal data, the fear of black boxing, filter bubbles and bias that users fear of being exposed to in

their media choices. From the service providers' perspective, provision of data legibility is considered to help with this concern as the users are now able to see their data and observe how it results in the service they are provided with, helping them understand the decisions underlying the media they are served.

6.5.1.2. Control

Closely aligned with data legibility were the conversations around data control. The requirement for enabling user control of data once the practices are made legible was emphasised upon.

> *Pat : Ok, so I think its very important that audiences, people we engage with can understand the consequences of what they are doing if data is being collected. If data isn't being collected, then its not an issue. As soon as you start to collect data so that you can provide better services then you have to allow people more control or feel that they have more control.*

Control within a media experience

Data control within a media experience becomes a challenging proposition as the demand for user engagement and involvement in the experience itself is high and the proposition of involving interface attributes that could potentially take away from this calls to be dealt with much nuance. Hence, one of the first priorities while unpacking data control in this context was to ensure that the options of control do not distract the users and detract their attention and engagement with the media experience itself. So, "*whatever mechanisms will make it apparent, but not disturbing. Disturbing in the sense disturbing the engagement [with] the experience*". The second is the need for quick and easy data control measures, which ensures effective user control of data by overcoming the privacy paradox (Norberg, Horne and Horne, 2007) of users not engaging with privacy enabling technologies due to the time and effort they demand.

Control within the LRoTF

Within the LRoTF, the use of the DataBox was seen as a method of improved data control. It's local processing of data with no sharing with external servers and deletion of the data at the end of the experience was thought to eradicate the need for user engagement in extensive granular control, which could be time consuming and distracting, while allowing the option for more engaged and active user control, to those requiring it.

> *Isaac : And the advantage of it is that whatever happens in that experience, stays local and is not shared and you may have stuff that is unique to that audience, but no one else*

needs to know. I think that became very clear with the DataBox.

Also, having one physical tangible data management device that helped negotiate multiple data sources meant that users did not have to exercise data control on many different interfaces and devices, but just on the box to help them with their data sharing options. This was compared to a 'gatekeeper' a 'metaphorical filter' that safeguarded the data and which could be included to optimise the experience or excluded, as per the users' choice, making the herculean task of ubiquitous data control easier and more manageable for the average users.

Pastor : This is like your gatekeeping. There aren't a lot of levers you got to pull on many different things. It all is here in this box. So, if you think of all the items around a living room that are Internet of Things connectable they all could be giving out data. And you just want to…. got to switch it off and on…. but if you switch it on it plays into this data flow and optimises the experience.

6.5.1.3. Balancing Value versus Trust

It was stated within the interviews that 'T.V. and film as we know it now is increasingly boring. And one of the ways to keep creating media that isn't boring will be to make more emotional'. To involve this emotional element, there is the need for experiences to be more contextually aware, which in turn requires user personal data.

Thus, the innovation and value personal data brings into traditional media consumption scenarios is highlighted again here, from the service providers' perspective. But while personal data provides much value to media experiences by making them emotional, rich, immersive and contextualised to the audiences, this shift calls to be balanced with the questions around user trust in the services and service provider, which is expected to be built through data legibility and control.

Patrick : The main challenge is getting the right balance between the enhancement of the experience, the tailoring of the experience, finding the places where the tailoring is really valuable. But doing that in a way that preserves the trust that you feel like you are engaging in the experience, you are not engaging in a transaction where your data could be flowing anywhere.

Nature of trust

As a media service provider, the BBC often re-iterated a 'duty of care' towards its' contributors, which becomes a part of their principles. When shifting to personal data leverage, this duty is also expected to shift to the users who in turn become data subjects or data contributors in the digital economy.

Bass : As a journalist, I have to express a duty of care to contributors. So the BBC has some principles. They have a duty of care to anyone who appeared on radio, on television, on the website or whatever. We have to look after them basically. Not misrepresent them, not put them in danger, all sorts of things. If somebody requests anonymity and they have good reason to request anonymity then we have to keep their identities secret. All of those things. Well I also have a duty of care to anyone who engages with BBC online services. I think it is just as extreme, we have to look after them. And safeguarding their personal data is a part of that.

This duty of care is pioneered by ensuring preservation of user trust in this ecosystem so that "*data and trust and value of service provided to customer can flow without disturbing any one of the three.*" While traditionally the BBC might be considered as a trusted organisation, this trust is in its' role as "*a trusted intermediary, the BBC sits between you and the film and the world and it tells you about the world, it informs you, it educates you, it also makes you feel good about your place in the world, it entertains people.*" But, it was acknowledged that trust is not 'a single index' and so because the BBC is trusted by the audiences to give objective news, high quality entertainment and to not upset the children, does not automatically make it a trusted data broker.

Particularly owing to its position in the society as a provider of services that fall into the 'pleasure spectrum', where engagement with media services is not a necessity for any individual but a choice made by them. To encourage this choice, the users have to be convinced with both the quality of the service through innovation powered by personal data leverage and by engendering user trust in this process of personal data collection and use.

Maintaining this Balance within the LRoTF

On the one side, the LRoTF experience itself was a first of its kind, thus providing value through innovation and immersion. It was termed '*a forward looking experience*' that pushed the boundaries of current popular media services by providing contextualised immersion in novel ways. On the other hand, while providing this innovation, it also ensured preservation of user trust in the following ways :

Firstly, the inclusion of ethical design is key when it comes to creation of data driven media experiences. Here, the need to reason why data is being collected, identifying the need for the same, minimising it's collection, highlighting and exposing it to the users and safeguarding its transfer, storage and sharing, throughout the design, production and dissemination process, in the most effective manners, is of paramount importance.

M : Ethically I think that we need to be open and for trust, we need to… so, by open I mean we need to be clear what we are collecting and by way of collecting it, give options and clarity. Basically the data dialogue in your words is kind of a mature one and involves kind of informed choice. We are not, I guess, these would be prejority? words, we are not just trolling for data because it might be useful. We are not just ramming things and storing things and retaining things because the way we have written our experience code means that we can. You know head-mounted experience for example we can track every little head movement over time for the duration of the experience. If we want to do something like that, we should be clear that we are doing it and we should be clear with our audience why we are doing that. Because I think trust then comes in, in terms of the transactionality of that.

Secondly, enabling this ethical design to engender trust through legibility and control. As explained previously, the multi-modal 'data dialogue' methods of ensuring effective legibility of data practices within the experience to the users. This is expected to help promote awareness around data practices and also help build accountability into the system which is expected to help users trust the process more.

DataBox as a Response

The DataBox provided the opportunity to innovate around data driven services while engendering trust. It opened up the space for providing value to the audiences while eliminating the need for collection and storage of data by the service provider, thus adopting design that is more ethical and considerate of user trust.

The approach of the DataBox alleviated concerns at multiple levels. It helps sustained delivery of value through innovation using personal data. While the benefits of the data is extracted, it ensures this data is not collected, stored or shared server side, thus reducing any risks associated with the same. This becomes a response to the concern of reputation damage and loss of user trust stemming from it. All while providing better accountability to the users through not just the capabilities of control presented by the box but also the physical presence of the box within one's home, which becomes a constant reminder of ethical data use.

Peyton : For us its is better we don't have the data. That is much better, we can hold the trusted position if we don't have the data. We can provide value if you let us send our applications and our services to you. So that delivers a better experience to craft those things. Just as much as it is important for the audiences that they feel safe, we also don't want to feel in any way compromised that we put ourselves in a position of exposure to the accusation that we could

have done something the audience wouldn't like, with their data. The box in the home is a nice physical token. To say, the physicality of it is a token. This is where your data is and this is the point of contact. This is like your gatekeeping.

The inherent characteristics of the DataBox enabled data legibility, user control and user agency within this scenario. The principles of the DataBox were involved within the LRoTF as a means for enabling this control. Local processing and deletion of user data post experience meant that data sharing with external entities was removed thereby eliminating the very need to control the same. This method of control by default proved to be quick, easy and effortless to the user with minimum distraction from the media experience itself. The constraining of data flow to stay within the room and on the box made the client side processing legible to the audiences while additional methods of data dialogue (like the data monitor, data receipt and the UV reveal) supported this initiative further through provision of more granular data legibility.

This experiment was expected to set a high standard in the creation of data driven media experiences, challenging any other 'mainstream' options that actively promoted user data collection and transfer to external servers.

Clint : The academic approach, the architecture and the philosophical approach of the DataBox is what is the best we can possibly do? And if we do all these things and we integrate it properly, then we have something that is as good as it can be. Therefore, if you choose to do anything less than that, to compromise, is therefore to say actually I don't think I care as much as I think I do. It is that message forces people to answer the question of why do you want this data? Why do you want to be able to use it in this way? So, I think in that sense it is very valuable.

While the efficacy, implementability and scalability of such an alternative method of personal data management is yet to be explored, this implementation within the LRoTF was expected to be a vehicle for further study, while being a benchmark for future practices within media that leverage personal data.

6.6. Study III Conclusions

This study is a bridging study that uses the viewpoints of the service providers and the users, uncovered in the first two studies to design a disruptive future media experience that leverages user personal data while addressing some of the previously identified challenges. The experience is presented to the audiences to provoke them and elicit their feedback and reasonings around the personal data leverage by a media experience and the 'data dialogue' responses included. The service providers' feedback on these priorities are also collected by

involving them right from the design and making of this experience to the deployment and experience of the same. They are then interviewed to then discuss their feedback on the LRoTF leveraging personal data, the inclusion of the 'data dialogue' responses and the future of such responses within data driven media.

The audience responses to the LRoTF as an experience was primarily positive with the respondents considering the experience as 'mad', 'magical' and 'intense'. Here, while considering the data collection that supported the experience customisation, they expressed their expectation of value and benefit in return for their personal data being collected. This perceived value depended on a number of factors. These included either dependencies on the data itself, like the type of data and it's context of use or was based on support measures provided to them through this data leverage process like the transparency afforded by the data practices and the option for user choice and control. Within the LRoTF, the audiences voiced their appreciation of the legibility provided by the 'data dialogue' techniques involved and the use of the DataBox that made the data collection seem more ethical, making the exchange within the LRoTF seem more reasonable to them.

Despite the service enhancement presented by the experience and the active responses to data leverage included within the LRoTF, when considering adoption of such data driven media experiences into one's life, the users were apprehensive about the same. Here, they presented a number of concerns that form a range of adoption challenges that call to be considered when designing future media experiences that leverage user personal data.

These adoption challenges could be broadly divided into two :

- User reasonings about physically immersive data-driven experiences
- User reasonings about broader concerns.

User reasoning about immersive data-driven experiences could be unpacked as follows. Firstly, privacy concerns and dystopia projections of external parties being able to 'snoop' into one's private life and personal space. Secondly, the need for agency within the media experience, particularly to be able to know how the experience is being shaped by the data and having the option to make changes here. Thirdly, agency outside of the experience, to have control over devices in the home when they are leveraged as sensors and actuators, thus maintaining autonomy over one's personal life. Fourthly, the need for symmetrical exchange of value if one's personal data is to be exchanged for a media service. This value ranges from entertainment and service enhancement to the need for creation of utility and convenience. Fifthly, the current lack of and the want for

accountability, particularly when considering user personal data collection and use within the home.

The rest of the adoption challenges voiced by the respondents, form a range of more broader concerns, presented as follows. Firstly, the fear of loss of privacy beyond just the experience to projected fears of being manipulated and influenced by the customised media that is presented. Secondly, the concern of being constrained in filter bubbles and echo chambers through the personalisation of recommendations, thereby restraining the possibility of discovering novel, diverse media content. Thirdly, the lack of data transparency and awareness that raises questions regarding the nuances of the experience, thus lowering user engagement and also instilling the feeling of loss of control. And finally, the concern of trust raised as a result of all of the challenges described above. This challenge also highlights the need for building and sustaining user trust in future media experiences that leverage user personal data through active responses that are sensitive to the other challenges highlighted here.

The **service providers' responses** to the LRoTF as a media experience that used personal data and the 'data dialogue' associated with it, primarily fell into one of three topics. The first was the provision of data legibility which focused on the importance of communicating the need and implications of the data leverage to the users. Here, the dialogue was rooted in the everyday practical delivery of data legibility within media experiences by using the 'data dialogue' responses used within the LRoTF as examples. The second was the importance of enabling agency through control to the audiences while exposing them to data driven media experiences. Here again, the conversations were grounded in the effective delivery of this goal through provision of practical, tangible solutions like the DataBox. Finally, the importance of balancing the value delivered to the audiences with the concern of trust highlighted within this scenario. Here, the nature of trust when considering personal data leverage by a media organisation and being sensitive to this uniquely novel dimension of trust that is presented as an emerging challenge in this shift towards personal data leverage was focused upon.

The conversations concluded by pointing out the potential of the DataBox and its underlying principles as designable solutions to a number of challenges highlighted here by both the service providers and the users. The innate nature of the DataBox to constrain data processing locally, provide granular control of user data (if deemed necessary) and the data legibility presented are seen as promising responses to some of the heavy challenges affecting both major stakeholders in this scenario. With the view of exploring the potential of these promises of the DataBox further in this domain and reflecting the findings of the three studies reported in this thesis, the BBC has very recently released the BBC Box (BBC, 2019), which initiates a

first-of-its kind response from industry to the challenges presented by the use of personal data in media.

Limitations of Methods Chosen : The LRoTF study was designed and conducted using the design-led approaches of design fiction, probes and interviews. While this works well for providing a tangible, real-life experience that is disruptive, thus eliciting user thought and dialogue around the future of data driven media experiences, the question of how such an experience would fare in the everyday lived world is one that still stands to be empirically answered. Here, there is the potential to conduct design ethnographies (Crabtree, Tolmie and Rouncefield, 2012) that would help observe and study user behaviours around data driven media technologies. This would not only help showcase the implications of a similar media experience in the mundaneness of the audiences' everyday lives but could also help further open up the adoption challenges discussed in the following findings section, thus including and evaluating more solutions for the challenges.

7. Discussion

This discussion chapter contextualises the results presented by the studies in the previous chapters within the literature reviewed in the beginning of this thesis. The following chapter brings together the findings of the three individual studies of this thesis to present the challenges of **value, trust, privacy, transparency, control** and **accountability** as reasoned by the users and service providers. These challenges are explained in detail by comparing and contrasting them with the findings of the previously discussed literature to present the unique conceptual and design contributions this thesis makes.

7.1. Introduction

The recent turn towards personal data in media coupled with previous research highlighting the challenges confronting the use of personal data in diverse domains puts out an explicit call for exploration of the socio-technical challenges presented by this shift within media experiences. The work of this thesis explores this research space through three studies that explore these challenges from the viewpoints of the users and service providers. While the first study helped affirm and introduce the reason for concern within media as it starts leveraging user personal data from the users viewpoint, the second study helped explicate the challenges faced by service providers in this context. Using the data from these two studies and combining it with previous research around Human Data Interaction (Mortier et al., 2014), a media experience was built that included responses to these challenges. This intervention was used to probe for feedback on these responses and the use of personal data by this media experience, again from the perspective of both the users and the service providers.

The results from these studies show the complexities associated with value as a challenge confronting the turn towards personal data in media experiences. This value is furthermore faced with other challenges like privacy, trust, transparency, control and accountability, which further exacerbates the complexities presented.

The following sections unpack these concepts, drawing together the results of the three studies here, while simultaneously comparing and contrasting them to previous literature. Thus, it firstly presents the unique **conceptual** contribution of this thesis within the scope of data driven media. Once this detailed conceptual contribution is presented, the discussion then summarises these challenges with a focus on **design** implications. Thus, the discussion concludes by presenting a set of sensitivities to be considered when designing media experiences that leverage user personal data.

7.2 Conceptual Contribution

The following section presents the unique conceptual contribution of this thesis. It ties together the findings of the three studies reported here, to present the following challenges in a conceptual fashion :

- **Value**
- **Trust**
- **Privacy**
- **Transparency**
- **Control**
- **Accountability**

Here, these challenges combine the viewpoints of both the users and service providers, thereby presenting their response to the turn towards personal data in media experiences where they reason out their particular challenges associated with this shift. These challenges are thus unpacked, compared and contrasted against previous literature to open up those discussions to include the complexities that arise when the practical reasonings of users and service providers are included.

7.2.1. Value

Personal data and its metaphorical comparisons to oil and fuel is a much debated topic in today's digital economy (Kalapesi, 2012). This is particularly because of the unique properties of personal data (Section 2.3) that makes the terms for its' exchange for services different from exchange of traditional assets like oil and fuel. Irrespective of the outcome of this ongoing debate, what is agreed upon is the position of personal data as an 'asset class' (Schwab *et al.*, 2011) that brings with it the notion of value or the promise to produce value.

Here, while personal data might be considered a currency that is exchanged in return for enhanced services, it is also the very driver that enables these services themselves. Previous work around the value of personal data has briefly looked into its scope within media (Rose, Rehse and Röber, 2012). Here, the recommendations made could be seen as high level, leaning more onto financial value, often abstracted and much in need of explication to be able to be applied to the practicalities of creating and serving media experiences. The results of this study show that while these high level recommendations might hold true, the concept of value associated with personal data in media experiences is complex. The complexities in this scenario arise in the form of multiple actors, forms of value

created, value distribution between actors, the challenges to value created and the balancing of these challenges to maximise the value produced, all of which are detailed in the following account.

While value associated with personal data within media is spoken about mostly in terms of fiscal value for the service providers in previous industry reports (Rose, Rehse and Röber, 2012), the results of this thesis show that value manifests in many forms and is actually a multi-partner proposition within this landscape. Thus, while personal data offers several new avenues of value creation to media service providers, following on the notion of being an asset, personal data also offers the end users value in return for their data.

To the end users, value from personal data manifests as **new forms of content**. This could be a new specific experience that is customised to reflect a user's particular context and preference, a richer media experience built up by adding in more layers that make the content more relatable to the users or even the generation of entirely new genres of media experiences created using user personal data, like an EPG that helps send recommendations of TV programmes to friends thus making 'orchestrated shared viewing' easier. Users also see value in **personalisation** of media experiences. Here, personalisation begins right from content discovery where the provision of accurate recommendations is seen as valuable to quicker, faster and simpler user journeys to even personalisation of the media content itself to make it more relatable and engaging to the users. Value is also seen in the possibility of personal data making **the delivery of appropriate content** easier. Here, delivering age appropriate content is of particular interest to households with children. Also of value is the notion of 'social' media experiences being made easier through collaborative profiles that moderate content to suit everyones' interest, viewing history and age to deliver content that engages everyone alike. Also, personal data use fundamentally **changes the relationship between the user and the service provider** by giving the user an increased influence on the future of the media experiences they would consume. This shift allows for user preferences, viewing patterns, history, engagement and feedback to be a part of the making of future media experiences and allows for media experiences that are no longer generic but customised to fit their life and context, thus making the media creation process in itself more user-centric.

To the service providers, value from personal data manifests in quite different ways. The BCG report of 2011 (Rose, Rehse and Röber, 2012) had suggested that media "*companies have access to a vast amount of personal information mainly relating to telephony usage or TV-viewing data, but so far make relatively little use of it.*" and that when it is leveraged successfully, it would be *"a huge growth driver in a sector where growth and competition is closely linked to the ability to innovate and enrich existing services*" (Rose, Rehse and Röber, 2012). By starting to leverage this data to the fullest, service

providers expect to help themselves out with their **business mandates** like identifying consumption gaps and trends, spotting underserved audiences, adopting measures to overcome these flaws and improve their services to provide content of value to all audiences equally by helping their product development through enabling growth and matching with competition. Media service providers expect that by leveraging the full potential of personal data, they would be **achieving greater audience satisfaction** by understanding the user response better through uncovering trends and patterns in consumption, identifying systemic biases and resulting gaps in consumption and using triangulation of various variables (eg., location, preferences etc.) to possibly understand user intention behind each visit to provide products that resonate with the users' expectations. They also expect personal data to help them **remain salient** in the users' everyday lives by understanding the changing user interests, improving upon any current gaps between the service delivered and audience expectations and keeping up with competition in terms of what the audiences expect on an everyday basis. It helps the service providers gain more granular **insight and feedback on the content** delivered, helping them improve their responses to issues, flaws and enhancement of services. Just as discussed for the users, the inclusion of personal data changes the nature of the relationship between the users and the service providers. Having a more granular picture of the user preferences and interests results in more customised media being delivered to them. Thus, the organisation effectively shifts from a one to many model to a **one to one model** with their audiences, which is seen as a valuable shift by the service providers.

Thus, the value proposition here is distributed among **multiple actors**, primarily the media service providers and the end users (Any other entities involved like third party data brokers, advertisers etc. had not been included as part of the research agenda and were not discussed as part of the studies. Thus, while they might have a role to play in this distribution, their role remains out of scope for this particular discussion). The properties of this distribution are yet to be resolved though technologists like Jaron Lanier (Lanier, 2014) suggest an asymmetry wherein the service providers derive substantially more value than the users.

In the results of this research, the users frequently question this **symmetry** when they exchange their personal data for media experiences. Resonating with Lanier's take on asymmetry in the personal data ecosystem (Lanier, 2014), users see an imbalance of value exchange between the personal data collected and the services offered to them. But, while the need for symmetry in the exchange of value is made explicit, the mechanics of this notion is yet unresolved. Currently, there is no metric for evaluating the symmetry. What might seem valuable to the users might not be valuable to the service providers and vice versa. Recent arguments against the notion of

personal data being the new ‘oil’ revolves around this very challenge presented by personal data (Martinez, 2019). While the value presented by massive amounts of personal data to a service provider might be significant, the value of a standalone piece of the same, to both the service provider and the data subject remains contested. Thus, while personal data affords different forms of value to both the service providers and the users with respect to media services, the value distribution and the symmetry of exchange of value presents complex challenges around the metric of exchange and distribution, which is yet to be resolved.

Hence, value in this scheme is far from empirical representations and substantiations. It is often maybe a sense or feeling from the users side, dependant on factors like the type of data collected, sensitivity of the data collected, the context of data collection and the type, relevance and significance of the media content received in return. Here, this ‘feeling’ is significantly affected in an adverse manner by various other **underlying challenges** presented by the use of personal data in media experiences. Thus, irrespective of the metrics associated with the exchange, when the challenges associated with the exchange of personal data is considered by the users the asymmetry of value exchange is seen as pronounced. This lack of symmetry does not just affect user engagement with media experiences, but it affects the perceived value exchange to such an extend that it confronts the very uptake of future media experiences that leverage personal data. This forms a spectrum of **adoption challenges** faced by the users, which, as seen in recent HCI literatures, whereas previously treated as “someone else’s problem”, is seen as a “material concern” if technologies have to “exist beyond its prototypical implementation” and be adopted into the users’ everyday lives (Lindley, Coulton and Sturdee, 2017).

These challenges, do not apply only to the users but extend to the service providers as they voice the need for balance of the value created with the challenges presented so that users engage with the personal data exchange in a sustained manner. The complexity of this balance is further added to, because when considering this balance the results show **varying priorities** between service provider teams. Here, on one end the data maximalists ask for the maximum amount of personal data to push for maximum value through innovation and the other end sees the data minimalists who focus on the challenges, the importance of responding to them diligently and using personal data only when it is deemed an absolute requirement for service delivery. Where one sits on this spectrum of value versus challenges depends upon the type of media service catered to, the audiences served by the service, the legal and regulatory frameworks applied to the service delivered, the demands and competition the service faces, the expectations of the audiences and the resulting ethos of the team considered. This complexity renders value as almost a fluid concept

that adapts to the context of the service, the service provider and the audience considered.

Thus, the previous account introduced the complexity of the notion of value within media experiences that leverage personal data. It has been established that personal data provides value to media experiences through multifarious manners. This **value is complex** in nature as firstly it is distributed among different actors within this eco-system. It provides an explicit set of possibilities of value creation to both the service providers and users within the scope of media. Here, this distribution of value between these actors, is almost seen as an exchange. But while the creation of value by personal data is an unrivalled verity, the absence of a metric for evaluation of this value created, where currently value is an abstract notion, makes the symmetry of this exchange a challenge to be resolved. Both users and service providers are sensitive to this balance in exchange which makes the call for symmetry a significant one. Thus, value in itself presents complex challenges in the use of personal data by media experiences. This complexity is further exacerbated by a number of other considerations that affect the symmetry of exchange further from the perspective of both the users and the service providers. To the service providers, these form a set of concerns that need to be constantly balanced with the value created by the service, wherein the properties of the balance itself are variant, often dependent upon the context of the media service provider. To the users, these challenges adversely affect their perceived value of a media service, to an extend where they are not just reluctant to engage with media experiences that leverage personal data but show significant reservations to the uptake of such experiences into their everyday lives, thus presenting a spectrum of adoption challenges.

These complexities show how the notion of value within media experiences that leverage personal data is in itself a significant challenge in this scenario, often exacerbated by other underlying challenges (Trust, Privacy, Transparency, Control and Accountability). To alleviate these tensions and successfully create positive value, there is the need to study, understand and respond to these underlying challenges, thus allowing for the multifarious pathways for value creation to benefit both the service providers and end users through sustained personal data leverage in media experiences. It is towards the presentation of these challenges that we now turn.

7.2.2.Trust

A ‘crisis in trust’ due to the introduction of personal data into digital technologies was discussed in detail in Section 2.3.1. Here, the adverse impact of this loss of user trust has been presented and substantiated empirically through statistics and forecasts that highlight the scale of this issue. While these figures demonstrate the enormity of the situation, the same literature recommends that in

order to find solutions, there is the need to dynamically study the particular manifestations of trust within different specific domains and contexts (Kalapesi, 2012).

While challenges pertaining to personal data use is itself in need for more critique within media research, the currently existent discussions around these challenges is often focused on privacy. Here, while there is a brief mention of trust, by Ijsselsteijn (IJsselsteijn, 2017), he leaves the space open for future research to study how trust plays into the field of media when personal data is leveraged.

This section collates the findings from the studies reported in this thesis to respond to this gap by unpacking the abstract concept of trust to demonstrate the subtleties it presents around media experiences that leverage user personal data.

Within media experiences, when personal data comes into play, the concept of trust plays out in two ways. First is the trust associated with the nature of broadcast media, where traditionally users trusted a media experience because of its very nature of serving uniform content to all. With that media being customised using personal data, the negotiation of user trust in such scenarios becomes a challenge to be considered. While this form of trust still continues to play a crucial role with respect to the organisation's reputation, with the shift of media towards personal data use, there is also the trust around the data practices.

Currently, there is a lack of trust in the data practices within media experiences which was often owed to concern spurred by reported security breaches (Arthur, 2011) and instances of data abuse (Cadwalladr and Graham-Harrison, 2018). Similar scenarios around media services that have been using personal data (Singel, 2009, 2010) further fueled this concern. Here, service providers have the added responsibility of being a 'data broker' where they have to understand and respond to the nature of trust and challenges associated with it.

The **track record** of the service provider often contributes to user trust in this scenario, but the users do not consider that to be a guarantee when their personal data is involved. Here, the users asked for explicit and effective responses that work towards building their trust for engaging with these data driven experiences. What is brought forth here is the current tendency to force the users to place their trust in these organisations through binary consent models and convoluted Terms and Conditions statements. Thus, users identify that at present the onus is on them if they are to engage in time-consuming transparency and control measures. Instead, they call for a shift from this pattern, to be enabled with data practices where the service providers invest in solid mechanisms that build user trust. The expectation from these mechanisms is to present the users with

engaging and user friendly data interactions that encourage them to organically build and sustain trust in these services.

Thus, even if a service provider has a reputation of a trusted media provider, the introduction of personal data use not only asks for active responses to build this newer form of trust but also challenges the trust built up by the organization over years of trusted media experience delivery, thus adversely affecting their reputation and legacy.

When considering media services, the consideration of trust as a core value is central also because media services fall into the 'pleasure spectrum' of the users' lives. Hence, it presents the possibility of the user easily choosing not to engage with these services if they do not trust them.

The service providers unpacked this challenge of trust within media into the following three considerations

- Legal Accountability
- Social Accountability
- Ethical Considerations

Sensitivity towards these challenges is expected to help move closer to effective responses to the current lack of user trust.

7.2.2.1. Legal Accountability

The close correlation between the use of personal data and the law has been discussed in detail previously. Many of these regulations, particularly the GDPR in Europe (EU GDPR, 2016) are put in place to provide more means of agency to the users and as a means for employing more accountability by the service providers, with a particular focus on user trust. Service providers see the execution of such legal premises as being not just a mandatory checkbox but a method towards overcoming the challenge of loss of user trust. Users see such legal mandates and their realisation by a service provider as a reason to place trust in their services. Thus, the careful consideration and application of legal accountability becomes a challenge that when responded to presents a means for alleviating user trust.

7.2.2.2. Social Accountability

Once the legal requirements are satisfied, service providers saw the next step towards building trust was to go beyond the mandatory to provide more organisational accountability from a social perspective. This priority could be seen in line with previous incidents faced by media/ social media organisations like Netflix (Singel, 2009, 2010)

and Facebook(Cadwalladr and Graham-Harrison, 2018) that undermined user trust in this scenario. Here, by paying more attention to their influence as a powerful social entity and being responsible for the social influence of their experiences becomes a challenge faced by organisations, the alleviation of which could have helped prevent such incidents.

7.2.2.3. Ethical Considerations

From philosophers (Solove, 2008; Wittgenstein, 2009) to regulatory bodies like the EDPS (EDPS, 2015, 2018), the debate around the importance of ethics in the use of personal data is of much interest. Media service providers also acknowledge the importance of these ethical considerations, particularly when dealing with media services which exercise very high ranges of influence and reach among the average populations. Resonating with other dialogue around ethics in personal data (Rosner, 2014) (Wright *et al.*, 2009; Kudina and Verbeek, 2019), the findings here also show the importance of ethical considerations around the use of personal data when considering media experiences. Thus, within media experiences, ethics becomes a challenge, which when responded to would contribute towards building user trust in this digital economy.

While the discussion above explicates the challenge of loss of user trust within the context of media services that leverage personal data, as reported in the studies presented in this thesis, trust often also forms an umbrella for a number of other contributing challenges. While an understanding of the subtleties associated with trust is necessary for effective response to it, so is a deeper understanding of the contributing challenges involved. It is towards the explication of these contributing challenges of privacy, transparency, control and accountability that this discussion now turns.

7.2.3. Privacy

On a generic level, privacy has been marked as both a ‘threat’ and ‘challenge’ when user personal data is leveraged by organisations (Rose, Rehse and Röber, 2012). But, the contextual integrity of privacy, as elaborated by Nissenbaum (Nissenbaum, 2004), where she highlights the contextual dependance of the notion of privacy, warrants that the nature and implications of privacy calls to be explored in a context sensitive manner within the domain of media.

Within media literature, privacy was noted as a ‘predominant’ issue in various instances. In advertising, privacy compromises were found to be off-putting and “creepy” (Rapp *et al.*, 2009; Stone, 2010). In social media studies taxonomies were developed around privacy to understand what privacy means to different people within an SNS context (Jamal, Coughlan and Kamal, 2013). As personal data leverage moved to more ‘mainstream’ media platforms like the recommendations on Netflix, the privacy concerns around these

technologies were also seen to be further exacerbated (Singel, 2009, 2010).

When considering personal data use within the experience content, studies have again touched upon privacy in the context of interactive and Social TVs (Tullio, Harboe and Massey, 2008). Here, while Tulio's work talked about user privacy concerns regarding knowing who one's conversation partner is, Bernhaupt et al's ethnographic studies around Interactive TV showed that "the majority of the households were not concerned about security and privacy issues" (Bernhaupt *et al.*, 2007) as they did not see the interactive TV as very much different from a traditional one.

As the findings of the studies reported in this thesis show, quite in contrast to Bernhaupt's findings, privacy was identified as a challenge while considering the use of personal data by media experiences, when the involvement of personal data is explicitly made clear to the users. Both Trepkeviciute (Trepkeviciute, 2017) and Hook (Hook, 2018) mention the importance of 'complex issues' like privacy within media experiences that involve user personal data, but there is still the need to unpack this concept to draw out what it means within the context of media. Therefore, the following paragraphs present the various manners in which this challenge of privacy expresses itself within the scope of media experiences.

Concern around the **security** of the data collected was one that contributed considerably to the challenge of privacy. Users raised concerns over their data being hacked while service providers talked about the importance of adopting efficient security measures (like anonymisation, encryption and data minimisation) to avoid data breaches and losses. While the technicalities of security issues and responses are beyond the scope of this thesis, Section 2.3.2. presents the increasing interest in security research and how it is not the same as privacy but becomes a contributing factor to the challenge of privacy when working with personal data. The results of this study reflect the same ideology wherein security was often mentioned as a matter of concern when discussing privacy, thus making it an important aspect of this challenge to be responded to. But, privacy as a challenge, branches out beyond security to other dimensions (World Economic Forum, 2018) that calls for a more sophisticated understanding.

Owing to previous data breaches (Arthur, 2011) and data abuse (Cadwalladr and Graham-Harrison, 2018) that compromised user privacy, even within the context of media services (Singel, 2009, 2010), users in Study I and particularly Study III were concerned about who had **access** to their data. They were apprehensive about external parties gaining access to their data. They worried if their privacy would be compromised by their data being broadcast to others, which could lead to others forming opinions about them and affecting their very behaviours. Thus, effective communication of the

data flows, particularly the identification of the parties who have access to the personal data is a challenge that requires response if user privacy concerns around media experiences using personal data is to be assuaged.

Such fears were further exacerbated when the **social contexts** in which media is often consumed and hence personal data would be collected is considered. Here, two main considerations were put forward with respect to the data being collected. One was the context of the data collection. Reflecting Bernhaupt's media research (Bernhaupt, Wilfinger and Mirlacher, 2010), which associates media consumption with the living room, the results from the studies in this thesis also drew similar associations with the home. Here, the home was referred to explicitly as a 'safe space' where users did not want to be concerned about their privacy. Second was that the data collected in such contexts could **involve more than one person** because of the very inherent nature of media consumption being a social activity. This relational nature of data (Crabtree and Mortier, 2015) means that data about one person could involve data about other associated individuals as well. Within the home, this could extend to any of the various cohorts of users present there including children, elders, guests and lodgers. Personal data collection in such complex social situations becomes increasingly sensitive as the privacy of all of these data subjects are challenged.

Peculiar to media experiences, mostly owing to the social nature of media consumption, privacy risks associated with personal data use also extend to certain social challenges that require consideration. These are concerns around the possibility of **social discomfort** that could arise from privacy compromises within media experiences. Examples of such instances include, a photograph with an ex-partner being used in the media that is being consumed on a date with the current partner, sharing one's Sport team preferences in a social context when the others in the room might not be in favour of that team, etc. These **'intangible risks'** that focus on the privacy implications of engaging with a media experience in a social setting have been sporadically mentioned in previous media research. For eg., Tullio's (Tullio, Harboe and Massey, 2008) work of Social Televisions where husbands of participants posing as their wives and someone mentioning a surprise Christmas present to it's recipient led to social discomfort. These were termed as 'disruptive misunderstandings' around privacy. The results here present empirical substantiation for such social repercussions to form a novel class of privacy challenges unique to media experiences that leverage user personal data, which calls to be responded to.

Privacy fears associated with personal data use by media experiences also escalated to **projections of dystopias**. One category of dystopias centered around the fear of personal data being used out-of-context of the media experience. Here, the users feared data about them being

used in manners that might adversely affect them. Quite strongly linked to the previously discussed concern around access was data being sent to external parties who might use it in unethical manners, like a burglar using the data to check when someone is away from the home or a recruiter using it to assess one's personality. Even when the data is not sent to external parties, there is still the fear of the media organisation using the data to form echo chambers filled with biased content that is intended to manipulate the users' attitudes. The second category of dystopias were around the fear of surveillance. The collection of personal data within personal spaces, without proper response mitigation, thus presenting the potential for compromise of user privacy were considered 'creepy' by the users, service providers and previous research (Stone, 2010) alike. These fears sometimes exacerbated to dystopias where users projected scenarios of regulatory bodies monitoring them or even the camera of their smart television recording their responses to certain content to evaluate their attitudes. If the challenge of user privacy is to be alleviated, substantial measures that circumvent the disruptive impacts of these dystopias are to be put in place around media experiences that leverage user personal data.

Thus, the challenge of privacy in itself displays a range of expressions within the scope of media that uses personal data. From the very technological concerns around security to the various considerations of privacy within social contexts to even the projection of dystopias, shows the significant scale of impact this challenge could unleash, if not carefully mitigated.

7.2.4. Transparency

Provision of transparency of data practices is constantly brought up as a means of alleviating concerns within the scenario of data driven technologies. Traditionally, the importance of transparency of data flows has been proposed and discussed much in law (EU GDPR, 2016), academia (Krishnan, 2016) and industry (Kalapesi, 2012; World Economic Forum, 2018) research.

Despite this longstanding interest in exercising transparency of data practices, mechanisms that serve the same to the users still seem to be ineffective. The results of the studies reported in this thesis confirm the user concern of a current lack of transparency that is felt when engaging with media experiences that leverage personal data. Transparency was discussed by both users and service providers, as a key to alleviating their concern of loss of user trust. When both parties highlight the same entity as the key to a better data exchange scenario, the question of why it still fails to be executed efficiently stands strong.

A classic example here would be the Terms and Conditions statements, which while familiar to most average technology users, was found to be ineffective in functioning as a tool for data

transparency (Jones, Sailaja and Kerlin, 2017; Sailaja and Jones, 2017). These long and legalese (McDonald and Cranor, 2008; Cate, 2010) agreements, while presenting all the information that is to be communicated to the users, often fail in their effectiveness at communicating the presented information to the users. Despite research in both industry (Zynga, 2011) and academia (Kelley *et al.*, 2009; Schaub *et al.*, 2015) there still remains the need for effective ways of information communication when it comes to Terms and Conditions agreements.

What could be seen here, is a gap between provision of transparency and its effectiveness. This is where research turns focus from just transparency provision to ensuring effective transparency provision. Human Data Interaction, a line of thought that sits at the confluence of the digital economy, computer science, psychology, sociology and HCI among other domains, talks about the concept of data legibility (Mortier *et al.*, 2014). Legibility, while focusing on the provision of transparency, subscribes to efforts that make transparency effective rather than just legally-abiding.

With the intention of realising legibility of data within a data driven media experience, the LRoTF in Study III was consciously designed to promote a 'data dialogue'. This initiative was also inspired by the WEF's report where 'effective dialogue' was seen as a solution to the loss of user trust in the digital economy (Kalapesi, 2012). While the realisation of such dialogue at a very practical level or its effectiveness was not in scope of their discussion, this research sought to look into exactly those through the design and inclusion of legibility mechanisms within the LRoTF. This dialogue was served through multiple modes and at multiple temporal points, reflecting previous research in useable privacy around data communication (Schaub *et al.*, 2015) while also leveraging the reflective nature of a media experience by interleaving this dialogue with the experience itself.

Response to these data dialogue initiatives and other conversations conducted in the previous two studies have helped add more detail to the concept of transparency in media experiences that use personal data, thus helping fill in the space opened up by Trepkeviciute's abstracted identification of responses to data transparency within media as recommended but 'complicated' (Trepkeviciute, 2017). The following paragraphs render this more nuanced definition of transparency within this context, thus shifting from a generic abstraction to the various manifestations of this 'complicated' challenge within the scope of media experiences leveraging personal data.

Firstly, transparency when considering a media experience that uses personal data was presented at three levels. The first was transparency of the data flows and data management outside of the experience to inform the users of who has access to the data, where it flows to and

where it is stored. The second was transparency of data flows within the experience wherein the users wanted to know what data was being collected, for what purposes and how it was being used to affect their media experience. The third was a mix of the two wherein transparency was expected to help surface issues within a media experience around personalisation like filter bubbles, echo chambers or personalisation of broadcast content which could be initiated by factors outside of the media experience leveraging user personal data to constrain the users' exposure and reach to diverse content and maybe even bias their attitudes.

Irrespective of the aspect of transparency discussed, audiences wanted to know more about the data collection, asking for data legibility at a granular level, so that it seeks their attention. But, particular to a media experience is the expectation of data legibility measures to not just stop at showing what data is being collected at a granular level but to extend it to present the **'effect' of that data on their media experience**. On a generic level, such requests are often made by users to know the purpose of the data collection, to check if their personal data was being collected in a superfluous manner. But, more uniquely, for a media experience, it is to know how they, the users, were affecting or rather even crafting the experience through their involvement. Thus, data legibility measures within a media experience becomes a source of communication for the users with the experience as they gain feedback about their data which in turn informs their future interactions with it.

When this causality is not legibly presented to them, within the experience users feel a 'gap'. And when this gap is responded to, thus effectively exercising data legibility within the experience, users expect their holistic media experience to improve in many ways. Firstly, revealing the underlying data practices within a media experience helps the users **understand** the experience better. Secondly, such knowledge equips the users better to **engage** in more **informed interactions** within the experience, making the media experience more **meaningful** to them. Thirdly, not knowing the 'cause' and 'effect' of the data flows leads to **confusion** and **stress** while consuming the experience, leading to user discomfort and lowered engagement as they question themselves about their interactions constantly, rather than be assured of the consequences of their actions. Fourthly, when such data legibility is not offered within the media experience, the users, being unable to understand the holistic nature of the data flows, their interactions and its effect on the media, confessed to feeling a lack of control. Fifthly, while a lack of effective data transparency causes all the aforementioned issues around interaction and engagement during the consumption of the media experience, the effect of the 'gap' pervades even post the experience, where the absence of the cause to effect data flow communication leaves the users short of understanding how the holistic experience was shaped for them, by them, through their

informed interactions and engagement with the data driven media experience.

One of the major challenges around transparency is the user notion of service providers not investing enough into effective data legibility measures. Here, they feel current measures are often convoluted, placing **the onus on the users** to expend extra effort to understand the data practices. Research has shown that such effort, if chosen to be invested by users is significant as an average man would be spending 244 hours a year reading privacy policies if they chose to do so (McDonald and Cranor, 2008). Such extensive demands demotivate users from engaging in such efforts, lowering engagement and trust in the services. In the case of media experiences, it may even affect user engagement with the experience itself due to its position in the 'pleasure spectrum' of the users' lives, where it is not a mandatory necessity.

Here, previous design-oriented research has proposed a number of alternatives to the current text-rich Terms and Conditions statements including labeling (Kelley *et al.*, 2009), warnings (Bauer, Bravo-Lillo and Cranor, 2013), iconography (McDonald and Cranor, 2008), controlled timings (Good *et al.*, 2007) and more customised presentations that suit the users' context and interest (Pinnick, 2011). But, as mentioned earlier, their widespread success is yet to be witnessed as common practice is still to present the traditionally long and legalese data agreements (McDonald and Cranor, 2008).

Thus, the turn to personal data in media experiences does not only present data transparency as a challenge. But, it also provides fertile grounds to incorporate and test these methods and expand into newer territories of data legibility wherein like in the LRoTF, rather than limit the user to upfront data agreements, users could be engaged in a 'data dialogue' that is interleaved into the media experience itself.

7.2.5. Control

When considering data transparency as a challenge within a personal data use scenario, user agency and control becomes a natural part of the discussion. In previous research within media Benson et al. (Benson, Saridakis and Tennakoon, 2015) found a negative correlation between information disclosure and perceived control, which is easily justifiable because if the users are made aware of the underlying data practices but not empowered with sufficient control mechanisms to exercise agency within that setting, the greater the knowledge they accumulate, the more powerless they will feel. Hence, user control occupies a premier position along with the challenges of data transparency when studying the use of personal data in media experiences.

Here, studies both in media (Trepkeviciute, 2017) and other digital economy domains (Rose, Rehse and Röber, 2012) have suggested

effective measures associated with the challenge of control as a potential response to user concerns as studies have shown that users are 52% more willing to share their personal data if they are given the means to manage and control it. But, the results of the studies reported here show that within media experiences that leverage personal data, users express a lack of control, which calls to be studied with diligence.

Jaron Lanier talks about an asymmetry of control in personal data ecosystems where the users are placed in "a structurally subordinate position" (Lanier, 2014). In order to overcome this reported challenge associated with user control (which is also echoed in the findings of Hui et al. (Hui, Hai Teo and Tom Lee, 2007)), this asymmetry calls to be balanced by **giving the users a more central role to play** in media experiences that leverage user personal data. Van Dyke supports this work towards more balance by handing over more control over user personal data to the users as they proposed the notion of 'privacy empowerment' wherein they show that empowering users with control lowers concerns around privacy and increases user trust (Van Dyke, Midha and Nemati, 2007).

The possibility of user centered consent models was proposed as part of the conversations here wherein the user engages more actively with the data, exercising choice and informed control throughout the data exchange, thus giving them more control within this space and making them a more integral part of the turn towards personal data in media. But to apply such an intervention calls for the understanding of the challenge of control in detail, explicating its subtleties as it unfolds within the emergent space of data driven media. It is towards presenting such nuances, as uncovered in the studies reported here that this discussion now turns.

Control within media experiences that use personal data were talked about from different perspectives. Firstly, the ability to exercise direct **control over the media experience** so as not to feel a sense of lack of agency. The users want this form of control to be effortless and in line with the experience as they asked for frictionless interactions using voice control. Secondly, the **control over the data** that in turn controls the experience. This control was deemed as currently unavailable to the users and they asked for it so that they have agency both within and without the media experience. Thirdly, when such personal data is collected from the home to drive media experiences, users asked for **a sense of autonomy** or overall control over the scenario. Here, the control includes the media experience, the data flows and also control over other external elements that form part of this eco-system like possible IoT elements, e.g., sofa cushions, window blinds, lights, fans and other everyday mundane items. Fourthly, the discussion **around filter bubbles and personalisation**, which forms a combination of having control over the data and through the data, the experience itself, became another prominent

topic of discussion where the fear of being constrained by previous choices, losing serendipity of finding new content, being repeatedly exposed to the same kind of content and and the fear of being biased, influenced and manipulated through such personalisation of media were highlighted as another aspect of the challenge of the users not having agency within data rich media experiences.

The home has been seen to be often associated with media consumption from previous literature (Bernhaupt, Wilfinger and Mirlacher, 2010). The findings of this research show that the concerns over data control are further amplified when such **social scenarios** arising from media consumption in the home is considered. The context of the home produces complex combinations of different cohorts of users with varying characterestics. They all have an equally diverse set of everyday mundane activities they exhibit within the personal space of the home, which is exacerbated by the call for effective user control in media experiences that leverage personal data. Here, control was required firstly with respect to the proper management of personal data collected from these individuals to ensure the right kind and grade of data practices is applied to each user, as per their preference. Secondly, control is also expected with respect to the translation of the personal data to the resulting media experience, to negotiate the best experience that would suffice the expectations of everyone present.

Here, control was seen as a three step process. The first being the understanding of the data processes and its effects on the audiences' experience and life (also discussed in Section 7.2.4.). The second being the capacity to negotiate in these circumstances by being able to make more granular decisions regarding data exchange in such scenarios. Third, the ability to exercise agency through control of the data and experience whenever necessary. These considerations are very much in line with the principles of HDI (Mortier *et al.*, 2014), thus confirming the applicability of the principles of data legibility, negotiability and agency within the context of media experiences.

Within broader research, the nature and definition of data control is an area that has seen quite some contribution in the recent past. These range from provision of opt-out (Malhotra, Kim and Agarwal, 2004), allowing for user permission and control over data sharing (Culnan and Armstrong, 1999; Hann *et al.*, 2002) and giving users the space to object to uses of their personal data outside of purposes it was collected for. (Culnan and Armstrong, 1999).

The findings of this research provide better resolution to these findings to define what is expected of data control, as it manifests within the scope of media experiences. For control of personal data within media experiences to be effective, it was sought to be **more granular** in nature, with the capacity to enable users with the ability to make decisions regarding even the smallest bits of their personal data. But while such granular control is an expectation, also to be

considered is previous research around privacy paradox which has shown lowered user engagement in privacy preserving scenarios (as opposed to user attitudes). Also, research has shown that individuals who found data control easier were 37% more willing to engage themselves in data exchange. Results from the studies reported in this thesis also resonate with these findings, expressing the user expectations of **simple, easy and quick interactions** with data so that they are able to maximise their time experiencing the media content itself.

These contradicting expectations of granular control versus the need for quick interactions is owed to the position of media consumption being in the 'pleasure spectrum' of the users' life where they engage with it for entertainment and relaxation. Hence, they are happy to welcome the enhancements provided by personal data but are not willing to expend significant amounts of their time managing their data for this purpose. Hence, the possibility of different levels of control becomes an alternative wherein the users are presented with the minimal amount of active user control upfront with the underlying more granular options being made available as per the user's discretion.

These practicalities of control, as they are manifest in media experiences that use personal data was expected to be applied through one of two approaches. The first is the traditional method of explicit data control while prioritising quick and easy control mechanisms for both the data and the experience. The second is to leverage the experience itself for data control, by interleaving the experience with options for control.

Emergent technologies like PIMS that allow for more nuanced user data control by giving the users a more central role in this context, thus responding to many of the challenges around data control presented here was discussed in Section 2.3.5. Such devices present the possibilities of accommodating these complex demands of control. Taking the example of the DataBox (Amar, Haddadi and Mortier, 2016) which was leveraged within the LRoTF, it was considered a 'gatekeeper' or a 'metaphorical filter' that kept the users' personal data safe from the rest of the internet. Here, the DataBox, through client side processing of personal data for media customisation and deletion of the personal data post the experience abstracted away the need for the users to granularly control what bits of data to share with others, while affording the enhancements offered by personal data. It also presented the possibility of a single device being able to form a hub for all personal data activities, thus constraining the idea of control to a central point rather than the user having to switch between individual providers and devices. Also, the presence of such a physical box was expected to be a constant reminder to the users of the control that is given to them in this shift towards personal data in their media experiences. While, this account

is not a recommendation of the DataBox as the solution for the future, it is used here as an example of a response that accommodates many of the nuances of control within data driven media experiences presented here.

7.2.6. Accountability

When considering personal data leverage, accountability has been a much debated challenge wherein interest from diverse disciplines have rendered different perspectives and definitions for it. Section 2.3.6. details these engagements with accountability, by laying out the approach taken towards this challenge from law and computer science.

While previous research around media is yet to explicitly identify accountability as a challenge, **the results of the studies reported in this thesis do highlight accountability as a challenge** when media experiences turn towards personal data leverage. Due to previous experiences with other data driven services, users reported not being satisfied with current levels of data transparency and control where they are forced to trust the service providers as opposed to being encouraged through adoption of appropriate responses to organically build trust. Media service providers also report accountability as a challenge by explicitly pointing out the different dimensions of accountability that must be considered if user trust is to be preserved. Here, to state at an abstracted level, a lack of accountability with respect to data practices is demonstrated. How this challenge of accountability plays into the scenario of media experiences driven by user personal data in multifarious ways is as follows.

In legal literatures, accountability is seen as one of two, internal or external accountability. Here, "**Internal accountability** refers to the policies and procedures that a 'data processing' organisation puts in place to demonstrate to itself that its data processing operations comply with the requirements of proposed data protection legislation" (Crabtree *et al.*, 2016). Service providers' conversations around using technology to secure the personal data collected through data minimisation, anonymisation, encryption and other data security measures, all fall into this form of accountability, which forms a mandatory basis for other accountability considerations to build upon. Owing to the nature of enquiry of this research, further technological detail associated with this form of accountability is out of scope for this discussion.

Within media experiences, internal accountability is prioritised particularly in personalisation and formation of user cohorts. In the U.K., not every user is to be subject to the same grade of personalisation. The same applies to the type of personal data that is collected from the users, which is dependant upon the user's age. As media experiences increasingly start leveraging user personal data for personalisation, both personal data collection and the resulting

delivery of personalisation calls to be managed with precision and judiciousness in order to be legally compliant. Formation of user cohorts based on age is currently applied as the means for managing this situation with age based cohorts deciding the right grade of personalisation to be made available to each user and managing age transitions automatically, to help maintain internal accountability.

"**External accountability** refers to the policies and procedures a data processing organisation puts in place to demonstrate to others outside the organisation that its operations are in compliance with the requirements of proposed data protection legislation" (Crabtree *et al.*, 2016). Most user conversations around accountability in the studies reported here primarily fall into this category of data accountability.

The first form of external accountability spoken in the studies presents the notion of 'providing an account of' the underlying data practises. HCI research shows much interest in the realisation of this form of legal accountability which focuses on the legibility provided by the computer system (Dourish and Button, 1998). Sometimes this form of accountability is termed 'computational accountability' (Crabtree *et al.*, 2016), to help mark the difference in perspective from that of other perspectives on accountability where the focus is on justification and responsibility. Bellotti and Edwards's (Bellotti and Edwards, 2001) work, which proposes a design framework for accountability extends the scope of this form of accountability further by including both informing the user and control as the two major aspects of provision of accountability.

With the introduction of the GDPR (EU GDPR, 2016) in the EU, such **computational accountability** also becomes a priority that makes organisations legally accountable to regulatory bodies. Thus, the discussion around provision of accountability through transparency and control speaks to both these notions of accountability previously considered within law and HCI. The results here ground these broader values from law and HCI to show that when considering data transparency and control provision, both these aspects of accountability is applicable within media experiences. Within media experiences, a clear chain of accountability of who has access to the data collected, the intended processing to be conducted on it and the use of such data collection, particularly the inferences from it and the effect of the personal data on the media experience is requested to alleviate current challenges around accountability. Such responses speak to both the need to cater to legal requirements to suffice certain mandates, as part of being internally (legally) accountable and the expectation of communicating these data practices effectively to the users to maintain external (legal) accountability, through provision of computational accountability.

Thus, as challenges of transparency and control within data driven media experiences are worked upon, they simultaneously respond to different aspects and perspectives of data accountability. Here, to

illustrate with the example of provision of transparency, when transparency is effective, the service providers are sufficing their requirement for internal legal accountability to the regulators. By doing so, they are also responding to the external legal accountability requirements by communicating data practices to the users and they are utilising the notion of computational accountability for this endeavour to realise effective methods of providing 'an account of' the data practices.

These findings show how the different siloed perspectives of accountability, as seen from diverse disciplines are manifest in the context of data driven media in complex manners. It also shows how these siloed perspectives on accountability, within the scope of media experiences, form overlaps, thus presenting opportunities for working towards a holistic response towards the overall challenge of accountability.

The aspect of accountability that involves responsibility introduces yet another dimension to this spectrum, that of **social accountability**. This view on accountability, while not mandatory, is an example of the service provider understanding of the need to go beyond just what is required, to ensure user trust within the digital economy in media. It refers to being accountable as a societal entity, respecting and mitigating the social implications of using personal data that could otherwise lead to loss of user trust in the organisation. This kind of social accountability has been spoken of by Nilsson et al (Nilsson *et al.*, 2019) where it is defined as the form of accountability that "*extends beyond expectations concerning the legibility of autonomous behaviour to expectations concerning the appropriateness of autonomous behaviour within a social setting*."

Within the context of media, social accountability holds a central position due to the inherent social nature of the activity of media consumption in users lives. Here, the importance of providing data driven inferences, recommendations and personalisation that is accommodating of these social circumstances, thus helping alleviate 'intangible risks' presented by the data collection is a key aspect of provision of social accountability. The management of user cohorts, while a tool for delivering improved services to the users becomes a point of tension when considering social accountability due to the possibility of making erroneous predictions and bad stereotyping. Here, careful consideration of socio-technical implications, such as the ones mentioned here and others like spreading awareness, and giving enhanced data control to the users are called forth to ensure better social accountability from the service providers' side.

This discussion unpacks how the challenge of accountability manifests with media experiences that use personal data. It presents the different perspectives on accountability, as described in previous work in law and computer science and grounds the findings of this research within this scheme to show how the different perspectives

on accountability apply here. It also shows how these siloed perspectives on accountability, when satisfied, form overlaps, thus working towards a holistic response towards the challenge of accountability.

The discussion further extends to present social accountability which goes beyond regulated mandates to consider the social challenges around media experiences leveraging personal data, that the service providers could be sensitive to. Due to the close association of media experiences with the home and the everyday mundane life settings of the users, the nature of the data here often tends to be considered social, sensitive and personal, making the users reluctant to engage in the use of their personal data in media experiences. Here, the consideration of these various manifestations of the accountability call is necessary to alleviate this multi-faceted challenge in a holistic manner.

7.3. Implications for Design

Having detailed the challenges of using personal data in media experiences from a conceptual perspective, the discussion now moves on to summarise these challenges to recommend the design implications presented by them.

When considering media experiences that leverage user personal data, the need for symmetry of exchange of value has been highlighted. The different ways in which personal data presents value to both the user and the service provider has been described in detail in Sections 5.6.1. and 7.2.1. But these multifarious forms of value are constantly confronted by a number of challenges arising from the use of personal data. The conceptual characteristics of these challenges which include value itself, trust, privacy, transparency, control and accountability have been unpacked and presented in Section 7.2. The following section utilizes these learnings to formulate a set of design sensitivities that respond to the challenges of using personal data in media experiences. These design sensitivities, reflecting the structure of the conceptual explication of challenges undertaken here, fall under the following categories :

- Trust
- Privacy
- Transparency
- Control
- Accountability

Thus, in the following section, the previously conceptually presented challenges are looked at from a design perspective. These learnings

also combine the viewpoints of the users and the service providers to present design recommendations that are sensitive to the challenges raised by both stakeholders. These guidelines are compared with current popular responses to these challenges within wider literature, to either build upon it or to highlight alternatives that present the possibility of more effective alleviation of challenges voiced by both the users and service providers within the scope of media experiences leveraging user personal data.

7.3.1. Trust

When involving user personal data in media experiences, loss of user trust has been highlighted as a significant challenge. Hence, while designing such experiences, investing in measures that contribute to building and maintaining user trust should be a consideration to be prioritised.

Within media experiences that use personal data, trust is not 'a single index' and hence the responses to loss of user trust cannot be unidimensional. These responses have to speak to the two forms of trust that is manifest here : that of being able to **deliver trusted, unbiased media content** and the newer role of a **trusted 'data broker'**.

The following three broad considerations were discussed within the studies to help respond to the challenge of user trust :

One of the key steps to being trusted as a data controller is being **legally accountable** by meeting legal measures mandated by the likes of the GDPR (EU GDPR, 2016). This includes data protection by design and default (Article 25), data impact assessment (Article 35), and the information to be provided to data subjects (Articles 13 & 14) that is typically contained in a service provider's privacy policy, which are all expected to help build user trust in the data practices. Also specific to media experiences is to go beyond the privacy policy, to realise requirements such as age specific data collection and media personalisation and effective use of data transparency even outside of such agreements such as the privacy policy or T&Cs were seen as specifically important guidelines to help build user trust within the space of media.

But, as mentioned previously, trust in media experiences also extends to the delivery of trusted content. For this trust to be enabled, responses that cater to more than what is mandated is called for. Hence the consideration of being **socially accountable** which involves being sensitive to the socio-technical challenges imposed by a media experience when it leverages user personal data is important. Such measures manifest in two categories, the **first** involves looking into the various possible social scenarios that the experience could lead to and minimizing the risk of situations of social tension and discomfort brought about here by the involvement of personal data.

This is done by eliminating the possibility of out-of-context decision making, erroneous stereotyping and delivery of irrelevant content in social scenarios. The **second** is for the organisation to take responsibility as a social entity, by building a reputation that explicitly pioneers data processes that minimises casualties to the users by promoting better user awareness and agency within the digital economy to build and maintain user trust. The LRoTF was an example of such an initiative, which through the inclusion of the 'data dialogue' and the DataBox built an environment of increased user awareness, agency and accountability with respect to the underlying data practices within the media experience.

Responding to both these aspects of trust is the possibility of including **ethical considerations** to both the media delivery and the data practices. Dialogue around ethics in personal data has mostly been high level, philosophical (Rosner, 2014) or mostly around privacy (Wright *et al.*, 2009; Kudina and Verbeek, 2019). The following accounts picks up this spirit of employing ethical considerations within data driven technologies to specifically respond to the socio-technical challenges arising from the use of personal data in media experiences.

Starting with media delivery, **bias** is a major ethical dilemma that was surfaced in the studies. Two kinds of biases are talked about in media experiences leveraging personal data. One is the possibility of **serving content that might be biased** to influence user attitudes, derived from the knowledge given by their personal data. Second is the possibility of constraining the users in biased recommendations and user journeys like **filter bubbles and echo chambers**, thereby limiting their diversity of exposure. Ethical consideration and response to both these aspects of bias within media is expected to contribute towards user trust.

Ethical responses also play out in other ways in this scope like the sensitivity towards **delivering age appropriate content** in shared social spaces and devices that might be exposed to vulnerable populations like children. Also, to consider the unique nature of personal data as an asset class that is not fungible and hence to exercise the ethical responsibility of **maximising the value returned** for the personal data collected through diversification of the feedback and experiences returned, so that the users feel a sense of symmetry in the value exchange. These include data feedback mechanisms, data dashboards that show the users information about their consumption and quantified self initiatives (Lupton, 2016).

Looking at the data practices, ethical responses that form active measures that seek to continuously empower the user through the turn towards personal data use is considered key in building user trust. Here, important steps would be prioritising ethical ways of handling personal data throughout the design cycle by applying data minimisation, justification, communication, user agency and use of

initiatives like Privacy by Design (Cavoukian, 2009) that **prioritise the user** throughout the making of the experience. Also identified is the potential of seemingly '**alternate**' approaches to data use that are more inclusive of the user's priorities, interests and agency and give the users a more central and participatory role in the data processes. These include a public service approach to data (BBC R&D, 2018) or more client side processing, which require minimal 'collection' of data by the organisations' servers. While the upsurge of such alternatives have been discussed previously (Section 2.3.5.) at high level, generic scopes, the application of the same in mainstream media delivery was seen here as an ethical possibility that would engender user trust.

While these handful of design sensitivities help initiate the process of building user trust, it was seen that trust as a challenge is in itself dependant on a number of different challenges like privacy, transparency, control and accountability. Responding to these challenges in turn contribute to building of user trust. Hence, it is towards the presentation of these challenges, in a way that speak to future design that this section now turns.

7.3.2. Privacy

Previous research has invested in responses to privacy challenges through initiatives like Useable Privacy (Iachello and Hong, 2007) and Privacy by Design (Cavoukian, 2009). While these responses provide more generic methods for working around privacy concerns, the following section presents the subtleties to be taken into consideration when designing to mitigate for privacy challenges within data driven media experiences in particular:

The close association of media consumption to the users' homes (Bernhaupt, Wilfinger and Mirlacher, 2010) makes the kind of data collected **sensitive** and capable of violating the users' '**safe space**'. Also to be considered is the **relational nature** of such data (Crabtree and Mortier, 2015) collected in the home, which could also involve information from other individuals, often belonging to different cohorts with respect to their role in the social setting. Therefore, considering privacy responses that are inclusive of these properties, which make them sensitive to the space, setting and people involved in the data processes and media consumption is important.

The combination of the **context of the home** and the **social nature of media consumption** within that space also introduces 'intangible' privacy risks associated with scenarios of discomfort produced by the media experience, due to context insensitive use of personal data. This type of privacy compromise, which is often local to the social setting experiencing the content is one that is unique to media experiences and one that calls for careful consideration when designing such experiences.

Privacy challenges also result in potentially **dystopian impacts** of surveillance, manipulation and danger which bring about reservations in the users minds against such technologies. Here, concrete assurances that help debunk such myths to alleviate privacy challenges and engender user trust are called forth to respond to this challenge.

7.3.3. Transparency

Data transparency is a much debated challenge when considering personal data use by technologies and responses to it has been looked into through law (EU GDPR, 2016), useable privacy (Iachello and Hong, 2007) within HCI and industry research (Zynga, 2011). But, effective transparency still remains a conundrum with the cliched example of the Terms and Conditions statements still being incomprehensible to the users (McDonald and Cranor, 2008). To alleviate this challenge of transparency, the following account presents a detailed set of characteristics of data transparency when considering the design of media experiences that leverage user personal data.

Provision of transparency within media experiences are more effective when data legibility (or measures for *effective* transparency) is enabled through building a '**data dialogue**' with the users rather than serving the traditional text rich data agreements. Within media experiences this dialogue can be enabled in many ways :

It could be served in **multi-modal** manners, like the data announcements displayed on the data monitor or a physical data receipt that is handed to the users. It could be served at different **levels of depth and granularity** to suit the context and preference of the users like the granular data announcements on the data monitor versus the overall summary provided by the data receipt. It could be served in the **most appropriate timings** rather than forcing the users to go through an upfront run-through, like the data monitor giving real time announcements of the data collection or the data receipt serving a summary at the end. Schaub et al.'s work on the design of Privacy Policies has highlighted similar reasonings around modes and timings when designing privacy policies (Schaub *et al.*, 2015). But, unique to media experiences is the opportunity to leverage the media experience itself to **interleave** the 'data dialogue' into it so as not to make the users feel like they need to expend extra energy to engage with their data interactions, like the data monitor's script being designed to flow with the users' interactions within the LRoTF.

Data legibility within a media experience extends beyond knowing just what data is collected. Here, the call for enhanced data legibility by also knowing what the collected data was used for, or **the impact the users (through their data) had on the experience** was put forth. Not knowing this connection between the cause and effect of the personal data collected on the resulting media was said to leave a

'gap' in the experience as knowing their impact was what brought more relatability and meaning to the personalisation that was applied to the experience.

Concrete and effective responses to data legibility are suggested because disruption of this communication does not just interfere with user engagement with the data but within a media experience it affects the users' understanding, engagement and immersion with the media content itself. Thus, a 'gap' in this flow or an ineffectively designed transparency measure that demands too much user attention, results in the users getting **confused** and **distracted** within the media experience, negatively affecting their holistic media consumption experience.

Within media experiences data legibility is also extended as a solution to the ethical concerns of **bias**, both around recommendations, filter bubbles, echo chambers and general curation of content that is presented to the users. By effectively communicating the data collection, the previously discussed cause to effect translation of data and the resulting data decisions that led to the media that is being consumed, user concern over such biases could be alleviated. Here, extending upon previous research around filter bubbles which have been looking into alternate ways of presentation (Park *et al.*, 2009; Munson and Resnick, 2010) and processing (Faridani *et al.*, 2010; Kriplean *et al.*, 2012) to check their adaptability to the design of media experiences that leverage personal data would be a good starting point when working towards such responses.

7.3.4. Control

Control within data driven technologies has been defined in many different ways in previous research (Section 2.3.5.). These include recommendations of opt-out (Malhotra, Kim and Agarwal, 2004), allowing for user permission and control over when their data is shared externally (Culnan and Armstrong, 1999; Hann *et al.*, 2002) and giving users the space to object to uses of their personal data when it is used for purposes other than what it was collected for. (Culnan and Armstrong, 1999)

While these notions of control *propose* for more space for user expression when it comes to data control, the following account describes the different dimensions of *practicalising* the said recommendations around the challenge of user control when designing media experiences that leverage user personal data.

User control manifests in four different forms, which have to be given due attention when considering personal data driven media experiences :

The first is the sense of agency **within the media experience** through direct and easy control like using remote control or voice actuated

interactions that help the users control the progress of the media experience.

The second is the sense of **control over the data** that is collected and used for the media experience, which is currently assumed as unavailable by many users.

Third is unique to media experiences that are consumed in the home where physical immersion using everyday mundane objects in the home is leveraged by the experience. Here, control over both the data and the media experience along with the different elements or actuators of the experience is deemed important for the users to have **a sense of autonomy** in the home.

Fourth, closely following legibility, control is also an important aspect of dealing with the concern over **filter bubbles and echo chambers**. Here, providing the users with ample opportunity to change the recommendations and personalisation served to them, if necessary becomes a key aspect of provision of control within media experiences that use personal data.

Control is also to be considered in **social scenarios** of media consumption. Options for engaging in appropriate methods of control over the personal data and to recommend or negotiate the right kind of content that is relevant and appropriate to everyone present becomes a design challenge to be explored.

While these different dimensions of control are surfaced and the warrant for inclusion of these considerations are made, media being part of the 'pleasure spectrum' of the users often simultaneously present the issue of users not being willing to expend significant amounts of their time exercising said control. Thus, when including these different manifestations of control, while there is the need for granularity of control, there is also the need for these data control interactions to be **easy, simple and quick**.

7.3.5. Accountability

The many subtleties associated with the challenge of accountability within the scope of media experiences driven by user personal data has been presented already. Continuing on from that discussion, the following sensitivities are to be taken into account when designing to respond to accountability challenges :

The provision for **internal legal accountability** by ensuring compliance to legal mandates. Within media experiences, these include measures like ensuring the right grade of personalisation through the use of user cohorts. (Section 7.2.2.1.)

The provision for **external legal accountability** by effective communication of the data practices to the users by being transparent about the data practices. (Section 7.2.2.1., 7.2.4. and 7.3.3.)

Using provision of **computational accountability** as a mechanism to enable external legal accountability by designing in data legibility measures. (Section 7.2.4. and 7.3.3.)

Working on provision of **social accountability** challenges by considering the social implications of the data use within a media experience and building in mitigation measures and control mechanisms to assuage challenges that arise here. (Section 7.2.2.2. and 7.3.1.)

To respond to the holistic challenge of accountability, a comprehensive understanding of the different user and service provider perspectives on accountability as detailed in the conceptual contribution is a starting point. This is then to be followed by careful consideration of these different aspects from a design perspective, as laid out above, to work towards alleviation of each of these different aspects thus resulting in a comprehensive response to this significant challenge.

The responses to data usage within the LRoTF were seen as effective design interventions that sought to respond to this challenge of accountability along with response to other challenges like transparency, control and privacy. The ‘data dialogue’, through the data monitor and the data receipt helped expose information about the data collection within the media experience. Here, the data monitor was seen as a part of the media experience, helping the users with their understanding of the experience and their interactions within it. This awareness helped respond to challenges around transparency and through it, concerns around user privacy. The **principles underlying the DataBox** were engaged with in the ‘data dialogue’ during the interviews. These conversations showed that the client side processing, with no data sharing with external servers, was seen as an effort to being **more accountable**.

This accountability was enabled from the **legal** perspective, by minimising the scope of casualties presented by the leverage of user personal data through client-side processing and post-experience deletion of user data. From the **social** perspective, the users felt the box as a representation of the service providers’ accountability as a social entity, towards the users. The concept of client side processing of data, with minimal sharing and deletion at the end of the experience also helped with alleviating concerns around privacy as data was no longer sent to anyone else. This also abstracted away the need for the users to actively **control** and limit the sharing of their data at a granular level, as the data was being disposed at the end of the experience after it was processed within the DataBox during the runtime of the experience. This combination of active measures that

responded to the challenges associated with personal data use made the data processing seem more **ethical** to the users.

These legal, social and ethically sensitive responses provided by the 'data dialogue' and the DataBox principles helped with engendering more user **trust** within the LRoTF space by responding to the challenges presented here, (though their response to adopting such technologies into their homes were different because they did not see such responses to the data practices within their homes currently), thereby making the experience seem valuable in return for user personal data.

While current mainstream responses around privacy, transparency, control and accountability (Sections 2.3.3 – 2.3.6) are designed around cloud based models, these frameworks form default triggers to many of the challenges presented here. They put the personal data and the users at risk from the point of collection of the data through movement of the data to external servers, thus adding in the need for enhanced user control of these shares and exacerbating concerns around privacy of the data collected and accountability of these processes. The responses to the experiments presented in these studies propose the possibility to move towards alternative models of data processing like the DataBox (Amar, Haddadi and Mortier, 2016) or other PIMS (Section 2.3.5.) that shift the locus of this process more to the users' side. Such a shift, as seen through this detailed discussion, helps alleviate challenges presented by both the users and service providers around trust, privacy, transparency, control and accountability, thereby making the data driven media experience received in exchange seem more valuable and the exchange more symmetric.

This proposal, while might have been considered radical in the past, is seeing increasing interest recently as technologies that shift the load of data practices to the client side are being taken up by industry giants. These include Google's use of federated learning (McMahan and Ramage, 2017) on their new phone OS, Apple's commitment to privacy through on-device intelligence that makes News recommendations on phones private, "never sharing it with others" (*Apple News - Apple*, 2019) or the BBC Box (BBC, 2019) which is essentially the DataBox, built to work with BBC applications, which could be placed in the audiences homes, thus helping the organisation deliver its audiences the benefits of personal data leverage while not having to collect, transfer or store personal data on external servers.

While these are not exact realisations of the design sensitivities identified in this thesis, this increasing shift towards on-device processing is a start towards service providers seeking alternatives that employ more user-centered data practices, which resonates with the core proposal of this thesis. Particularly for media experiences leveraging personal data, the conceptual contributions and design sensitivities outlined here could be tools to help navigate this

complex shift by unpacking and presenting the practical subtleties associated with it.

8. Conclusion

This thesis explores the challenges of using personal data in media experiences. Modern media is turning towards personal data to help enhance the possibilities for innovation in service provision. But while the capabilities presented by personal data are manifold, it has also been seen to introduce many socio-technical challenges. Currently, these challenges see limited critique within the context of media experiences. They are engaged with in broader literatures like technology, digital economy, law and ethics but often these are found to be separate, siloed and in need for richer context sensitive study. Thus, there is a call to explore this gap within the scope of media as it turns towards personal data leverage.

To undertake this research endeavour, three studies were planned, designed and conducted. Overall, the research approached the problem inspired by the viewpoints strategy in software engineering to include the perspectives of two significant stakeholders in this economy : users and service providers. This approach was adopted to overcome the limitations of previous work around the challenges of using personal data where the users are often treated as a symbolic presence where they are talked about rather than talked to, with respect to their response to such challenges. Thus, the studies here, by involving two major stakeholders in this scenario, was planned to elicit the challenges from the perspectives of both the population whose data is used and those who use the data. This was expected not only to ensure that the results reflect a more holistic view of the challenges confronting the turn towards personal data in media, but also contextualise the findings in the everyday user and service provider practicalities of this shift.

The first study here was a user study designed to understand the user viewpoints regarding the use of personal data in future media experiences. It helped understand that the users were positive about the service enhancement produced in media services by personal data use but there were significant barriers like lack of trust that led to users being unsure of this shift. The study helps unpack these emotions further by explicating the various factors like privacy, legibility, control and accountability that build up to form the overarching challenge of lack of trust.

The second study explores the landscape from the service providers' viewpoint wherein interviews with employees from a varied range of teams within the BBC helped surface the challenges as seen by media service providers during the turn to personal data. Here again the importance of delivering value and the key role of personal data in doing the same was surfaced. But along with it, balancing the risk of loss of user trust was highlighted as necessary. This balance was said to be maintained through measures that ensure legal accountability, social accountability and ethical responsibility from the service

providers' side while the potential of technology as a medium for these responses was also voiced.

The third study bridged the learnings from the first two to create a design fiction, an alternate future world that leveraged user personal data from mundane, everyday objects to customise an adaptive media experience while providing active and explicit responses that are intended to alleviate concerns associated with personal data use. This intervention was used to elicit both user response and service provider response to the experience as a whole, to the use of personal data in it, the inclusion of the data dialogue responses and the future of such experiences and responses. This helped evaluate the feedback to the various responses included and probe the space further to understand user reasonings around the challenges in a deeper, richer manner.

The results from the studies, followed by an engagement of the same with previous literature resulted in a set of contextually situated challenges facing the turn towards personal data in new media. These challenges are presented as both conceptual contributions and implications for design within this thesis.

The conceptual contributions explicate these challenges to bring out the subtleties associated with them as they manifest within the scope of media experiences. These contributions present the viewpoints of both the users and the service providers, as they respond to the turn towards personal data in media experiences and reason out their particular challenges associated with this shift. These viewpoints are made to speak to the broader literatures discussed previously to show how they contribute to those previous conceptualisations of these challenges to include the complexities involving the practical reasonings of users and service providers within the context of media experiences. These conceptual explications of challenges, as voiced by the users and service providers are as follows :

Value

This research presents the challenge of value when considering media experiences that use personal data by highlighting and explicating the complexities associated with it. These complexities manifest in the form of multiple actors, forms of value created, value distribution between actors, the challenges to value created and the balancing of these challenges to maximise the value produced, all of which are situated in the context of media experiences leveraging user personal data.

Trust

The research extends the previously established 'crisis in trust'(Section 2.3.1.) to describe its properties as it manifests in media experiences that leverage user personal data. Here, to rebuild the

current lack of user trust, consideration of trust as more than just a 'single index', to be able to respond to it in multifarious manners that tackle the different aspects of this challenge is highlighted. This challenge is further unpacked as legal accountability, social accountability and ethical considerations, the response to which would work towards alleviating the challenge of loss of user trust.

Privacy

Privacy as a challenge in personal data driven technologies is a generic topic of much debate and interest in research. This thesis explicates the practical nuances characterised by privacy as a challenge as personal data is leveraged by media experiences. This account reports a range of concerns around privacy starting from data security challenges to questions around the data flows and who has access to the personal data collected to the social implications arising from media experiences using personal data and projected fear of potential dystopias.

Transparency

Transparency as a challenge again is moved from generic abstractions to resolve the details of its properties as present in media experiences that leverage user personal data. Here, the different aspects of transparency within and outside of a data driven media experience is called out. Also introduced is the importance of communicating the 'effect' of the data on the media experience, the implications of transparency within a data driven media experience and the call to shift the onus for engaging with transparency measures from the users to present them with user-friendly alternatives.

Control

The challenge of control within a media experience that leverages user personal data is seen to have a multi-dimensional nature. The research details these various aspects of control while also highlighting how this challenge is exacerbated within social scenarios which is often closely associated with media consumption. It also presents the subtleties to be considered when realising control measures where the expectation of quick and easy control interactions are highlighted.

Accountability

The research presents how accountability forms a challenge when media experiences start leveraging personal data. It describes how the various challenges identified here speak to the different perspectives on accountability from previous literature in Law and Computer Science. It also shows how these siloed perspectives on accountability, when satisfied, form overlaps, thus working towards a holistic response towards the overall challenge of accountability.

The implications for design build upon the conceptualisations presented previously to recommend a set of practically resonant design sensitivities to be considered when working with media experiences that leverage user personal data. These recommendations again combine the viewpoints of the users and service providers to present design considerations that are sensitive to the challenges raised by both parties, to work towards alleviation of these challenges.

Trust

Responding to the challenge of loss of user trust by actively involving solid mechanisms that speak to this challenge is called forth. This includes being sensitive to the manifestations of trust within a data driven media experience which includes trust associated with the media content delivered and that associated with the data processing. Here, design that looks into responses that make these experiences legally accountable, socially accountable and ethically considerate, is suggested as a move towards overcoming the challenge of loss of user trust.

Privacy

When considering the challenge of privacy within media experiences using personal data, sensitivity towards the home as a context of service consumption, the social nature of this process which introduces different cohorts of users and the resulting relational nature of the data involved are highlighted as steps towards alleviation of this challenge. Also called forth is the need to debunk the projected fears of dystopias that surround loss of user privacy in these scenarios.

Transparency

The research proposes certain concrete recommendations that would contribute to overcoming the challenges of lack of data transparency. It encourages the inclusion of a rich ‘data dialogue’ with the users which is multi-modal, served at the appropriate levels of depth, granularity and timings and interleaved with the media experience itself in order to make it more engaging and less demanding on the users. Within media experiences, user expectation of transparency extending from what data is collected to include the impact the users (through their data) has on the experience is also put forth.

Control

When considering control mechanisms within media experiences using personal data, sensitivity towards the different aspects of control, within the experience, on the data processes and outside of the experience is highlighted. Also, to be considered is the impact of social scenarios on these control responses and the call for quick and

easy control from the users' side when considering media experiences.

Accountability

Sensitivity to the different notions of accountability when considering media experiences leveraging personal data is highlighted. Careful understanding and response to these different perspectives on accountability through being sensitive to other previously discussed challenges of transparency, control and trust is marked as important if the holistic challenge of accountability is to be responded to.

Within each challenge, these recommendations are compared with current response measures and recommendations within these scopes, using the recommendations of this thesis to add to these previous literatures while scoping these detailed findings to be used specifically when considering media experiences using personal data.

The conceptual explication of challenges here highlight the importance of enabling the users with a more central role in this scenario to work towards alleviating the challenges of both the users and the service providers identified here. The design contributions lay out sensitivities that help realise this shift by proposing measures that both support current practices while also recommending a shift from cloud based models to alternate methods of data processing that allow for increased user participation.

9. Key Contributions

The key contributions of this thesis is not intended solely for the media, technology, HCI or design communities. But it's multi disciplinary approach lends value to multiple domains in one way or another.

These key contributions could be categorised into two, conceptual contributions and implications for design. The conceptual contributions present the following challenges, as reasoned by both users and service providers, of using personal data in media experiences. They compare the findings of the studies reported here to speak to previous literature from multiple disciplines to show how they contribute to those previous conceptualisations of these challenges. The conceptual contributions thus explicate these challenges to include the complexities they embody when involving the practical reasonings of users and service providers within the context of media experiences leveraging user personal data. These conceptual contributions are :

- This research presents the challenge of value when considering media experiences that use personal data by highlighting and explicating the complexities associated with it.

- The research extends the previously established 'crisis in trust' (Section 2.3.1.) to describe its properties as it manifests in media experiences that leverage user personal data.

- This thesis explicates the practical nuances characterised by privacy as a challenge as personal data is leveraged by media experiences.

- This works contributes to dialogue around transparency as a challenge by moving from generic abstractions to resolve the details of its properties as present in media experiences that leverage user personal data.

- The challenge of control within a media experience that leverages user personal data is explicated to present the properties it embodies when considered within the scope of media experiences leveraging user personal data.

- The research presents the different aspects of the challenge of accountability manifest within media experiences that leverage user personal data by reflecting them back to the various siloed perspectives on accountability, as presented by literature from different disciplines.

The second set of contributions build upon these conceptual contributions to present the sensitivities to be taken into account when considering the design of media experiences that leverage personal data. These recommendations combine the viewpoints of the users and service providers to present design considerations that are sensitive to the challenges raised by both parties, to work towards responding to these challenges. These sensitivities are as follows :

- Responding to the challenge of loss of user trust by actively involving considerations that speak to this challenge is presented herewith.

- Sensitivities to be prioritised when considering the various concerns arising from the challenge of privacy is explained.

- Solid recommendations that would help alleviate the challenge of transparency by unpacking the various subtleties and opportunities it presents within data driven media experiences are brought forth.

- The nuances around control which help explain the different dimensions of this challenge to be responded to when designing to alleviate it is explained.

- Accountability and its manifestations in different forms, in line with different multi-disciplinary perspectives on it is highlighted to suggest the careful consideration and response to these different perspectives when considering response to this challenge.

The discussion continuously grounds these findings within literature that proposes current popular responses to these challenges, which are predominantly based on cloud-based data processing models with limited space for user participation in the data management processes. The conceptual contributions here highlight the importance of enabling the users with a more central role in this scenario to work towards alleviating the challenges of both the users and the service providers identified here. The implications for design lay out sensitivities that help realise this shift.

10. Future Work

Having concluded the key findings of this work, it is also important to acknowledge the potential it presents for further research. These possibilities are as follows :

Firstly, the research focuses predominantly on design as a means and response to the challenges faced in future data driven media experiences. Thus, the studies investigate design led solutions from a 'human' perspective of designers and users. There is the potential to extend this work to consider the technological challenges and requirements in more detail and also to include the viewpoints of other stakeholders like regulators, government bodies, data brokers etc.. Thus, expanding upon the viewpoints included and using the findings of this research to translate into technological frameworks that could easily and effectively be deployed to respond to the challenges outlined here is a possibility.

Secondly, the probe used in the final study was designed to reflect the principles of Human Data Interaction (Mortier *et al.*, 2014) in its response to the challenges outlined in the previous studies. There is the possibility of extending this work to other design frameworks like Privacy by Design or to be more inclusive to audiences by introducing accessibility into the scheme.

Thirdly, the studies here primarily use probes as the means to dialogue and discussion. Technology probes could be considered throwaways (Hutchinson *et al.*, 2003). Aligning with that line of thought, as media innovates to provide more data driven affordances, there is the possibility to design alternate versions of the experiments in the future, that are capable of including more data dependencies and innovation that stems from the same. This also presents possibilities of design of the 'data dialogue' that extends beyond the ones included here in the LRoTF. This work provides a starting point for what could be expected to be a plethora of further research in both media and data design that could help respond to the challenges presented here.

Lastly, the key findings of this study explicitly highlight the challenges in the emergent field of media experiences that leverage user personal data. While the final study does look into responses to some of these challenges, for them to become solutions there is the need for these responses to be implemented, investigated and evaluated further in real-world data driven media consumption contexts. Hence, just as mentioned in the previous article, the responses of 'data dialogue' designed for the LRoTF (including the use of the DataBox, the data monitor etc.) are only examples that help investigate this emergent scenario at its current state. Rather than advocate for them as 'the' solutions for the future, this work uses these responses as probes for exploration of this space and

recommends them as possible solutions, but more importantly, as vehicles for further exploration and study towards future solutions.

Bibliography

(Ec), E. C. (2011) 'SPECIAL EUROBAROMETER 359 Attitudes on Data Protection and Electronic Identity in the European Union', p. 330. Available at: http://ec.europa.eu/public_opinion/index_en.htm.

Ackerman, M. S. and Mainwaring, S. D. (2005) 'Privacy issues and human-computer interaction', *Computer*, pp. 381–400. Available at: http://cs.calstatela.edu/~wiki/images/b/bf/Privacy.pdf (Accessed: 11 January 2017).

Al-Fedaghi, S. (2007) 'How sensitive is your personal information?', *Proceedings of the 2007 ACM symposium on Applied computing*, pp. 165–169. doi: 10.1145/1244002.1244046.

Almuhimedi, H. (2015) 'Your location has been shared 5398 times - A field study on mobile app privacy nudging', *In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 787–796.

Amar, Y., Haddadi, H. and Mortier, R. (2016) 'Privacy-Aware Infrastructure for Managing Personal Data Personal Data Arbitering within the Databox Framework', *Proceedings of the 2016 conference on ACM SIGCOMM 2016 Conference*, pp. 571–572.

Amatriain, X. (2013) 'Big & personal: data and models behind netflix recommendations', in *Proceedings of the 2nd International Workshop on Big Data, Streams and Heterogeneous Source Mining Algorithms, Systems, Programming Models and Applications - BigMine '13*, pp. 1–6. doi: 10.1145/2501221.2501222.

Apple News - Apple (2019) *Apple*. Available at: https://www.apple.com/apple-news/ (Accessed: 15 August 2019).

Arthur, C. (2011) *Sony suffers second data breach with theft of 25m more user details*, *The Guardian*. Available at: http://www.guardian.co.uk/technology/blog/2011/may/03/sony-data-breach-online- entertainment. (Accessed: 6 July 2019).

Bacon, J. (2016) *Young viewers choosing VOD over pay TV*, *Marketing Week*. Available at: https://www.marketingweek.com/2016/02/11/young-viewers-choosing-vod-over-pay-tv/ (Accessed: 18 September 2017).

Baker, L. B. and Finkle, J. (2011) *Sony PlayStation suffers massive data breach*, *Reuters*. Available at: http://www.reuters.com/article/2011/04/26/us-sony-stoldendata-idUSTRE73P6WB20110426 (Accessed: 6 July 2019).

Bauer, L., Bravo-Lillo, C. and Cranor, L. (2013) 'Warning design guidelines', *Pittsburgh, PA:* Available at: http://ai2-s2-pdfs.s3.amazonaws.com/3f14/e9176d3382a2acbd959db8be7630f0ad37e0.pdf (Accessed: 11 January 2017).

Bauer, L., Bravo-Lillo, C., Cranor, L. F. and Fragkaki, E. (2013) 'Warning design guidelines', *Pittsburgh, PA: Carnegie Mellon University*.

BBC (2014) *Electronic Programme Guides (EPG)*, *BBC*. Available at: http://www.bbc.co.uk/webwise/helpplus/electronicprogrammeguides.shtml (Accessed: 14 July 2019).

BBC (2019) *Why the BBC does not want to store your data - BBC News*, *BBC*. Available at: https://www.bbc.co.uk/news/technology-48825413 (Accessed: 15 August 2019).

BBC R&D (2018) *Building A Public Service Internet - BBC R&D*, *BBC*. Available at: https://www.bbc.co.uk/rd/projects/public-service-internet (Accessed: 15 August 2019).

Bellotti, V. and Edwards, K. (2001) 'Intelligibility and Accountability: Human Considerations in Context-Aware Systems', *Human–Computer Interaction*, 16(2–4), pp. 193–212.

Benson, D. and Hughes, J. A. (1983) *The perspective of ethnomethodology*. Longman Publishing Group.

Benson, V., Saridakis, G. and Tennakoon, H. (2015) 'Information disclosure of social media users', *Information Technology & People*, 28(3), pp. 426–441. doi: 10.1108/ITP-10-2014-0232.

Berkvosky, S., Eytani, Y., Kuflik, T. and Ricci, F. (2007) 'Enhancing privacy and preserving accuracy of a distributed collaborative filtering', *RecSys '07: Proceedings of the 2007 ACM Conference on Recommender Systems*, pp. 97–104. doi: 10.1145/1297231.1297234.

Bernal, P. (2014) *Internet privacy rights: rights to protect autonomy*. Cambridge University Press. Available at: https://books.google.co.uk/books?hl=en&lr=&id=0TI4AwAAQBAJ&oi=fnd&pg=PR9&dq=internet+privacy+rights+bernal&ots=cDs2UoKkkY&sig=XwJ429XuXnTU0nXP-BkxG6Zlfr0 (Accessed: 7 July 2019).

Bernhaupt, R., Obrist, M., Weiss, A., Beck, E. and Tscheligi, M. (2007) 'Trends in the Living Room and Beyond', in *European Conference on Interactive Television*. Springer Berlin Heidelberg, pp. 146–155. doi: 10.1007/978-3-540-72559-6.

Bernhaupt, R., Wilfinger, D. and Mirlacher, T. (2010) 'Personalizing the TV Experience: Vocomedia–a Case Study from Interactive TV', in *Mass Customization for Personalized Communication Environments: Integrating Human Factors*. IGI Global, pp. 149–160. Available at: https://www.igi-global.com/chapter/personalizing-experience-vocomedia-case-study/38512 (Accessed: 15 August 2019).

Bing, J. (1972) 'Classification of personal information, with respect to the sensitivity aspect', in *Proceedings of the First International Oslo Symposium on Data Banks and Societies*. Oslo: Scandinavian Univ. Books.

Bisson, S. (2017) *Building my own Internet of Things ambient experience, one step at a time*, *zdnet.com*.

Black, J. (2001) 'Decentring regulation: Understanding the role of regulation and self-regulation in a "post-regulatory" world', *Current legal problems*, 54(1), pp. 103–146.

Bleecker, J. (2009) 'Design Fiction: A Short Essay on Design, Science, Fact and Fiction', *Near Future Laboratory*, (March), pp. 1–97. Available at: http://drbfw5wfjlxon.cloudfront.net/writing/DesignFiction_WebEdition.pdf.

Blippar (2019). Available at: https://www.blippar.com (Accessed: 5 July 2019).

Bonnici, S. (2003) 'Which channel is that on? A design model for electronic programme guides', *Proceedings of the 1st European conference on interactive television: from viewers to actors*, pp. 49–57.

Boyatzis, R. E. (1998) *Transforming qualitative information: Thematic analysis and code development*. Sage.

Bozios, T., Lekakos, G., Skoularidou, V. and Chorianopoulos, K. (2001) 'Advanced techniques for personalized advertising in a digital TV environment: the iMEDIA system', in *Proceedings of the eBusiness and eWork Conference*, pp. 1025–1031. Available at: http://users.ionio.gr/choko/pdf/Bozios_2001.pdf (Accessed: 5 July 2019).

Braun, V. and Clarke, V. (2006a) 'Using thematic analysis in psychology', *Qualitative research in psychology*. Available at: http://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa (Accessed: 18 September 2017).

Braun, V. and Clarke, V. (2006b) 'Using thematic analysis in psychology', *Qualitative research in psychology*, 3(2), pp. 77–101.

Brehm, J. W. (1966) 'A theory of psychological reactance'.

Brehm, J. W. (1989) ‘Psychological reactance: Theory and applications’, *ACR North American Advances*. Available at: http://www.acrwebsite.org/volumes/6883/volumes/v16/NA-16 (Accessed: 5 July 2019).

Brown, J. B. (1999) ‘The use of focus groups in clinical research’, in Crabtree, B. F. and Miller, W. L. (eds) *Doing qualitative research*. Sage Publications. Available at: https://ci.nii.ac.jp/naid/10018364132/ (Accessed: 1 July 2019).

Brownsword, R. (2008) *Rights, regulation and the technological revolution*. Oxford University Press. Available at: https://dl.acm.org/citation.cfm?id=1386907 (Accessed: 7 July 2019).

Bursztein, E., Moscicki, A., Fabry, C., Bethard, S., Mitchell, J. C. and Jurafsky, D. (2014) ‘Easy Does It: More Usable CAPTCHAs’, *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI ’14*, pp. 2637–2646. doi: 10.1145/2556288.2557322.

Cadwalladr, C. and Graham-Harrison, E. (2018) *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*, *The Guardian*. Available at: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election (Accessed: 28 August 2018).

Calo, R. (2011) ‘The boundaries of privacy harm’, *Ind. LJ*, 86, p. 1131.

Camp, J. and Osorio, C. (2003) ‘Privacy-Enhancing Technologies for Internet Commerce.’, *In Trust in the Network Economy SpringerVerlag Berlin*.

Carroll, J. (1995) ‘Scenario-based design: envisioning work and technology in system development’. Available at: http://dl.acm.org/citation.cfm?id=209227 (Accessed: 10 July 2016).

Carroll, J. M. (1999) ‘Five reasons for scenario-based design’, in *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences*. IEEE. Available at: https://ieeexplore.ieee.org/abstract/document/772890/ (Accessed: 3 July 2019).

Cate, F. (2010) ‘The limits of notice and choice’, *IEEE Security & Privacy*. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5439530 (Accessed: 10 September 2016).

Cavoukian, A. (2009) ‘Privacy by Design - The 7 foundational principles - Implementation and mapping of fair information practices’, *Information and Privacy Commissioner of Ontario, Canada*.

Cesar, P. and Chorianopoulos, K. (2009) *The Evolution of TV Systems,*

Content, and Users Toward Interactivity, *Foundations and Trends in Human-Computer Interaction*. doi: 10.1561/1100000008.

Cesar, P., Chorianopoulos, K. and Jensen, J. F. (2008) 'Social television and user interaction', *Computers in Entertainment (CIE)*, 6(1), p. 4. Available at: https://dl.acm.org/citation.cfm?id=1350847 (Accessed: 5 July 2019).

Chellappa, R. K. and Sin, R. G. (2005) 'Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma', *Information Technology and Management*, 6(2–3), pp. 181–202. doi: 10.1007/s10799-005-5879-y.

Cheng, J. (2014) *Analyzing Minard's Visualization Of Napoleon's 1812 March*, *thoughtbot.com*. Available at: https://thoughtbot.com/blog/analyzing-minards-visualization-of-napoleons-1812-march (Accessed: 23 July 2019).

Cherry, C. (1977) 'The telephone system: creator of mobility and social change', *The social impact of the telephone*, pp. 112–126.

Chollangi, N. (2017) *RYE Grads make Interactive Documentary about Immigrant Entrepreneurs*. Available at: https://theeyeopener.com/2017/01/rye-grads-make-interactive-documentary-about-immigrant-entrepreneurs/ (Accessed: 5 July 2019).

Churnside, T. (2013) *Object-Based Broadcasting*. Available at: http://www.bbc.co.uk/rd/blog/2013-05-object-based-approach-to-broadcasting (Accessed: 10 September 2016).

Cissée, R. and Albayrak, S. (2007) 'An agent-based approach for privacy-preserving recommender systems', *Proceedings of the International Conference on Autonomous Agents*, 5, pp. 319–326. doi: 10.1145/1329125.1329345.

Clara Rodriguez Fernandez (2017) *This Bioartist is Exhibiting Living Portraits of her own Daughters*, *labiotech.eu*. Available at: https://labiotech.eu/bioart/gina-czarnecki-heirloom-bioart/ (Accessed: 5 July 2019).

Clarke, R. (2006) *Roger Clarke's 'What's Privacy?'* Available at: http://www.rogerclarke.com/DV/Privacy.html (Accessed: 7 July 2019).

Clee, M. A. and Wicklund, R. A. (1980) 'Consumer behavior and psychological reactance', *Journal of consumer research*, 6(4), pp. 389–405. Available at: https://academic.oup.com/jcr/article-abstract/6/4/389/1790960 (Accessed: 5 July 2019).

Constine, J. (2016) *Augmented Reality For Trying On Makeup Is A Booming*

Business, *TechCrunch.com*.

Coulton, P. (2015) 'Playful and Gameful Design for the Internet of Things', in, pp. 151–173. doi: 10.1007/978-981-287-546-4_7.

Coulton, P. (2017) 'Sensing Atoms and bits', *Sensory Arts and Design*, 189. Available at: https://books.google.co.uk/books?hl=en&lr=&id=KZWWDQAAQBAJ&oi=fnd&pg=PA189&dq=Sensing+atoms+and+bits.+in+I+Heywood+PAUL+coulton&ots=kdY8sXXuzL&sig=2AktMuGrz2lBXlkhG6Vu0WoHpM8 (Accessed: 27 August 2018).

Coulton, P., Lindley, J. and Cooper, R. (2018) *The Little Book of Design Fiction for the Internet of Things*. Available at: https://www.petrashub.org/the-little-book-of-design-fiction-for-the-internet-of-things/.

Crabtree, A., Chamberlain, A., Grinter, R. E., Jones, M., Rodden, T. and Rogers, Y. (2013) 'Introduction to the special issue of "The Turn to The Wild"', *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(3), p. 13.

Crabtree, A., Hemmings, T., Rodden, T., Cheverst, K., Clarke, K., Dewsbury, G., Hughes, J. and Rouncefield, M. (2003) 'Designing with care: Adapting cultural probes to inform design in sensitive settings', *Proceedings of the 2004 Australasian Conference on Computer-Human Interaction (OZCHI2004)*, pp. 4–13. Available at: http://www.academia.edu/download/46591699/Designing_with_Care_Adapting_Cultural_Pr20160618-6462-1p69z8b.pdf (Accessed: 3 July 2019).

Crabtree, A., Lodge, T., Colley, J., Greenhalgh, C. and Mortier, R. (2016) 'Building Accountability into the Internet of Things'. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2881876 (Accessed: 27 August 2018).

Crabtree, A. and Mortier, R. (2015) 'Human Data Interaction : Historical Lessons from Social Studies and CSCW', in *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway*, pp. 3–21.

Crabtree, A., Tolmie, P. and Rouncefield, M. (2012) *Doing Design Ethnography*, *Springer Verlag*. London.

Creegan, C. L. (1989) *Wittgenstein and Kierkegaard: Religion, individuality, and philosophical method*. London: Routledge.

Crisp, S. (2017) *Hands-on: Mardles storybooks pop up with augmented reality*, *New Atlas*.

Culnan, M. J. and Armstrong, P. K. (1999) 'Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation', *Organization Science*, 10(1), pp. 104–115. doi: 10.1287/orsc.10.1.104.

Culnan, M. J. and Bies, R. J. (2003) 'Consumer Privacy: Balancing Economic and Justice Considerations', *Journal of Social Issues*, 59(2), pp. 323–342. doi: 10.1111/1540-4560.00067.

Czarnecki, G. (2018) *Heirloom*. Available at: https://www.ginaczarnecki.com/heirloom (Accessed: 5 July 2019).

Damiani, J. (2017) *Meta Debuts Groundbreaking Interactive AR at Sundance*, *VR Scout*. Available at: https://vrscout.com/news/meta-interactive-ar-journey-natural-machine-sundance/ (Accessed: 5 July 2019).

David Bowie's "Heroes" Inspires Quest into New Technologies at Sundance 2017 (2017) *BMI*.

Davies, N., Langheinrich, M., Clinch, S., Elhart, I., Friday, A., Kubitza, T. and Surajbali, B. (2014) 'Personalisation and Privacy in Future Pervasive Display Networks', in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pp. 2357–2366. doi: 10.1145/2556288.2557287.

DCMS (2013) *Amendments to the Communications Act 2003 (Electronic Programme Guides)*, *Sustainable Development*.

Dean, D., Digrande, S., Field, D., Lundmark, A., O'day, J., Pineda, J. and Zwillenberg, P. (2012) 'The 4.2 Trillion Opportunity: The Internet Economy in the G-20', *Boston Consulting Group*, (March), p. 57.

Dennedy, M., Fox, J. and Finneran, T. (2014) 'Privacy Engineer's Manifesto', *Apress, New York*.

Denzin, N. K. and Lincoln, Y. S. (eds) (2011) *The Sage handbook of qualitative research*. Sage.

Van Doorn, J. and Hoekstra, J. C. (2013) 'Customization of online advertising: The role of intrusiveness', *Marketing Letters*, 24(4), pp. 339–351. Available at: https://link.springer.com/article/10.1007/s11002-012-9222-1 (Accessed: 5 July 2019).

Dourish, P. and Anderson, K. (2006) 'Collective Information Practice: Exploring Privacy and Security as Social and Cultural Phenomena', *Human-Computer Interaction*, 21(3), pp. 319–342. doi: 10.1207/s15327051hci2103_2.

Dourish, P. and Button, G. (1998) 'On "Technomethodology" : Foundational

Relationships Between Ethnomethodology and System Design', *Human-Computer Interaction*, 13(4), pp. 395–432. doi: 10.1207/s15327051hci1304_2.

Dugas, A. F., Hsieh, Y.-H., Levin, S. R., Pines, J. M., Mareiniss, D. P., Mohareb, A., Gaydos, C. A., Perl, T. M. and Rothman, R. E. (2012) 'Google Flu Trends: Correlation With Emergency Department Influenza Rates and Crowding Metrics', *Clinical Infectious Diseases*, 54(4), pp. 463–469. doi: 10.1093/cid/cir883.

Dwork, C. (2008) 'Differential privacy: A survey of results', in *International conference on theory and applications of models of computation*. Springer, Berlin, Heidelberg.

Van Dyke, T., Midha, V. and Nemati, H. (2007) 'The Effect of Consumer Privacy Empowerment on Trust and Privacy Concerns in E-Commerce', *Electronic Markets*, 17(1), pp. 68–81. doi: 10.1080/10196780601136997.

Eclipse wiki (2009) *Higgins Personal Data Service*. Available at: http://www.eclipse.org/higgins/ (Accessed: 7 February 2017).

EDPS (2015) 'Opinion 3 / 2015 Europe ' s big opportunity', 2015(July), pp. 1–13.

EDPS (2018) 'Public Consultation on Digital Ethics'.

Egelman, S., Harbach, M. and Peer, E. (2016) 'Behavior Ever Follows Intention?: A Validation of the Security Behavior Intentions Scale (SeBIS)', in *In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 5257–5261.

Egelman, S., Kannavara, R. and Chow, R. (2015) 'Is this thing on?: Crowdsourcing privacy indicators for ubiquitous sensing platforms', in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 1669–1678.

Etherington, D. (2017) *Niantic's follow-up to Pokemon Go will be a Harry Potter AR game launching in 2018*, *TechCrunch.com*.

EU (1995) *EU Data Protection Directive 95/46/EC*, *European Parliament and of the Council*. Available at: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A31995L0046 (Accessed: 7 July 2019).

EU GDPR (2016) *Home Page of EU GDPR*. Available at: http://www.eugdpr.org/ (Accessed: 7 July 2019).

European Commission (2012) 'Proposal for a regulation of the European

Parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)', *COM (2012) 11 final, 2012/0011 (COD), Brussels*, 0011, p. 118. doi: 2012/0011 (COD).

European Data Protector Supervisor (2015) 'Opinion 4 / 2015 Towards a new digital ethics', (September).

European Union (2018) 'Charter of Fundamental Rights (EU)', *Encyclopedia of Big Data*. doi: 10.1007/978-3-319-32001-4_35-1.

Evans, M., Ferne, T., Watson, Z., Melchior, F., Brooks, M., Stenton, P. and Forrester, I. (2016) 'Creating object-based experiences in the real world', in *IBC 2016 Conference*. Institution of Engineering and Technology, pp. 34–8. doi: 10.1049/ibc.2016.0034.

Faridani, S., Bitton, E., Ryokai, K. and Goldberg, K. (2010) 'Opinion space: a scalable tool for browsing online comments', *Proceedings of the 28th international conference on Human factors in computing systems*, pp. 1175–1184. doi: 10.1145/1753326.1753502.

Franke, N., Keinz, P. and Steger, C. J. (2009) 'Testing the Value of Customization: When Do Customers Really Prefer Products Tailored to Their Preferences?', *Journal of Marketing*, 73(5), pp. 103–121. doi: 10.1509/jmkg.73.5.103.

FTC (1998) *Privacy Online : A Report To Congress*.

Garfinkel, H. (1996) *Studies in Ethnomethodology*. London, United Kingdom: Polity Press.

Gaver, W., Dunne, T. and Pacenti, E. (1999) 'Cultural Probes- Probing People for Design Inspiration', *interactions*, 6(1), pp. 21–29.

Gavison, R. (1980) 'Privacy and the Limits of Law', *The Yale Law Journal*, 89(3), pp. 421–471. Available at: https://www.jstor.org/stable/795891 (Accessed: 6 July 2019).

GDPR (2018) *GDPR Article 4*. Available at: https://gdpr-info.eu/art-4-gdpr/.

GDPR, A. 7 (2016) 'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da', *Official Journal of the European Union*, pp. 1–88.

Gershgorn, D. (2015) *Google Translate adds 20 languages to Augmented Reality App*, *popsci.com*. Available at: https://www.popsci.com/google-translate-adds-augmented-reality-translation-app/ (Accessed: 5 July 2019).

Given, L. M. (ed.) (2008) 'Ethnomethodology', in *The SAGE Encyclopedia of Qualitative Research Methods*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. doi: 10.4135/9781412963909.n152.

Goldfarb, A. and Tucker, C. (2011) 'Online Display Advertising: Targeting and Obtrusiveness', *Marketing Science*, 30(3), pp. 389–404. doi: 10.1287/mksc.1100.0583.

Goldreich, O. (2006) 'Foundations of cryptography - A primer', *Foundations and Trends in Theoretical Computer Science*, 1(1), pp. 1–116. doi: 10.1561/0400000001.

Good, N. S., Grossklags, J., Mulligan, D. K. and Konstan, J. a. (2007) 'Noticing notice: a large-scale experiment on the timing of software license agreements', *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, p. 607. doi: 10.1145/1240624.1240720.

Goodner, S. (2016) *Augmented reality atom kit makes hands-on chemistry easier to grasp*, *New Atlas*. Available at: https://newatlas.com/happy-atoms-augmented-reality-chemistry/44074/ (Accessed: 5 July 2019).

Google (2019) *Google Sky*. Available at: https://www.google.com/sky/ (Accessed: 5 July 2019).

Gradinar, A., Burnett, D., Coulton, P., Forrester, I., Watkins, M., Scutt, T. and Murphy, E. (2015) 'Perceptive Media – Adaptive Storytelling for Digital Broadcast', *Human-Computer Interaction*, pp. 586–589. doi: 10.1007/978-3-319-22723-8_67.

Grant, C. (2016) *Pokémon Go has been downloaded 500 million times*, *Polygon.com*. Available at: https://www.polygon.com/pokemon-go/2016/9/7/12836898/pokemon-go-downloads-500-million (Accessed: 5 July 2019).

Green, A. (2018) *What is User Behavior Analytics?*, *varonis.com*. Available at: https://www.varonis.com/blog/what-is-user-behavior-analytics/ (Accessed: 24 July 2019).

Hachman, M. (2016) *Intel's Authenticate tech brings simple-but-powerful security to Skylake chips*, *pcworld.com*. Available at: https://www.pcworld.com/article/3024314/intels-authenticate-tech-brings-simple-but-powerful-security-to-skylake-chips.html (Accessed: 7 July 2019).

Halcomb, E. and Hickman, L. (2014) 'Mixed Methods Research', *Nursing Standard*, 29(32), pp. 41–47. doi: 10.4018/IJAVET.2016040103.

Hanley, P. and Viney, R. (2001) 'Pressing the red button: Consumers and digital television', *Cultural Trends*, 11(43–44), pp. 35–60. doi: 10.1080/09548960109365165.

Hann, I.-H., Hui, K.-L., Lee, T. and Png, I. (2002) 'Online information privacy: Measuring the cost-benefit trade-off', *ICIS 2002 proceedings*, p. 1. Available at: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1093&context=icis2002 (Accessed: 23 July 2019).

Harris, J. and Kamvar, S. (2006) *We Feel Fine*. Available at: http://wefeelfine.org (Accessed: 5 July 2019).

Harris, J. and Kamvar, S. (2008) *I Want You to Want Me*. Available at: http://number27.org/iwytwm (Accessed: 5 July 2019).

Hautea, S., Dasgupta, S. and Hill, B. M. (2017) 'Youth Perspectives on Critical Data Literacies', in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. New York, New York, USA: ACM Press, pp. 919–930. doi: 10.1145/3025453.3025823.

Hazzard, A., Benford, S. and Burnett, G. (2015) 'Sculpting a mobile musical soundtrack', in *Proceedings of the 33rd annual ACM conference on Human factors in computing systems*. ACM, pp. 387–396. Available at: https://dl.acm.org/citation.cfm?id=2702236 (Accessed: 23 July 2019).

Hemmings, T., Clarke, K., Rouncefield, M., Crabtree, A. and Rodden, T. (2002) 'Probing the Probes', *PDC*, (June), pp. 42–50.

Henry, S. (2013) *A New Story : Purposeful Storytelling and Designing with Data*. Available at: https://www.academia.edu/4248593/A_New_Story_Purposeful_storytelling_and_designing_with_data.

Holladay, R. (2013) *To Hear this Music You Have to be There. Literally*, *TED*. Available at: https://www.ted.com/talks/ryan_holladay_to_hear_this_music_you_have_to_be_there_literally (Accessed: 5 July 2019).

Hong, E. (2016) *Why some Koreans make $10,000 a month to eat on camera*, *Quartz*. Available at: https://qz.com/592710/why-some-koreans-make-10000-a-month-to-eat-on-camera/ (Accessed: 5 July 2019).

Hook, J. (2018) 'Facts , Interactivity and Videotape : Exploring the Design

Space of Data in Interactive Video Storytelling', *Proceedings of the ACM TVX International Conference on Interactive Experiences for Television and Online Video. ACM TVX '18 International Conference on Interactive Experiences for Television and Online Video*, pp. 43–55. doi: 10.1145/3210825.3210826.

House, W. (2012) 'Consumer data privacy in a networked world: A framework for protecting privacy and promoting innovation in the global digital economy', *White House, Washington, DC*, pp. 1–62. Available at: https://www.whitehouse.gov/sites/default/files/privacy-final.pdf.

Hui, K.-L., Hai Teo, H. and Tom Lee, S.-Y. (2007) 'The value of privacy assurance: an exploratory field experiment', *Mis Quarterly*, pp. 19–33. Available at: https://www.jstor.org/stable/25148779 (Accessed: 23 July 2019).

Hutchinson, H., Mackay, W. E., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., Eiderbäck, B. and others (2003) 'Technology probes: inspiring design for and with families', in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '03*, pp. 17–24.

Iachello, G. and Hong, J. (2007) 'End User Privacy in Human Computer Interaction', *Foundations and Trends in Human Computer Interaction*, 1(1), pp. 1–137. Available at: https://www.gartner.com/newsroom/id/3114217.

IJsselsteijn, W. (2017) 'Here's Looking At You, Kid', in *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video - TVX '17*. New York, New York, USA: ACM, pp. 1–1. doi: 10.1145/3077548.3077562.

Ioannidis, Y., Balet, O. and Pandermalis, D. (2014) *Tell me a story: augmented reality technology in museums*, *The Guardian*.

Jamal, A., Coughlan, J. and Kamal, M. (2013) 'Mining social network data for personalisation and privacy concerns: a case study of Facebook's Beacon'. Available at: https://bura.brunel.ac.uk/handle/2438/8080 (Accessed: 5 July 2019).

Jeckmans, A., Beye, M., Erkin, Z., Hartel, P., Lagendijk, R. and Tang, Q. (2013) 'Privacy in Recommender Systems', *Social media retrieval*, pp. 263–281. Available at: http://www.csrquest.net/uploadfiles/1D.pdf.

Jones, O. D. (2000) 'Navigating your TV : The Usability of Electronic Programme Guides', *UsableiTV 3 (2002): 2-6*, 3, pp. 2–6.

Jones, R., Sailaja, N. and Kerlin, L. (2017) 'Probing the Design Space of Usable Privacy Policies: A Qualitative Exploration of a Reimagined Privacy Policy', in *Proceedings of the 31st British Computer Society Human Computer*

Interaction Conference. BCS Learning & Development Ltd, p. 50. Available at: http://hci2017.bcs.org/wp-content/uploads/BHCI_2017_paper_62-1.pdf (Accessed: 15 September 2017).

Kalapesi, C. (2012) *Rethinking Personal Data : Strengthening Trust*.

Kelley, P. G., Bresee, J., Cranor, L. F. and Reeder, R. W. (2009) 'A "nutrition label" for privacy', in *Proceedings of the 5th Symposium on Usable Privacy and Security SOUPS 09*, p. 4. doi: 10.1145/1572532.1572538.

Kennedy, H. and Moss, G. (2015) 'Known or knowing publics? Social media data mining and the question of public agency', *Big Data & Society*, 2(2), p. 205395171561114. doi: 10.1177/2053951715611145.

Krauss, C., George, L. and Arbanowski, S. (2013) 'TV predictor: personalized program recommendations to be displayed on SmartTVs', in *Proceedings of the 2nd international workshop on big data, streams and heterogeneous source mining: Algorithms, systems, programming models and applications. ACM, 2013.*, pp. 63–70.

Kriplean, T., Morgan, J., Freelon, D., Borning, A. and Bennett, L. (2012) 'Supporting reflective public thought with considerit', *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, p. 265. doi: 10.1145/2145204.2145249.

Krishnan, M. S. (2016) 'The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization Author(s)':, 30(1), pp. 13–28.

Krueger, R. (2014) *Focus groups: A practical guide for applied research*. Available at: https://books.google.co.uk/books?hl=en&lr=&id=tXpZDwAAQBAJ&oi=fnd&pg=PT7&dq=focus+groups+a+practical+guide+krueger&ots=Psq61Wbt1y&sig=cjFXgQwUvqPZsnnXGTGNZ_VamDQ (Accessed: 1 July 2019).

Kudina, O. and Verbeek, P.-P. (2019) 'Ethics from within: Google Glass, the Collingridge dilemma, and the mediated value of privacy', *Science, Technology, & Human Values*, 44(2), pp. 291–314.

Lam, S. K. "Tony", Frankowski, D. and Riedl, J. (2006) 'Do You Trust Your Recommendations? An Exploration of Security and Privacy Issues in Recommender Systems', in G., M. (ed.) *Emerging Trends in Information and Communication Security. ETRICS 2006. Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, pp. 14–29.

Lanier, J. (2014) *Who owns the future?* Simon and Schuster.

Last Hijack (2015). Available at: http://lasthijack.com (Accessed: 5 July 2019).

Lazar, J., Feng, J. H. and Hochheiser, H. (2017) *Research methods in human-computer interaction*. Morgan Kaufmann.

Lee, H. and Kwon, J. (2013) 'A Smart TV Application for Personalized Recommender Services', 29(Ca), pp. 174–176.

Lennon, M. (2011) *Massive Breach at Epsilon Compromises Customer Lists of Major Brands*, *securityweek.com*. Available at: http://www.securityweek.com/massive-breach-epsilon-compromises-customer-lists-major-brands (Accessed: 6 July 2019).

Levine, R., Locke, C., Searls, D. and Weinberger, D. (2009) *The cluetrain manifesto*. Basic books.

Li, H., Edwards, S. M. and Lee, J.-H. (2002) 'Measuring the Intrusiveness of Advertisements: Scale Development and Validation', *Journal of Advertising*, 31(2), pp. 37–47. doi: 10.1080/00913367.2002.10673665.

Lindley, J. (2018) 'A thesis about design fiction', (July). doi: 10.17635/lancaster/thesis/449.

Lindley, J., Coulton, P. and Sturdee, M. (2017) 'Implications for Adoption', *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, pp. 265–277. doi: 10.1145/3025453.3025742.

Lindley, J., Sharma, D. and Potts, R. (2014) 'Anticipatory Ethnography: Design Fiction as an Input to Design Ethnography', *Ethnographic Praxis in Industry Conference Proceedings*, 2014(1), pp. 237–253. doi: 10.1111/1559-8918.01030.

Luger, E., Urquhart, L., Rodden, T., Golembewski, M., New, U. S. A. and York, N. Y. (2015) 'Playing the Legal Card: Using Ideation Cards to Raise Data Protection Issues within the Design Process', *In Proceedings SIGCHI Conference Human Factors in Computer Systems (CHI'15) ACM Press*, pp. 457–466.

Lull, J. (1980) 'The social uses of television', *Human communication research*, 6(3), pp. 197–209.

Lupton, D. (2016) *The quantified self*. John Wiley & Sons.

Lynch, M. (1997) *Scientific practice and ordinary action: Ethnomethodology and social studies of science*. Available at:

https://books.google.co.uk/books?hl=en&lr=&id=aMu7oIdtnSQC&oi=fnd&pg=PP1&dq=lynch+business+of+talk+scientific+practice+and+ordinary+action&ots=oLTVRkjVTf&sig=oAqcMqW83erFn-qXuJUsxGTBRSs (Accessed: 8 January 2018).

Malhotra, N. K., Kim, S. S. and Agarwal, J. (2004) 'Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model', *Information Systems Research*, 15(4), pp. 336–355. doi: 10.1287/isre.1040.0032.

Manheim, K. (1952) *On the interpretation of Weltanschauung*. From Karl Manheim.

Markus, M. L. (1987) 'Toward a "critical mass" theory of interactive media: Universal access, interdependence and diffusion', *Communication Research*, 14(5), pp. 491–511. doi: 10.1177/009365087014005003.

Markussen, T. and Knutz, E. (2013) 'The poetics of design fiction', in *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces*. ACM. Available at: https://dl.acm.org/citation.cfm?id=2513531 (Accessed: 1 July 2019).

Martinez, A. G. (2019) *No, Data Is Not the New Oil | WIRED*, *Wired.com*. Available at: https://www.wired.com/story/no-data-is-not-the-new-oil/ (Accessed: 15 August 2019).

McDonald, A. and Cranor, L. F. (2008) 'The Cost of Reading Privacy Policies', *A Journal of Law and Policy for the Information Society*, 4(3), pp. 543–568.

McMahan, B. and Ramage, D. (2017) *Federated Learning: Collaborative Machine Learning without Centralized Training Data*, *Google AI Blog*. Available at: https://ai.googleblog.com/2017/04/federated-learning-collaborative.html (Accessed: 8 July 2019).

McReynolds, E., Hubbard, S., Lau, T. and Saraf, A. (2017) 'Toys that Listen: A Study of Parents, Children, and Internet-Connected Toys', in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 5197–5207.

McSherry, F. and Mironov, I. (2009) 'Differentially private recommender systems: Building privacy into the netflix prize contenders', *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 627–635. doi: 10.1145/1557019.1557090.

Miller, M. H. (2011) *Sony data breach could be most expensive ever*, *csmonitor.com*. Available at:

http://www.csmonitor.com/Business/2011/0503/Sony-data-breach-could-be-most- expensive-ever (Accessed: 6 July 2019).

Miller, W. and Crabtree, B. (1992) 'Primary care research: A multimethod typology and qualitative road map'. Available at: http://psycnet.apa.org/psycinfo/1992-97742-001 (Accessed: 5 January 2018).

de Montjoye, Y.-A., Shmueli, E., Wang, S. S. and Pentland, A. S. (2014) 'openPDS: Protecting the Privacy of Metadata through SafeAnswers', *PLoS ONE*. Edited by T. Preis. Public Library of Science, 9(7), p. e98790. doi: 10.1371/journal.pone.0098790.

Mortier, R., Haddadi, H., Henderson, T., Mcauley, D. and Crowcroft, J. (2014) 'Human-Data Interaction : The Human Face of the Data-Driven Society'.

Moser, C., Chen, T. and Schoenebeck, S. (2017) 'Parent's and Children's Preferences about Parents Sharing about Children on Social Media', *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 5221–5225. Available at: http://dl.acm.org/citation.cfm?id=3025587 (Accessed: 14 September 2017).

Munson, S. A. and Resnick, P. (2010) 'Presenting diverse political opinions: how and how much', in ACM (ed.) *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1457–1466. doi: 10.1145/1753326.1753543.

My Sky is Falling (2013). Available at: http://www.myskyisfalling.com (Accessed: 5 July 2019).

Mydex (2017) *Mydex charter*. Available at: https://pds.mydex.org/mydex-charter (Accessed: 7 February 2017).

Narayanan, A. and Shmatikov, V. (2006) 'How To Break Anonymity of the Netflix Prize Dataset', *arXiv preprint cs/0610105*. Available at: http://arxiv.org/abs/cs/0610105.

Nelson, D. (2001) *Pursuing Privacy in Cold War America*. Columbia University Press.

Nilsson, T., Crabtree, A., Fischer, J. and Koleva, B. (2019) 'Breaching the future: Understanding human challenges of autonomous systems for the home', *Personal and Ubiquitous Computing*, 23(2), pp. 287–307. doi: 10.1007/s00779-019-01210-7.

Nissenbaum, H. (2004) 'Privacy as contextual integrity', *Wash. L. Rev.*, pp. 101–139. doi: 10.1109/SP.2006.32.

Norberg, P. A., Horne, D. R. and Horne, D. A. (2007) 'The privacy paradox: Personal information disclosure intentions versus behaviors', *Journal of Consumer Affairs*, 41(1), pp. 100–126.

O'Sullivan, D., Smyth, B. and Wilson, D. (2003) 'Explicit vs implicit profiling-A case-study in Electronic Programme Guides', in *IJCAI*, p. 1351. Available at: https://pdfs.semanticscholar.org/d867/63fdf12bd4ee9a3bc0acb8f2a0df777cc316.pdf (Accessed: 1 November 2018).

Obrist, M., Cesar, P., Geerts, D., Bartindale, T. and Churchill, E. F. (2015) 'Online video and interactive TV experiences', *Interactions*, 22(5), pp. 32–37. doi: 10.1145/2799629.

Okazaki, S., Li, H. and Hirose, M. (2009) 'Consumer Privacy Concerns and Preference for Degree of Regulatory Control', *Journal of Advertising*, 38(4), pp. 63–77. doi: 10.2753/JOA0091-3367380405.

Oliver, D. (2014) *Privacy engineering: A dataflow and ontological approach*. Available at: https://dl.acm.org/citation.cfm?id=2686009 (Accessed: 7 July 2019).

Opinion 3/2010 on The Principle of Accountability (2010) *Article 29 Data Protection Working Party*. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2010/wp%0D173_en.pdf.

OwnCloud (2017) *ownCloud*. Available at: https://owncloud.org/features/ (Accessed: 1 January 2017).

Park, S., Kang, S., Chung, S. and Song, J. (2009) 'NewsCube: Delivering Multiple Aspects of News to Mitigate Media Bias', *Proceedings of CHI '09, the SIGCHI Conference on Human Factors in Computing Systems*, pp. 443–453. doi: 10.1145/1518701.1518772.

Phelan, C., Lampe, C. and Resnick, P. (2016) 'It's Creepy, But it Doesn't Bother Me', in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. New York, New York, USA: ACM Press, pp. 5240–5251. doi: 10.1145/2858036.2858381.

Pierson, J. and Heyman, R. (2011) 'Social media and cookies: challenges for online privacy', *info*. Edited by L. Van Audenhove, 13(6), pp. 30–42. doi: 10.1108/14636691111174243.

Pike, M., Ramchurn, R., Benford, S. and Wilson, M. L. (2016) '# scanners: exploring the control of adaptive films using brain-computer interaction', in *Proceedings of the 2016 CHI Conference on Human Factors in Computing*

Systems. ACM, pp. 5385–5396.

Pinnick, T. (2011) *Layered Policy Design | TRUSTe*, *trustarc.com*. Available at: http://www.truste.com/blog/2011/05/20/layered-policy-and-short-notice-design/ (Accessed: 7 July 2019).

Polat, H. and Du, W. (2005) 'SVD-based collaborative filtering with privacy', *Proceedings of the ACM Symposium on Applied Computing*, 1, pp. 791–795. doi: 10.1145/1066677.1066860.

Pollner, M. (2010) *Mundane reason: Reality in everyday and sociological discourse*. Cambridge University Press.

Project, L. (2012) *Locker Project*. Available at: https://github.com/quartzjer/Locker (Accessed: 7 February 2017).

Quevillon, F. (2011) *Derive*, *Docubase*. Available at: https://docubase.mit.edu/project/derive/ (Accessed: 5 July 2019).

Rada, R. (1995) 'Hypertext', in *Interactive Media*. New York, NY: Springer New York, pp. 21–47. doi: 10.1007/978-1-4612-4226-0_3.

Ramchurn, R., Martindale, S., Wilson, M. L. and Benford., S. (2019) 'From Director's Cut to User's Cut: to Watch a Brain-Controlled Film is to Edit it.', in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, p. 148.

Ramzan, N., Zwol, R. van, Lee, J., Clüver, K. and Hua, X. (2012) *Social media retrieval*. Available at: https://books.google.co.uk/books?hl=en&lr=&id=p81-pnyjODAC&oi=fnd&pg=PR5&dq=1.+Social+Media+Retrieval,+Naeem+Ramzan&ots=F2bpKHGlrV&sig=DjAUtxa055Q87AihCn5yV7oDy3c (Accessed: 5 July 2019).

Randall, D., Marr, L. and Rouncefield, M. (2001) 'Ethnography, ethnomethodology and interaction analysis', *Ethnographic Studies*, 6(1), pp. 31–44.

Randall, D. and Rouncefield, M. (2012) *Ethnography | The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Available at: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/ethnography (Accessed: 5 January 2018).

Rapp, J., Hill, R. P., Gaines, J. and Wilson, R. M. (2009) 'Advertising and Consumer Privacy', *Journal of Advertising*, 38(4), pp. 51–61. doi: 10.2753/JOA0091-3367380404.

Reding, V. (2012) 'The EU Data Protection Reform 2012 : Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age', *Innovation Conference Digital, Life, Design*, (January), pp. 1–6. Available at: http://europa.eu/rapid/press-release_SPEECH-12-26_en.htm.

Redmiles, E. M., Kross, S. and Mazurek, M. L. (2017) 'Where is the Digital Divide?', in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. New York, New York, USA: ACM Press, pp. 931–936. doi: 10.1145/3025453.3025673.

Reich, J. E. (2015) *Art Piece Fires Off Cap Guns Whenever There Is An IRL Drone Strike*, *Tech Times*. Available at: https://www.techtimes.com/articles/65726/20150703/art-piece-fires-cap-guns-whenever-irl-drone-strike.htm (Accessed: 5 July 2019).

Riedl, M. O. and Bulitko, V. (2013) 'Interactive narrative: An intelligent systems approach', *Ai Magazine*, 34(1), pp. 67–67. Available at: https://www.aaai.org/ojs/index.php/aimagazine/article/view/2449 (Accessed: 5 July 2019).

Robertson, A. (2017) *A Hidden Figures augmented reality app puts invisible scientific geniuses all over America*, *The Verge*. Available at: https://www.theverge.com/2017/1/8/14201852/outthink-hidden-figures-augmented-reality-app-ibm-nyt-ces-2017 (Accessed: 5 July 2019).

Robertson, T. and Wagner, I. (2015) 'CSCW and the Internet of Things', in *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway*. Cham: Springer International Publishing, pp. 285–294. doi: 10.1007/978-3-319-20499-4_15.

Robson, C. (1993) *Real world research: A resource for social scientists and practitioners-researchers*. 3rd Editio, *West Sussex: John Wiley & Sons*. 3rd Editio. West Sussex: John Wiley & Sons.

Robson, C. (2011) *Real world research*. Malden, MA: Blackwell Publishing. Available at: https://fenix.tecnico.ulisboa.pt/downloadFile/3779573194578/DEA_MetodologiasInvestigacao_programa_0910_v6.pdf (Accessed: 1 July 2019).

Rogers, Y. and Marshall, P. (2017) 'Research in the Wild', *Synthesis Lectures on Human-Centered Informatics*, 10(3), p. i-97. doi: 10.2200/S00764ED1V01Y201703HCI037.

Rogers, Y., Sharp, H. and Preece, J. (2002) *Interaction design*. John Wiley & Sons.

Room 202 (2016). Available at: https://www.oculus.com/experiences/rift/1074845635897855/ (Accessed: 5 July 2019).

Rose, J., Rehse, O. and Röber, B. (2012) *The Value of our Digital Identity*, *Boston Cons. Gr*.

Rosner, G. (2014) 'Who owns your data?', in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM.

Rosson, M. B. and Carroll, J. M. (2002) *Usability Engineering: Scenario-based development of human-computer interaction*. Morgan Kaufmann.

Ruoti, S., Andersen, J., Heidbrink, S., O'Neill, M., Vaziripour, E., Wu, J., Zappala, D. and Seamons, K. (2016) 'We're on the Same Page', in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, pp. 4298–4308. doi: 10.1145/2858036.2858400.

Sacks, H. (1984) 'Notes on Methodology', *Structures of social action: Studies in conversation*, pp. 21–7.

Sacks, H. (1992) *Lectures in Conversation*. Oxford, United Kingdom: Wiley-Blackwell.

Sailaja, N. and Jones, R. (2017) 'Industry Ideals and Barriers in Using Alternative Privacy Policies', in *Proceedings of the 31st British Computer Society Human Computer Interaction Conference*. BCS Learning & Development Ltd, p. 12. Available at: http://hci2017.bcs.org/wp-content/uploads/BHCI_2017_paper_64-1.pdf (Accessed: 15 September 2017).

Sarajevo Unlimited (2019). Available at: http://weareunlimited.ba/ (Accessed: 20 August 2019).

Sawaya, Y., Sharif, M., Christin, N. and Kubota, A. (2017) 'Self-confidence trumps knowledge: A cross-cultural study of security behavior', in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2202–2214. doi: 10.1145/3025453.3025926.

Schaub, F., Balebako, R., Durity, A. L. and Cranor, L. F. (2015) 'A Design Space for Effective Privacy Notices', in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pp. 1–17.

Schwab, K. (2016) *The Fourth Industrial Revolution: what it means, how to respond*, *World Economic Forum*. Available at: https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-

what-it-means-and-how-to-respond/ (Accessed: 6 July 2019).

Schwab, K., Marcus, A., Oyola, J. O., Hoffman, W. and Luzi., M. (2011) *Personal Data : The Emergence of a New Asset Class*.

Shklovski, I., Mainwaring, S. D., Skúladóttir, H. H. and Borgthorsson, H. (2014) 'Leakiness and creepiness in app space', in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pp. 2347–2356. doi: 10.1145/2556288.2557421.

Shotton, M., Melchior, F., Evans, M., Churnside, T., Brooks, M. and Armstrong, M. (2014) *Object-based broadcasting - curation, responsiveness and user experience*, *BBC*. Available at: https://www.bbc.co.uk/rd/publications/whitepaper285 (Accessed: 21 January 2019).

Si-soo, P. (2017) *Daily users of 'Pokemon Go' reach 4.2 million in S. Korea: survey*, *The Korea Times*. Available at: https://www.koreatimes.co.kr/www/news/tech/2017/01/134_222903.html (Accessed: 5 July 2019).

Silverberg, D. (2016) *How Two Pen Pals Learned About Each Other Through Data Mapping*, *VICE*.

Silverman, D. (2006) *Interpreting qualitative data: Methods for analyzing talk, text and interaction*. Sage.

Singel, R. (2009) *Netflix Spilled Your Brokeback Mountain Secret, Lawsuit Claims*, *Wired*. Available at: https://www.wired.com/2009/12/netflix-privacy-lawsuit/ (Accessed: 19 September 2018).

Singel, R. (2010) *NETFLIX Cancels Recommendation Contest after Privacy Lawsuit*. Available at: https://www.wired.com/2010/03/netflix-cancels-contest/ (Accessed: 19 September 2018).

Smyth, B. and Cotter, P. (2000) 'A personalised TV listings service for the digital TV age', *Knowledge-Based Systems*, 13(2), pp. 53–59.

Solove, D. J. (2008) *Understanding privacy*. Cambridge, MA: Harvard University Press. Available at: http://www.academia.edu/download/43383679/Understanding_Privacy.pdf (Accessed: 7 July 2019).

Sondhi, J. (2015) *Possibilia*. Available at: https://www.shortoftheweek.com/2016/08/03/possibilia/ (Accessed: 5 July 2019).

Spiekermann-Hoff, S. (2012) ‘The challenges of privacy by design’, *Communications of the ACM (CACM)*, 55(7), pp. 34–37. Available at: http://epub.wu.ac.at/id/eprint/5494 (Accessed: 7 July 2019).

Spiekermann, S. (2012) ‘The RFID PIA – Developed by Industry, Endorsed by Regulators’, in *Privacy Impact Assessment*. Dordrecht: Springer Netherlands, pp. 323–346. doi: 10.1007/978-94-007-2543-0_15.

Spiekermann, S. and Cranor, L. F. (2009) ‘Engineering Privacy’, *IEEE Transactions on Software Engineering*, 35(1), pp. 67–82.

Sterling, B. (2012) *Bruce Sterling Explains the Intriguing New Concept of Design Fiction (Interview by Torie Bosch)*, *Slate.com*. Available at: http://www.slate.com/ blogs/future_tense/2012/03/02/bruce_sterling_on_design_fictions_.html (Accessed: 9 February 2014).

Steuer, J. (1992) ‘Defining Virtual Reality: Dimensions Determining Telepresence’, *Journal of Communication*, 42(4), pp. 73–93. doi: 10.1111/j.1460-2466.1992.tb00812.x.

Stewart, K. A. and Segars, A. H. (2002) ‘An Empirical Examination of the Concern for Information Privacy Instrument’, *Information Systems Research*, 13(1), pp. 36–49. doi: 10.1287/isre.13.1.36.97.

Stone, B. (2010) *Ads Posted on Facebook Strike Some as Off-Key*, *The New York Times*. Available at: https://www.nytimes.com/2010/03/04/technology/04facebook.html.

Strauss, A. and Corbin, J. (1990) *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Supervisor, E. D. P. (1995) *Data Minimization Article 6.1(b) and (c) of Directive 95/46/EC*. Available at: https://edps.europa.eu/node/3080.

Supervisor, E. D. P. (2017) *Data Minimization Article 4.1(b) and (c) of Regulation EC (No) 45/2001*. Available at: https://edps.europa.eu/taxonomy/term/133.

Tam, K. Y. and Ho, S. Y. (2006) ‘Understanding the impact of web personalization on user information processing and decision outcomes’, *MIS quarterly*, pp. 865–890. Available at: https://www.jstor.org/stable/25148757 (Accessed: 5 July 2019).

Tanenbaum, J., Tanenbaum, K. and Wakkary, R. (2012) ‘Steampunk as design fiction’, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. Available at:

https://dl.acm.org/citation.cfm?id=2208279 (Accessed: 1 July 2019).

Taylor, A. and Harper, R. (2002) 'Switching on to Switch Off: an analysis of routine TV watching habits and their implications for Electronic Programme Guide Design', *UsableiTV*.

The Silent History (2012). Available at: http://www.thesilenthistory.com (Accessed: 5 July 2019).

The White House (2012) *We Can't Wait: Obama Administration Unveils Blueprint for a "Privacy Bill of Rights" to Protect Consumers Online | whitehouse.gov*, *The White House*. Available at: https://obamawhitehouse.archives.gov/the-press-office/2012/02/23/we-can-t-wait-obama-administration-unveils-blueprint-privacy-bill-rights (Accessed: 7 July 2019).

Thota, S. C. and Biswas, A. (2009) 'I Want to Buy the Advertised Product Only!: An Examination of the Effects of Additional Product Offers on Consumer Irritation in a Cross-Promotion Context', *Journal of Advertising*, 38(1), pp. 123–136. doi: 10.2753/JOA0091-3367380108.

Trepkeviciute, D. (2017) *Internet of Things (IoT) Driven Media Recommendations for Television Viewers. The Concept of IoT TV*. KTH Royal Institute of Technology School.

Tucker, C. E. (2014) 'Social Networks, Personalized Advertising, and Privacy Controls', *Journal of Marketing Research*, 51(5), pp. 546–562. doi: 10.1509/jmr.10.0355.

Tullio, J., Harboe, G. and Massey, N. (2008) 'Investigating the Use of Voice and Text Chat in a Social Television System', in *European Conference on Interactive Television*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 163–167. doi: 10.1007/978-3-540-69478-6_21.

UKI (2009). Available at: http://u-k-i.co/html/intro.html (Accessed: 5 July 2019).

UN (2018) *United Nations Global Pulse*. Available at: https://www.unglobalpulse.org/about-new (Accessed: 6 July 2019).

Ur, B., Habib, H., Johnson, N., Melicher, W., Alfieri, F., Aung, M., Bauer, L., Christin, N., Colnago, J., Cranor, L. F., Dixon, H. and Emami Naeini, P. (2017) 'Design and Evaluation of a Data-Driven Password Meter', in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pp. 3775–3786. doi: 10.1145/3025453.3026050.

Urquhart, L. (2016) *Towards User Centric Regulation : Exploring the*

Interface Between Information Technology Law and Human Computer Interaction. University of Nottingham.

Urquhart, L. (2018) 'Ethical dimensions of user centric regulation', *ACM SIGCAS Computers and Society*, 47(4), pp. 81–95.

Urquhart, L., Sailaja, N. and McAuley, D. (2017) 'Realising the Right to Data Portability for the Internet of Things', *SSRN Electronic Journal*. doi: 10.2139/ssrn.2933448.

Ursu, M. F., Kegel, I. C., Williams, D., Thomas, M., Mayer, H., Zsombori, V., Tuomola, M. L., Larsson, H. and Wyver, J. (2008) 'ShapeShifting TV: Interactive screen media narratives', *Multimedia Systems*, 14(2), pp. 115–132. doi: 10.1007/s00530-008-0119-z.

Ursu, M. F., Thomas, M., Kegel, I., Williams, D., Tuomola, M., Lindstedt, I., Wright, T. and Al., E. (2008) 'Interactive TV narratives: Opportunities, progress, and challenges', *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 4(4), p. 25. Available at: https://dl.acm.org/citation.cfm?id=1412198 (Accessed: 21 January 2019).

Usmani, W. A., Marques, D., Beschastnikh, I., Beznosov, K., Guerreiro, T. and Carriço, L. (2017) 'Characterizing Social Insider Attacks on Facebook', in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. New York, New York, USA: ACM Press, pp. 3810–3820. doi: 10.1145/3025453.3025901.

Verbeek, P.-P. (2011) *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.

Wang, Y., Leon, P. G., Acquisti, A., Cranor, L. F., Forget, A. and Sadeh, N. (2014) 'A field trial of privacy nudges for facebook', *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pp. 2367–2376. doi: 10.1145/2556288.2557413.

Warren, S. D. and Brandeis, L. D. (1890) 'Right to privacy', *Warren, Samuel D., and Louis D. Brandeis. 'Right to privacy.' Harv. L. Rev.*, 4, p. 193. Available at: https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/hlr4&section=31 (Accessed: 6 July 2019).

Weiser, M. and Brown, J. S. (1997) 'The coming age of calm technology', *Beyond calculation*, pp. 75–85. Available at: https://link.springer.com/content/pdf/10.1007/978-1-4612-0685-9_6.pdf (Accessed: 7 July 2019).

Westin, A. F. (1968) 'Privacy and freedom', *Washington and Lee Law Review*,

25(1), p. 166. Available at: http://scholarlycommons.law.wlu.edu/cgi/viewcontent.cgi?article=3659&context=wlulr (Accessed: 11 January 2017).

White, T. B., Zahay, D. L., Thorbjørnsen, H. and Shavitt, S. (2008) 'Getting too personal: Reactance to highly personalized email solicitations', *Marketing Letters*, 19(1), pp. 39–50. doi: 10.1007/s11002-007-9027-9.

Wittgenstein, L. (2009) *Philosophical Investigations*. John Wiley & Sons.

World Economic Forum (2018) 'Data Policy in the Fourth Industrial Revolution: Insights on Personal Data', (November).

Wright, D., Gutwirth, S., Friedewald, M., Hert, P. De, Langheinrich, M. and Moscibroda, A. (2009) 'Privacy, trust and policy-making: Challenges and responses', *Computer law & security review*, 25(1), pp. 69–83.

Wright, D. and Hert, P. De (2011) *Privacy Impact Assessment*. Available at: https://books.google.co.uk/books?hl=en&lr=&id=I9-NlkBGQjYC&oi=fnd&pg=PR5&dq=privacy+impact+assessment&ots=WG38LXjbcy&sig=6Se1-ikShgK6DUHwn91OMPrvCCk (Accessed: 7 July 2019).

Xin, Y. and Jaakkola, T. (2014) 'Controlling privacy in recommender systems', *Advances in Neural Information Processing Systems*, 3(January), pp. 2618–2626.

Zhang, M. (2010) *Museum of London Releases Augmented Reality App for Historical Photos*, *PetaPixel*.

Zynga (2011) *PrivacyVille*. Available at: https://www.zynga.com/privacy/privacyville (Accessed: 18 September 2017).

Appendices

Appendix I : Questionnaires

Study I : Questionnaire I (Filled in before the session begins)

To be filled in before the discussions begin. For each of the following questions, please choose one or more of the choices as applies.

Which of these devices do you watch television on?
- TV
- Laptop
- Desktop
- Phone
- Tablet
- Other_____

At what times of the day do you watch television in the weekdays?
- Morning
- Afternoon
- Evening
- Night
- Other_____

At what times of the day do you watch television in the weekends?
- Morning
- Afternoon
- Evening
- Night
- Other_____

Do you use Video on Demand :
- Daily
- Weekly
- Monthly
- Never
- Other_____

Do you use Electronic Programme Guides :
- Daily
- Weekly
- Monthly
- Never
- Other_____

Do you mainly watch television alone or with others?
- Alone

- With partner
- With children
- With friends
- With Others_____

Where do you mostly watch television?
- Living room
- Bedroom
- Kitchen
- Other_____

What other activities do you do while watching TV?
- Cooking
- Reading
- Social Media
- None
- Other_____

Do you use a Smart TV to do any of the following?
- Social networking
- Video On Demand
- Games
- Other_____

How do you recommend TV programmes to others?
- Word of mouth
- Social media
- Other_____

Study I : Questionnaire II (Filled in after the session is closed)

To be filled in after the discussions are completed. For each of the following functionalities proposed in the discussion, please indicate how useful you find them.

Essential *Unnecessary*

User profiles |_____|_____|_____|_____|

Essential *Unnecessary*

Single user profile|_____|_____|_____|_____|

Essential *Unnecessary*

Multiple user profile|_____|_____|_____|_____|

Essential *Unnecessary*

Categories|_____|_____|_____|_____|

Essential *Unnecessary*

Setting priority |_____|_____|_____|_____|

Essential *Unnecessary*

Variable length videos |_____|_____|_____|_____|

Essential *Unnecessary*

Device selection |_____|_____|_____|_____|

Essential *Unnecessary*

Recommender option |_____|_____|_____|_____|

Essential *Unnecessary*

Multichannel integration|_____|_____|_____|_____|

Essential *Unnecessary*

Social Groups |_____|_____|_____|_____|

Essential *Unnecessary*

Social networks |_____|_____|_____|_____|

Essential *Unnecessary*

Granular data control |_____|_____|_____|_____|

Essential *Unnecessary*

Video on Demand |_____|_____|_____|_____|

Three Electronic Programme Guide variations were presented to you today. Which of these would you prefer?

- o Current
- o Calendar
- o Recommender

Would you like to add or omit any functionalities proposed in your chosen version of EPG?

Study III : Questionnaire (Filled in before the session begins)

Living Room of the Future

Pre-Experience Questionnaire

Age :

Gender :

Ethnicity :

Occupation :

Why did you come to FACT today?

_____
_____
_____

How did you find out about the Living Room of the Future today?

_____
_____
_____

What interests you about the Living Room of the Future?

_____
_____
_____

Have you heard about the Internet of Things (IoT) before today?
□ Yes □ No

Do you own any of the following Internet of Things (IoT) devices?

□ Alexa □ SmartWatch □ SmartPhone □ GoogleHome

□ Other. If so, please specify_____

Does data privacy concern you?

□ Yes □ No □ Maybe □ Don't Care □ Not Aware

Do you usually watch TV alone or with others?

_____
_____
_____

For Researcher Use:

Participant Number :_____Group Number :_____Consent Form Signed :_____
Notes:
_____
_____

Information Sheet and Consent Form used for EPG Focus Groups

Scenario-based Exploration of Adaptive EPG

Participant Information Sheet and Study Consent Form

The study is a short evaluation of a new type of Electronic Programme Guide (EPG) being explored by researchers at BBC R&D and the University of Nottingham. You will be asked to complete a short questionnaire on your current viewing habits and then be introduced to the novel "Adaptive EPG" and its use of personal data to deliver content tailored to you, your family, and friends. The study will take no longer than 45 minutes. You are encouraged to ask questions and comment on the Adaptive EPG as we present it and you will be interviewed about your views on the technology and its possible ramifications after the presentation is completed. You will also be asked to complete another short questionnaire on completion of the interview rating features of the Adaptive EPG.

The presentation and interview will be audio-recorded to enable us to capture your views in detail and to preserve them for analysis. The audio data will be transcribed and analysed to identify issues and themes raised by the participants. Extracts from the data – i.e., quotations – may be used to elaborate issues and themes in written reports, research presentations, and scientific publications including the researcher's PhD thesis. All transcripts and quotations will be anonymized, and participants' identities will not be disclosed in any research presentations, written reports or publications.

The data will not be put on the Internet. It will be kept securely on a password and firewall protected drive at the University of Nottingham in accordance with the requirements of the 1998 Data Protection Act. The raw (non-anonymized) audio data and questionnaires will be stored for a maximum of 7 years after the

publication of any works based on it and then be destroyed. The questionnaire data and anonymized transcriptions may be reused in subsequent research activities. You are free to withdraw from the research at any time, in which case any data provided by you will also be withdrawn from the research. Should you wish to withdraw, or have any queries, please contact:

Neelima Sailaja,
Horizon Centre for Doctoral Training,
University of Nottingham,
Jubilee Campus, Wollaton Road,
Nottingham, NG8 1BB
Email: neelima.sailaja@nottingham.ac.uk

Please tick the check boxes below as appropriate and sign if you agree to take part in the research.

❑ I have read and understand the information sheet, including the data to be recorded.

❑ I understand that I can withdraw at any time by contacting the researcher at the address provided, and that my data will not be used in the research.

❑ I confirm that I am over the age of 18.

Participant's Details:

Signature: _____

Date: _____

Consent Form used for BBC Interviews

Participant Study Consent Form

The interview is an effort to understand service provider views on personal data capture, use and barriers to further use across various teams at the BBC. It is done as part of digital economy research being conducted at the University of Nottingham and BBC R&D. You will be asked a few questions regarding the methods, motivations and orientations associated with current personal data capture and use trends. The interview will take no longer than 1 hour. It will be an open-ended discussion which aims to explore service provider views on the collection and use of personal data in the delivery of innovative media services.

The interview will be audio-recorded to preserve your feedback for subsequent analysis. The audio data will be transcribed and analysed to identify issues and themes raised across all the participants in this study. Extracts from the data – i.e., quotations – may be used to elaborate issues and themes in written reports, research presentations, and scientific publications including the researcher's PhD thesis. All transcripts and quotations will be anonymized, and participants' identities will not be disclosed in any research presentations, written reports or publications.

The data will not be put on the Internet. It will be kept securely on a password and firewall protected drive at the University of Nottingham in accordance with the requirements of the 1998 Data Protection Act. The raw (non-anonymized) audio data will be stored for a maximum of 7 years after the publication of any works based on it and then be destroyed. You are free to withdraw from the research at any time before publication, in which case any data provided by you will also be withdrawn from the research. Should you wish to withdraw, or have any queries, please contact:

Neelima Sailaja,
Mixed Reality Lab,
School of Computer Science,
University of Nottingham,

Jubilee Campus, Wollaton Road,
Nottingham, NG8 1BB
Email: neelima.sailaja@nottingham.ac.uk

Please tick the check boxes below as appropriate and sign if you agree to take part in the research.

❑ I have read and understand the information sheet, including the data to be recorded.

❑ I understand that I can withdraw at any time before publication of the research by contacting the researcher at the address provided, and that my data will not be used in the research.

❑ I confirm that I am over the age of 18.

Participant's Details:

Name: _____

Signature: _____

Date: _____

Researcher's Details:

Name: _____

Signature: _____

Date: _____

Information Sheet used for BBC Interviews

Participant Information Sheet

The interview is an effort to understand service provider views on personal data capture, use and barriers to further use across various teams at the BBC. It is done as part of digital economy research being conducted at the University of Nottingham and BBC R&D. You will be asked a few questions regarding the methods, motivations and orientations associated with current personal data capture and use trends. The interview will take no longer than 1 hour. It will be an open-ended discussion which aims to explore service provider views on the collection and use of personal data in the delivery of innovative media services.

The interview will be audio-recorded to preserve your feedback for subsequent analysis. The audio data will be transcribed and analysed to identify issues and themes raised across all the participants in this study. Extracts from the data – i.e., quotations – may be used to elaborate issues and themes in written reports, research presentations, and scientific publications including the researcher's PhD thesis. All transcripts and quotations will be anonymized, and participants' identities will not be disclosed in any research presentations, written reports or publications.

The data will not be put on the Internet. It will be kept securely on a password and firewall protected drive at the University of Nottingham in accordance with the requirements of the 1998 Data Protection Act. The raw (non-anonymized) audio data will be stored for a maximum of 7 years after the publication of any works based on it and then be destroyed. You are free to withdraw from the research at any time before publication, in which case any data provided by you will also be withdrawn from the research. Should you wish to withdraw, or have any queries, please contact:

Neelima Sailaja,
Mixed Reality Lab,
School of Computer Science,
University of Nottingham,

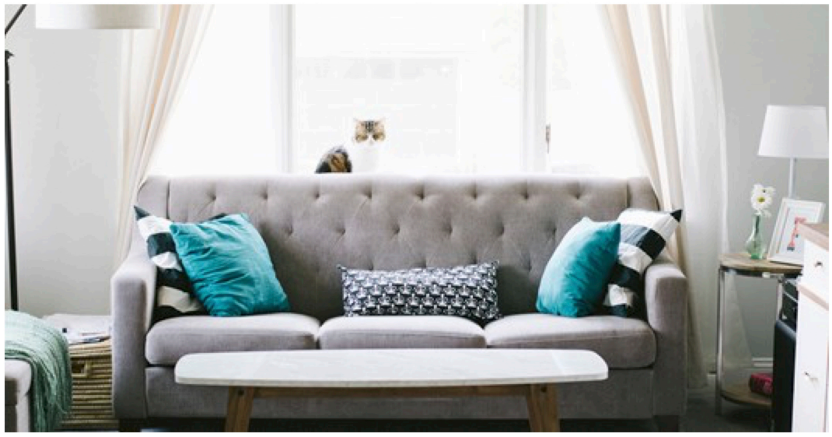
Jubilee Campus, Wollaton Road,
Nottingham, NG8 1BB
Email: neelima.sailaja@nottingham.ac.uk

Consent Form used for LRoTF Study

Participant Information Sheet [/ /]

The Living Room of the Future Evaluation Study

University of Nottingham
UK | CHINA | MALAYSIA



The purpose of this study is to evaluate the user experience of the Living Room of the Future, which exploits data from Internet-enabled Things situated in the living room to deliver a new immersive media experience.

The purpose of this information sheet is to inform you what you will be required to do should you choose to participate.

- **1.** We would like to know more about the kind of people who have taken part in the experience and what motivates them to do so. To find out about you we want you to fill in a short questionnaire.
- **2.** We would like to know how people interact and react within The Living Room of the Future. To find out we need several volunteers who are willing to let us record their experience on audio/video so we can see what happens in The Living Room of the Future.
- **3.** We would like to know what you think and feel about The Living Room of the Future. To find out we invite you to participate in a short interview after you have experienced The Living Room of the Future, which will be audio recorded.

Your participation in our evaluation of The Living Room of the Future is entirely voluntary. You are not obliged to participate in any part of this study. Alternatively, you may only wish to engage with parts of it, e.g., parts 1 & 3 but not part 2. If you decide to participate, it is entirely your choice what you contribute. You can tell us what you want to do in the consent form.

Should you choose to take part in the study, you have the right to withdraw from it at any point without explanation and to have the data you have contributed destroyed, or your identity redacted if it involves other people who have also consented to take part in the study but have not withdrawn their consent.

The raw (audio/visual) data we collect from you will be stored on an encrypted hard drive. Processed (transcribed) data will be stored on our secure servers, along with questionnaire data. At no time will raw data be put online and made publically available.

Extracts from the data may be used in research works, including presentations at research events and in research publications that are hosted online. Wherever possible data extracts will be anonymous and will not identify you. However, if you agree to participate in part 2 of this study, data extracts may include potentially identifiable visual images of yourself.

Contact details are provided at the bottom of this document, should you wish to exercise your rights, and please feel free to ask us any questions you might have.

Information Sheet used for LRoTF Study

INFORMATION TO BE PROVIDED WHERE PERSONAL DATA ARE COLLECTED FROM THE DATA SUBJECT (GDPR ARTICLE 13)

Identity and Contact Details of the Data Controller (1a/b). Should you have any queries about your data or wish to exercise your statutory rights please contact the data controller.

The University of Nottingham is registered as a Data Controller under the Data Protection Act 1998 (Registration No. Z5654762, https://ico.org.uk/ESDWebPages/Entry/Z5654762). The University of Nottingham, University Park, Nottingham, NG7 2RD, is committed to protecting your personal data and informing you of your rights in relation to that data. Accordingly, the University will process your personal data in accordance with the General Data Protection Regulation (GDPR) and the Data Protection Act 2018. If you require advice on exercising any of your rights, please contact the University's data protection team on data-protection@nottingham.ac.uk.

Purposes of Data Processing (1c). Your data is being processed for the purposes of evaluating the user experience of The Living Room of the Future.

Legal Basis of Data Processing (1c/d). The legal basis for processing stands on your consent.

Recipients of the Data (1e). Your data may be archived and disclosed to other researchers for historical, scientific or statistical purposes. Extracts of your data may be disclosed in published works posted online for the scientific community.

Data Transfer (1f). Your data will not be transferred outside of the United Kingdom for processing.

Period for which Personal Data will be Stored (2a). Your data may be stored securely for an indefinite period in accordance with GDPR Article 5(b) for scientific, historical or statistical research purposes, insofar as data storage complies with Article 5(e) and appropriate safeguards are put in place; see 'further processing of the data' for specification of appropriate safeguards.

Your Rights (2b). You have the right to object to data processing and to request from the controller:

- Confirmation as to whether or not personal data are being processed and, if so, to access the data.
- Rectification of inaccuracies without undue delay.
- Restriction of further processing pending investigation of any complaint.
- Data portability and to receive a copy of the data in an accessible format.
- Erasure of the data and to be 'forgotten'.

Some caveats do apply including providing proof of identity and paying an administration fee for data access requests; see Articles 15, 16, 17, 18, 19 and 20 for full details https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:119:FULL&from=EN

The Right to Withdraw Consent (2c). Where the processing is legally based on consent, you have the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal.

Your Right to Lodge a Complaint with a Supervisory Authority (2d). You have right lodge a complaint with the Information Commissioner's Office at any time regarding the controller's treatment of your personal data. You can contact the ICO on 0303 123 1113 or online https://ico.org.uk/global/contact-us/.

Provision of Personal Data (2e). You are not obliged to provide data for this study. The data is not part of a statutory or contractual requirement, or requirement necessary to enter into contract.

Automated Processing and Profiling (2f). Your data will not be subject to automated processing or profiling.

Further Processing of the Data (3/4). Your data may be archived and reused in accordance with Article 5(b) for other scientific, historical or statistical purposes; these purposes are not considered to be incompatible with the initial purposes for which the data were gathered and further information as specified in Article 13(4) is therefore not required prior to further processing. However, if your data is archived appropriate safeguards must be put in place in accordance with Article 5(e). Safeguards in this case consist of storing identifiable data on a non-networked hard drive in a secure location and applying 256-bit encryption to protect the data at rest. Identifiable data will be not be archived if consent has not been given.

For any research questions, comments, requests or enquiries please contact the researcher at :

neelima.sailaja@nottingham.ac.uk