Artificial Intelligence, Museum Environments and their Constituents: A Cross-disciplinary Study Using Recommender Systems to Explore Digital Collections

A thesis submitted to The University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Humanities

2022

Lukas B. M. Hughes-Noehrer

School of Arts, Languages and Cultures

Table of Contents

Word Count: 50,941

List of Tables

List of Figures

Abbreviations

AAT	Art & Architecture Thesaurus
AI	Artificial Intelligence
BIPOC	Black, Indigenous, and people of colour
CC0	Creative Commons 0
CH	Cultural Heritage
CMS	Collection Management System
CNN	Convolutional Neural Network
DS	Data Science
GIS	Geographical Information System
HCI	Human-computer interaction
IPCC	Item-Based Pearson Correlation
MAG	Manchester Art Gallery
ML	Machine Learning
MVP	Minimum Viable Product
NASA TLX	NASA Task Load Index
NLP	Natural Language Processing
NMAAHC	National Museum of African American History and Culture
NMAH	National Museum of American History
NG	National Gallery (London)
OCR	Optical Character Recognition
OWL	Ontology Web Language
PIS	Participant Information Sheet
PoI	Point of Interest
RDF	Resource Description Framework
RFID	Radio-frequency identification
RS	Recommender System
SAAM	Smithsonian American Art Museum
SI	Smithsonian Institution
SLT	Senior Leadership Team
SOTA	State of the Art
UGC	User-generated content
UE	User Engagement
UI	User Interaction
UPCC	User-Based Pearson Correlation
UX	User Experience
W3	World Wide Web

Glossary

Big Data	Datasets exceeding usual processing powers or very complex datasets
Edge Intelligence	Defines data capture and processing close to the devices used to gather the data, such as IoT devices or even smartphones and postulates that data should be kept close to its origin (the “edge device”)
Explainability	Term used in AI/ML that refers to models that are comprehensible to humans; often used together with trust and privacy
Folksonomy	User-generated system of classification and organisation of online content
Interactive (noun)	A system that takes in user input and returns an output accordingly
ICONCLASS	An image classification system
Spectrum Standards	Collection management standards set by the Collections Trust
Web of Data	Also known as Semantic Web or Web 3.0 describes a web-standard that aims to make data on the Internet machine-readable
WordNet	Lexical database hosted by Princeton University

Abstract

This thesis explores the contingent relationships between museum environments, their constituents, and AI technologies. It considers the challenges, issues and effects that arise when such technologies are applied to a museological context, through practice-based and empirical research on the museum sector. It asks questions about the role and potential of AI technologies in museums and their data practices and investigates in what ways AI challenges and/or enhances the public and museum professionals' perception of AI.

To answer these questions, the thesis applies a novel combination of methods, including the development of a new recommender system to research AI technologies as experienced through their cultural engagement in public museums, positioning the institutions as interactive laboratories. This approach investigates the applicability and usability of algorithmic outputs in museum settings, addressing trust issues, testing new strategies, exploring content creation and the implications of its future use in a technically informed society. During this process, the recommender system is not perceived as definitive, but as an evolving object that is transient, changing and question-generating, establishing the RS both as a system for curating online museum experiences and a method in its own right. The research is informed by an empirical-philosophical framework forged out of a postphenomenological vocabulary which enables investigation of the socio-technical and cultural roles of the recommender system and its constituents through the relation humans have with technological artefacts.

The thesis argues that AI needs to create value and become significant to constituents to become a more sustainable practice within museum environments and to translate its full potential onto the practices of institutions. Such practices afford collaborative approaches to harness the power of AI and address the challenges of pervasiveness and ubiquity of those technologies inherit, which can lead to mistrust and avoidance. The research concludes by confirming the contingency between museum environments, their varied constituents, and AI technologies; its findings have implications for museum practice and present a unique contribution to a developing interdisciplinary field.

Declaration

I hereby declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Acknowledgements

I would not have been able to write this thesis without the support and guidance of many great individuals over the last three years. This project has not just contributed to my knowledge and strengthened my research abilities, but also provided me with the opportunity to develop strong networks with people I now call friends and collaborators.

First and foremost, I would like to sincerely thank my two supervisors, Dr Abigail Gilmore and Prof Caroline Jay, for their continuous support, encouragement, and guidance throughout this PhD. This project would not have been possible without their belief in me and my work, and their uncountable hours of motivation and inspiration. Most importantly though, thank you, Abi and Caroline, for listening to me, for understanding me, and for establishing a safe space to conduct my research in - you were invaluable to this thesis.

I further want to greatly thank my internal reviewer, Prof Uli Sattler, for her ongoing advice and our thought-provoking *Kaffeeklatsch* meetings as well as Dr Jonathan Carlton, who I am honoured to not just call a best friend, but also a mentor - I will always appreciate your support and help.

I would also like to say thank you to the members of the Interaction Analysis and Modelling lab and the team of the Institute for Cultural Practices for giving me a scholarly home and a space of fruitful discussion and learning, as well as the Engineering and Physical Sciences research council and the University of Manchester for their financial support.

Further thanks and gratitude go to the professionals working at the Smithsonian, Manchester Art Gallery and Manchester Museum Partnership, the National Gallery in London, and Art UK. Thank you for your trust in me and for opening your digital stores, but also thank you for your guidance and advice, the discussions we had, and the friendships that grew out of this PhD over the years.

I would also like to thank my family and friends for their unconditional outside-of-PhD support and the times so valuable to unwind and let my mind breathe. I know that most of you are still confused about the nature of my PhD - Museology, Computer science, a Netflix for museums? - do not worry, so was I at times, but I hope this thesis will add clarity.

Lastly, I want to thank my husband Craig with all my heart. He is the bedrock of so many things in my life as he was for this thesis. Thank you for putting up with me and my night time reading light, my early mornings to get up and write, and most importantly thank you ever so much for always being there for me. Diolch yn fawr, o waelod fy nghalon – dwi’n dy garu di.

1. Introduction

This thesis concerns the role that AI technologies can play in enhancing the value of museums and art galleries[1] as both producers and custodians of knowledge. As society takes on an ever more digitally connected semblance, where technology is ubiquitous, museums face increasing demands to offer content online, particularly since the start of the SARS-CoV-2[2] pandemic (ICOM, 2020a). This progression provides an opportunity to make collections more accessible and to engage a broad audience for education, entertainment, and scholarly purposes inside and outside of the physical institution. Nevertheless, given their vast collections, the task of creating individually tailored online pathways that capture the tastes and interests of users, but also facilitate new ways of exploring collections and support new meaning-making in a networked society, poses an additional challenge for museums (S. Anderson, 2020). A potential solution is the enhancement of browsing online collections through the application of Artificial Intelligence (AI) technologies, such as recommender systems (RS).

This thesis therefore investigates the application of an RS in museums to facilitate access to online collections through personalised pathways and to explore digital collections with AI. Anchored in practice-based and empirical research, the thesis aims to shed light on relations between the triad of constituents, technologies, and the museum environment through the development of a web application content-based RS. The application functions as the core of research interventions conducted for this thesis and interaction with it builds the main node for theoretical discussions and empirical examination. The development of the system is seen as a form of practice research through its iterative design and implementation process, which is both informed and critically interrogated by the accompanying research interventions consisting of a survey, focus groups and an online user study bringing the system's development into close encounter with museum professionals and other constituents, who have a stake in the organisation of, and access to, online collections as discussed below.

[1] The terms 'museum', 'art gallery', and 'institution' are used interchangeably throughout the thesis.

[2] Commonly known as COVID-19.

A research field on their own, although generally very closely associated with AI and machine learning (ML) technologies, RS are founded in approximation theory, cognitive science, information retrieval, and forecasting theories (Adomavicius and Tuzhilin, 2005). They are used on a daily basis by millions of people in well-known media applications, such as Netflix (Gomez-Uribe and Hunt, 2016), Spotify (Jacobson et al., 2016), Instagram (Medvedev et al., 2019) and YouTube (Covington et al., 2016), in e-commerce (Rastogi, 2018; Smith and Linden, 2017), broadcast and news platforms (Boididou et al., 2021), and in medical sciences (Oumaima et al., 2020). With such a varied field of application, RS have become a major driving force of learning and knowledge discovery, as a means for filtering through abundant information, and as invaluable marketing and sales tools. Gaining popularity since the 1990s and the rise of the Web, they are now the bedrock of many personalised user experiences.

It is the possibility of individual, personalised access to information that separates RS from traditional information retrieval techniques and search engines (Burke, 2002), enabling new ways to explore the wealth of digital data held by institutions. Access is both important to the Internet and the museum field alike. However, many museums struggle to offer ways to engage online with collections beyond search bars, making it hard for constituents to fully enable the wealth that lies beneath them (Falkowski, 2016). Furthermore, the majority of research on the searchability of museum collections concerns *what was done* without asking *does it work* (Solas, 2010). This thesis aims to address this critical gap by investigating the interaction of constituents with a museum recommender system for online collections and assessing its implications and usefulness to their wide range of values and interests.

The main aim of RS is to serve recommendations to users based on some form of user interaction. This can be explicit, requiring users to give direct feedback, such as ratings or likes, or devised on the basis of implicit interaction events, such as clicks and dwell times. With their ability to choose items out of abundant forms of data and multiple options to combine algorithms, RS present a means not just to explore museum online collections, but also to facilitate the creation of new knowledge and support meaning-making beyond traditional methods that lack means to take users' preferences and advances in AI research into account. Apart from being a scalable solution to support the rapid advancement of access to digital collections, RS gather data on users' behaviour to infer valuable metrics and insights, informing future technological developments and strategies of museums. Being a point of interaction that processes, but also collects data, RS could provide the means to explore the broader relationships between museum environments, their constituents, and the technologies that are shaped by and constitutive of each other through their networked relationships, suggesting that the RS is not just a technological artefact, but a research method in its own right.

RS have been applied in the museum and heritage sector before. However, investigating the history of recommender system development for museums, it is evident that most systems were designed for in-house user experiences and on-site guided tours, focussed "either on gallery-like presentations and/or linear narratives" (Pavlidis, 2019, p. 193). Whereas RS are suitable approaches to enhance on-site visitor experiences, those systems do not necessarily cater to "hybrid spaces" (Henning, 2006, as cited in Kidd, 2016, p. 3) which blend in-person with digital experiences and have affordances different to those concerned solely with the physical space. Digital media offer new opportunities to engage with collections and create alternative experiences and novel modes of exploration (Newell, 2012) and the digital medium, just as the museum itself, needs scrutinising in terms of what consequences its use has for constituents as the "meanings (the agency) associated with any given media are never fixed, but rather that they change according to the experiences and knowledge of an individual and the shifting values and discourses of any given community or society" (Parry, 2007, p. 10).

To embed the research in a theoretical framework, the thesis draws upon postphenomenology, a philosophy of technology and technoscience, mainly founded in hermeneutical phenomenology and American pragmatism, which attributes to technologies an active role in mediating and shaping the relationships between human beings and the world around them (Verbeek, 2005). Postphenomenology asserts that "technology does something" (Verbeek, 2005, p. 66) and provides tools to develop a way of rendering technical or technological objects as mediators between human beings, reality, and the entanglement of humans with technology.

This thesis therefore aims to draw a holistic picture of the RS application and its implications, addressing all stages, from early design to implementation and evaluation, situated in a postphenomenological framework that enables HCI research that does not render technologies as passive entities, but as actively shaping the environments they are situated in (Verbeek, 2005). The introduction of the system into the environment of the museum and its constituents conceives the RS as a relational artefact that aims to stimulate discussions whilst functioning as a conversational junction linking various threads together, aiming to be an igniter of thoughts and a placeholder for wider discourse about AI technologies in general and their agency in museum spaces.

There has been extensive discussion of the agency of AI technologies and the notion of technologies not being neutral (Floridi et al., 2018; Mittelstadt et al., 2016), but being active mediators of the relations between human beings and reality, altering their environment and therefore coproducing the material realities and perceptions of the spaces they are active in (Verbeek, 2005). The notion of agency adds another layer of context to this thesis as AI technologies, their explainability and transparency and issues of data trust and privacy are major concerns (Veale and Binns, 2017). The pivotal role of museums as trusted institutions responsible for stewarding, holding, and communicating knowledge, providing education (Falk and Dierking, 2018) and creating new knowledge offers an opportunity to test technologies with the public and educate constituents about AI. Such interventions could contribute towards more explainable and transparent technologies and an informed usership as users often do not know about their existence or functionality (The Royal Society, 2017).

The research which supports this thesis is practice-based, and this thesis is accompanied by a portfolio (please find the portfolio[3] here) providing a personal,

[3] For readers of a hardcopy, please follow: https://tinyurl.com/lhnthesisportfolio

reflective account alongside presentation and documentation of the development of the recommender system *MuseREC*, first introduced in Noehrer et al. (2021a). The system's pre-processing, processing and post-processing stages are outlined in the *Five-step recommender life cycle* (see *Portfolio*, p. 7), which was developed for this thesis and creates the scaffolding of the portfolio and the practical work. The life cycle draws upon established data science pipelines and lays out the necessary steps and areas of consideration to render a satisfying picture of the application in the museum sphere. The life cycle was further informed by guidelines and forms of standardisation (Wilkinson et al., 2016) to ensure rigorous development, use, and evaluation of the system and guarantee careful consideration of ethical concerns, pitfalls, and potential harms of such systems from an ML (boyd and Crawford, 2012; Suresh and Guttag, 2020) and a museum perspective (S. Anderson, 2020).

To enable an audience to explore collections personalised to their tastes and recent activity, along with investigating the recommendations, *MuseREC* is a content-based RS using collection data, which includes metadata and images. The system, formed of two components - the content-based model and the web system (with the latter creating a wrapper around the former) - serves recommendations to users based on their interactions with items in the collection, enabling personalised pathways. Items in this specific case are metadata and images of objects from the Art UK[4] collection.

To construct the recommendation model, an NLP technique, *word2vec,* was applied to the rich information in the collection metadata, such as the titles and descriptions (please refer to *Portfolio,* p. 26). Further, a Convolutional Autoencoder - a deep learning method that outputs a reconstructed input - was trained using high-resolution digital images of the objects to learn low-dimensional feature representations (see *Portfolio*, p. 29). Both methods output feature vectors for every item in the collection - a feature vector can be seen as a mathematical representation - a list of numbers - of an image or the metadata or a concatenated version of the two (see *Portfolio,* p. 32). As there are no targets

[4] For more information about Art UK, please visit https://artuk.org

to optimise for, nor user preference information to construct a collaborative or hybrid model, a similarity search algorithm was applied (see *Portfolio,* p. 33) to produce recommendations of similar items.

The thesis and its portfolio together form an attempt to bring the discrete disciplinary fields of recommender systems and research on museum environments and practices together, and to render the ensuing knowledge comprehensible for audiences from both computer science and museology. This is no easy feat, and to help with the translation some of the key terms which are central to this thesis are explained below.

1.1. Coming to terms with the terminology

The very first section has introduced a specific terminology that will reappear throughout this thesis. The following paragraphs shall therefore establish a common understanding of these words to contribute to clarity and avoid misunderstandings. The established terminology should cater to how these terms are understood and are not intended to be exhaustive definitions with universal applicability.

Museum environments

The term *museum environments* is used here to encompass all those factors that establish the various embodiments and perceptions of a museum, rather than rendering it as a fixed-in-time institution. These environments are considered “multistable” (Ihde, 2012) in the sense that they are subjected to change, are shaped by the constituents who act inside and around them and other mediating factors such as technologies. Etymologically, the term also caters to the technological environment of the thesis as it is deeply ingrained in the networked structures of modern computation, the Internet, and W3, which all further constitute museum environments through their connectedness to a variety of media. This establishes the museum environment as a sociotechnical one with infrastructures that are more or less visible, and that even operate in the background. It is understood as an environment with “no clean edges, firm boundaries, or absolutely exterior positions" (Wiltse, 2017, p. 10). First and foremost, albeit the same essence or fabric, they can be perceived and used differently and appear in variations. Multistability of those environments, however, is not infinite as museums are subjected to specific value systems, practices, and transactions and they are organised in frameworks of micro levels, such as their strategies, missions, and visions and macro levels, such as laws, regulations, and sector standards.

Constituents

The terms audience, visitor, user, consumer, professional, and creator, amongst others, describe the wide field of humans interacting with museums and are often interchangeably used throughout scholarship and practice with varying definition and levels of exclusivity. This thesis uses the term *constituents* as an all-encompassing term to reflect all human beings entering, residing, and departing the museum environment. This includes not just various types of individuals traditionally classed as *visitors*, but also all types of professional roles in the sector. Constituents make up the very fabric of the museum and are an integral part in determining its structure as well as its being. Constituents further have powers vested in them that may or may not be exercised and they can take on active as well as passive forms of agency. They are as Byrne summarises

> *"fluid, mutable, protean. They grow, change, adapt, hybridize and reform according to circumstance and need. As such, constituencies, as well as the status of being constituent, are always in the process of both becoming and unbecoming - constituencies result from a process of social production whose mediums and vehicles are, of necessity, collaborative"* (Byrne, 2014, p. 27).

The multifaceted nature of constituents further accepts that one and the same constituent can take on various roles and viewpoints that can rapidly change due to spatiotemporal factors as well as extrinsic and intrinsic motivations. The term also refers to museum professionals, policy makers, or funding bodies, as they are all involved in meaning-making in relation to their stakes in the museum as constituents.

Artificial Intelligence (AI)

In this thesis, AI is understood as computational operations that are able to learn from data and perform problem-solving tasks that usually require capabilities expressed by humans. The field of AI is deliberately defined widely and incorporates symbolic (e.g., rule-based, semantic technologies, and knowledge graphs) and non-symbolic systems (e.g., machine learning, deep learning, neural networks, natural language processing, and data mining). It also acknowledges that participants in any of the thesis' research interventions might have a different understanding and their definition may include data science in general and other computationally-intensive methods or machines that were attributed with some form of agency. All of these are valid and help to support their arguments made and it is the context that is important rather than finding an accurate definition.

1.2. Background to the research and rationale

This PhD project received funding from an Engineering and Physical Sciences Research Council (EPSRC) Doctoral Training Partnership award. It caters to the remit of digital economy, especially touching upon the priority areas of trust, identity, privacy & security and content creation & consumption. The original impetus for using an RS was to facilitate community co-production and engagement at Platt Hall in Manchester as the hall had been closed since 2017 for refurbishment and was undergoing an extensive community engagement exercise. This included also thinking about engaging the neighbouring communities and schools with the collections in two ways. First, to engage constituents who would not necessarily visit the museum in person and second, to engage constituents online during closure of the physical site. At the beginning of the PhD, a Participatory Action Research (PAR) workshop was attended and after fruitful discussions and talks with Platt Hall's Senior Curator Liz Mitchell the idea of using a recommender system materialised.

However, as preparation and planning of further workshops with members of the communities got underway, they were brought to a sudden halt by the COVID-19 pandemic that forced the UK into its first lockdown in March 2020, six months after the start of the PhD. Accommodating the new circumstances meant rethinking the project as research interventions in-person were heavily restricted or not feasible at all for most of 2020 and 2021. Driven by the opportunities arising out of the establishment of partnerships with two other institutions, namely the Smithsonian Institution (SI) in the United States of America (USA) and the National Gallery (NG) in London, and the openness and interest of staff to contribute to the research, the thesis' research methodology was adapted to its current form.

As already noted above, RS are not completely novel applications used in museums, but it soon became obvious - through discussions with professionals and a literature review - that neither early design nor actual implementation of the systems included a holistic approach that takes inputs of constituents into account from an early stage and looks beyond the technology, touching upon implications of future use and human perception of it. Studies focused on how well the system works from technical aspects at the cost of inclusion of various constituencies, and systems were mostly evaluated from a technical perspective or lacked varied user input, not just in testing the technologies, but also throughout the design and development stages. There also seemed to be little theoretical research about the perception of AI technologies, especially RS, in museums and a broader discussion about their usage and implications. Preliminary research identified further that there exists a gap in understanding the barriers and implications of using AI technologies in the museum sphere, lacking considerations about the environments such technologies constitute as well as around the data practices they afford.

Data-led activities in museums have been declining in the UK since 2013 (Arts Council England and Nesta, 2020), with museums not keeping pace with the rapid rollout of new technologies or indicating that developed software and products do not cater to the needs of cultural heritage institutions. A development that might risk museums becoming sought-after data providers without gaining any reasonable output for themselves, this can contribute to a widening gap between data-driven technologies and museums. However, AI has the potential to transform museums and the data they are holding, as well as enable computational research to be informed by them as they offer a wealth of data, various user groups to research, and a platform to facilitate interaction whilst educating the public. It is therefore important to take a bigger picture approach to the usefulness and application of AI and data science in museums to identify future symbiotic relationships and barriers.

Although there has been some scholarship of RS used to explore museum visiting (see e.g., Albanese et al., 2011; Kuflik et al., 2011; Pavlidis, 2019; Stock and Zancanaro, 2007), there has been scant research which explores how they might improve and change museum practice and professional workflows. Furthermore, there are few occasions where museum professionals are included in the research at early stages, leading to missing reflection and valuable contributions from those stakeholders. Thus, it is important to collect information about how professional workflows can be informed by and inform the development and usage of those applications to better comprehend their implications in relation to museum data.

In terms of a wider usership, user interaction studies were mainly conducted as small-scale lab experiments or, if tested in-the-wild, attracted a small number of participants often recruited out of one population, e.g., a convenience sample of students. This can cause a lack of generalisability, a failing to address a wider demographic, and not enough data to conduct sound statistical analysis. Large-scale user studies of online museum RS collections are yet to be conducted. In order to test the efficacy and perception of using RS to access and widen engagement with online museum collections, the research developed a novel methodology that combines intrinsic and extrinsic interaction data with in-depth accounts of professionals from an early development stage and onwards. Evaluating such a system is an important means to foster further research and yield valuable insight around user experiences and AI technologies in museum environments.

An aim to address the gaps in scholarly research and threading together input from professional accounts led to this highly interdisciplinary thesis, which is situated at the intersection of HCI and museum studies, forging a framework out of both disciplines to draw a holistic picture.

1.3. Aims and objectives

The above gaps in research drove the desire to establish a wider account of AI, specifically RS, in museums and to investigate how such systems can support professionals and other constituents alike. “AI […] plays a crucial role in culture, increasingly influencing our choices, behaviours, and imaginations” (Manovich, 2019, p. 2). Thus, the thesis further aims to contribute understanding of the wider field of how AI technologies in museums are perceived and the considerations necessary to address the implications of their future use and its broad usership. It is therefore important to examine how constituents work alongside AI to co-create meaning and understanding within public museums. It further tests professional constituents’ reactions to new agents in the museum environment and if they prove to be yet another threat to their “expertise and their social and institutional authority” (Perin, 1992, p. 188).

Through identifying benefits and potential disadvantages to the usage of AI, especially where there are gaps in constituents’ perceptions versus real life application, this research will counter misconceptions that can lead to fear and mistrust (Russell, 2016). The thesis also questions how an RS may inform museological practice regarding bias in collection data and yet unpredictable and currently unknown patterns of misrepresentations or harmful content. This aims to push the museum discourse beyond common boundaries, gathering knowledge with the help of algorithms and creating new connections between objects, their meanings, and their place within online museum collections, especially as museums have often been seen as institutions where “social inequalities have been constituted, reproduced and reinforced” (Sandell, 2005, p. 185). Besides, it gives the constituents the rare possibility to have a discussion about AI technologies and their algorithms, especially in situations where they can “intervene[] in the conceptual foundations of culture [...] and insinuate themselves into long-established routines” (Hallinan and Striphas, 2016, p. 118). This project will significantly contribute to the fields of HCI and digital museology to critically reflect and understand “how these technologies operate to structure the world around them, and in doing so transform humanities knowledge and practice” (Berry and Fagerjord, 2017, p. 104). Furthermore, it explores how AI might help to foster the social mission of museums for the public - away from a “disciplinary museum” (Hooper-Greenhill, 1992) towards a diverse museum that is digitally fit and aware of its social responsibilities. This can establish museums as open and useful places where constituents can gain familiarity with AI, enabling scholars to research interactions, and to provide explanations that “allow people to question and critique a system in order to develop appropriate reliance, rather than blind faith” (Rader et al., 2018, p. 1). Importantly, it aims to bridge the gap between theory and practice in digital heritage research that is prevalent since postmodernism and the early Nineties (Bonacchi and Krzyzanska, 2019). The project therefore aims to consider the practical outcomes and challenges of placing an RS at the heart of museum practice, within digital collections, and responses and perceptions of their many different constituents. The mixed methods selected provide ways to track and identify the different interpretations of success of the interactions between the RS and these constituents in promoting items within collections enhancing their value, and creating new data which can transform professional practice.

To investigate the system and its environment holistically, the following three main research questions have been established:

RQ-1: What is the role and potential use of Recommender Systems in museum settings?

RQ-2: How can RS be used to access, describe, interpret, and enhance existing collections, and meaningfully translate data?

RQ-3: In what ways will the use of RS in museums challenge and/or enhance the public and professionals’ perception of AI?

The above main research questions were divided into further sub-questions and research interventions were drafted accordingly to render a coherent picture of the system itself and the practices around it beyond a mere technical remit of testing the system. The studies and their related sub-questions, including reasons for undertaking them, are explained below.

MAIA survey: Understanding the current AI environment in museums

To understand and situate how museums and the professionals working in them are embracing AI technologies, a questionnaire survey was undertaken during the second half of 2021. The aim of the survey is to investigate AI technologies in the wider museum environment, and it specifically aimed to include accounts of a variety of job roles operating in it, ranging from SLT, curatorial, and marketing to research and restoration, amongst others. Situating an RS in a museum requires an understanding of the current usage of technologies in museums, and so the survey collected information on a broad spectrum of topics related to AI - from current institutional set-ups to future predictions.

To understand the environment a future recommender system will operate in, a survey was conducted to investigate the following three research questions:

RQ-S1: How broad is the uptake of AI in museums?
RQ-S2: What are the current applications and use cases?
RQ-S3: What are the barriers of using AI technologies in museums?

Focus groups: Museum professionals can contribute to the system development and give insights of future needs and implications to consider

The aim of this study is to gather valuable information about how recommender engines can possibly support museum professionals and how this user group perceives such a system, especially considering their workflows. Conducting the focus groups further aims to contextualise the RS and to investigate a shared understanding rather than individual accounts. The focus groups intend to explore the way social and cultural knowledge and meaning are produced around the

system, investigate ethical considerations, and use professional opinions to translate their insight to a wider usership. Further, they aim to identify areas for further development and inform future research and applications for museums.
To include professionals from an early stage, four focus groups were held covering the following research questions:

RQ-F1: Can a recommender web app, based on concatenated vectors from an Autoencoder for images and word2vec for metadata, aid the professional workflows and if so, how?

RQ-F2: What are the requirements to such a system by professionals and what are their considerations?

RQ-F3: How do professionals see other constituents such as visitors, interact with this system?

Online user study: User-centric evaluation of the recommender system

To gather in-the-wild user interaction data and insights, an online user study was conducted to determine if different models of recommender engines are perceived differently in terms of user experience and engagement, as well as investigate if some models are more suitable to serve meaningful suggestions to users of museum online collections when compared to others. With a clear focus on user-centric evaluation, another aim of this study was to better understand how users perceive the RS through gathering intrinsic and extrinsic user feedback; this can inform future research and system development, adding context and insight about the system. It further tests general perceptions of users around technologies and their usage, hoping to elicit opinions around ethical considerations and general usage of the system.

To achieve this, the study addresses the following research questions:

RQ-UXI1: What is the difference in subjective recommendation quality between different models?

RQ-UXI2: Does a system that provides recommendations enhance the user experience of browsing online collections compared to random suggestions?

RQ-UXI3: How satisfied were users with the recommendations?

1.4. Open sourcing

The following code repositories and datasets were established in course of this PhD and were made openly available:

- *MAIA Survey: Code for Data Analysis* (Hughes-Noehrer, 2022)
- *MAIA Survey: Dataset* (Hughes-Noehrer et al., 2022b)
- *Museum Recommender (MuseREC) Web App* (Hughes-Noehrer, 2022c)
- *Museum Recommender (MuseREC). Models and Data* (Hughes-Noehrer, 2022a)
- *Museum Recommender (MuseREC). Data and Analytics* (Hughes-Noehrer, 2022g)
- *Smithsonian SAAM Metadata Cleaner* (Hughes-Noehrer, 2022e)

1.5. Outline of chapters

This introduction outlined the context of the research and the related research questions including the interventions to answer them. It also introduced the theoretical framework of postphenomenology and the aims and objectives of this thesis and the rationale for pursuing this PhD.

Chapter two, the literature review, synthesises scholarly, professional publications, and grey literature which have informed this thesis. The chapter is therefore split into three parts, which broadly reflect postphenomenology's human-technology-world relations. 2.1. establishes the space the RS is situated in (i.e. the world), followed by 2.2., which investigates contemporary constituencies in museums (i.e. humans) and then 2.3. which introduces related works of recommender systems for museums and cultural heritage institutions (i.e. the technologies). The summary of chapter 2 further relates the findings of the review to the research questions.

Chapter three outlines the methodology used to answer this thesis' research questions. It gives a more thorough account of postphenomenology and introduces its main concepts and how it is applied. This is followed by an overview of the partner museums, with a focus on the history, their digital strategies, and operational set-up. After this, the qualitative and quantitative methods used in this thesis are explained.

Chapter four presents the results of the Museum and AI Applications (MAIA) survey, which was chosen as the first research intervention to be introduced as it is the broadest of the three interventions in terms of technologies and establishes an overview of the current AI and museums environment.

Chapter five demonstrates the results of the four focus groups and discusses the research findings in a chapter summary. It is structured according to the three postphenomenological pillars of humans, technologies, and the world - as already used to structure the literature review.

Chapter six presents the online user study, beginning with section 6.1. explaining the apparatus and study design, followed by 6.2. demonstrating the results of the post-study survey and the user interaction metrics, and concludes with 6.3., the chapter summary.

Chapter seven, the main discussion chapter, synthesises the findings according to the postphenomenological framework and spans a trajectory from the formulation of the main research questions to the end of the substantial part of the thesis.

To conclude this thesis, chapter eight incorporates the main findings and presents a future outlook, which includes consideration for further research, suggestions for professionals and constituents interested in the field, as well as lessons learned whilst pursuing this PhD.

2. Literature review

This literature review aims to situate the recommender system in the environment it is operating in through synthesising relevant academic and professional literature in the areas of data, AI, museums, and their constituents. It is written as a review of scholarship on the contemporary practices relating to data-intensive methods for engagement in and management of digital collections, and whilst it acknowledges foundational contributions of digital heritage and museology scholars, it is not intended to be an account of the historical development of either discipline.

It aims to lay a foundation for the thesis by synthesising the current discourse in museum studies and computer science, with a focus on data and AI, its applications, and perception in society.

It identifies and critically reviews publications that investigate the applicability, usability, and evaluation of recommender systems in a museum setting - such as machine learning or data-intensive methods to co-create new knowledge and meaning, spanning a trajectory from visitor engagement and behaviour to audiences' perceptions, understanding and use of digital technologies, to the generated content as experienced through their cultural engagement in art galleries and museums.

The review is therefore structured according to three main strands informing the thesis. Section 2.1. *Museum environments,* introduces the discourse around data, digital technologies and the spaces and materialities they are constituting. Second, 2.2. *Constituents,* gives an overview of the contemporary constituent landscape exploring possible userships and their implications for the recommender. Lastly 2.3. *Systems and evaluation,* discusses recommender system applications for museum and heritage sites and their forms of evaluation.

2.1. Museum environments

This section gives an overview about the environment recommender systems are operating in by summarising key points around data-intensive methods and cultural practices.

AI technologies for museums have seen broader uptake in recent years. French and Villaespesa (2019) identify opportunities for museums to use AI not just to enhance visitor experiences or enable new ways of exploring collection data, but also to gather valuable insights about the visitors themselves. Merritt (2017) identifies AI as an essential tool for museums to cope with the ever-growing amount of data in the 21st century and a means of making collections more accessible and usable for the public. The interplay of AI and museums is a young field and most possibilities have not yet been explored but promise powerful and new ways to investigate collections, objects and their creators. However, enabling the power of AI still requires significant resources, tools, time, and expertise (Ciecko, 2020).

Investigating how we work together with machine-learning infrastructures, how this interplay effects the creation of new knowledge and meaning, and how we engage with them, also provides a challenge to reform established institutional structures (Bassett et al., 2017) in a world where "authority is increasingly expressed algorithmically" (Pasquale, 2016, p. 8).

Data in general is the prerequisite for AI technologies to serve successful recommendations to users, helping them to find information that fulfils criteria such as being novel, interesting, diverse, relevant, and meaningful (Aggarwal, 2016; Pavlidis, 2019). The digitisation of museum collections has created vast quantities of data, providing the basis for both computer scientists and museum professionals to design data-intensive and data-led models (Giannini and Bowen, 2019). Harnessing the power of computational methods provides the means to create, process, modify, transfer, visualise, and store information in machine-readable formats. All of these tasks involve data.

The database therefore has become a predominant and key form of cultural expression of the modern age (Manovich, 1999) leading to a “cultural reconceptualization” that is dominated by the “computer’s ontology, epistemology, and pragmatics” (Manovich, 2002, p. 47). Striphas observes an “enfolding of human thought, conduct, organization and expression into the logic of big data and large-scale computation, a move that alters how the category *culture* has long been practised, experienced and understood” (Striphas, 2015, p. 396). Together with Hallinan, he formalises these thoughts further, giving a specific outlook on the influence of recommender systems and how they possibly lead to a reinterpretation of culture and a new understanding of it shaped by algorithms (Hallinan and Striphas, 2016). Those now “algorithmic cultures” have taken over a major part of cultural production through “human beings delegating the work of culture [...] to data-intensive computational processes” (Striphas, 2015, p. 396), revealing, first and foremost Western culture as a “project that seeks to transform itself into an apparatus” (Flusser, 2013, p. 9).

Kitchin (2014) defines apparatuses as, amongst others, practices, organisations, places or forms of knowledge, that form complex data assemblages through their various elements, e.g., institutions, curators or public and political opinions, situating them in a web of relations determined and co-shaped by data. He also identifies data as fundamental for today’s knowledge production, highlighting that data is neither naturally neutral nor static, but strongly influencing and influenced by the actors and communities of practice using and producing data (Kitchin, 2014). Museums find themselves amidst this “data revolution” that radically transforms the way we leverage insight and value (Kitchin, 2014). With the advent of big data, institutions were subjected to the power of metrics (Beer, 2016) and were driven to capture data and evaluate cultural performance through it (Arvanitis et al., 2016). It is evident that this revolution and the outlook of a “digital democracy” (Jenkins, 2006, p. 208) are continuously but slowly emerging, characterised by difficulties in finding a common approach across the sector and unevenly spread resources and capacities.

The COVID-19 pandemic had a profound effect on institutions and their data practices, evidently speeding up digital data practices and strategies with several implications for this thesis, too. Digital data practices refer to the ways digital data are used, perceived, and handled within and by museums and their constituencies using digital technologies and computational methods.

To continue operating in the context of multiple lockdowns and ongoing restrictions, museums were forced to change and/or accelerate their digital practices and processes, and many institutions, and their audiences, were thrown into the “digital deep end” (Finnis and Kennedy, 2020, p. 11), making the internet the default form of engagement. There was a significant rise in online content production, and republishing and repackaging content (Finnis and Kennedy, 2020), shifts which in turn necessitated changes in internal processes, such as data cleaning, cataloguing, or getting collections online (Art Fund, 2020). Those institutions with digital strategies in place reported a smoother transition than those who had previously struggled to incorporate digital products into their operations or keep pace with rapid technological developments before COVID-19 (Finnis and Kennedy, 2020; Merritt, 2021) particularly, smaller organisations with fewer resources and capacities (Travkina and Sacco, 2020). The pre-existence of organisational digital culture and capacity (Newman et al., 2020) enabled museums to bounce back faster post-COVID with “strategic foresight” (Merritt, 2021, p. 3) through the ability to “speak machine” (Maeda, 2020).

Evidently, data-intensive techniques and ever-increasing computational powers are affecting museums, their practices and audiences and what used to be a relatively slow but steady process substantially gained traction over the last three years, driven by a global pandemic. This trend is traceable back to the paradigm shift towards a “New Museology” (Vergo, 1989) and in foundational publications of digital museology and heritage scholars (see, amongst others, M. L. Anderson, 1999; Arvanitis, 2004; Bowen et al., 1998; Cameron and Kenderdine, 2010; Parry, 2005, 2007, 2013; Parry and Sawyer, 2005). This further highlights that the implementation of a museum recommender system is not a purely technical task, but has a far-reaching, often reflexive impact on the various apparatuses and the assemblages evolving out of it. The use of data leads to a reshaping of culture and new ways of knowledge production and access to it, but also how performance is measured and justified - the computer has established a new space of producing and consuming culture. The next paragraphs explore this space and the various notions of how culture and knowledge are produced within it.

New radical approaches driven by technologies offer new ways of displaying not just objects, but also ideas, enabling new forms of how knowledge is understood (Hooper-Greenhill, 1992) and can further help museums to create new knowledge, around objects and beyond, supporting a “meaning making that engages affordances unique to data” (Padilla, 2017, p. 1). Data does not simply allow access to information, but “allows users to find new meanings” (Kenderdine, 2016, p. 24). Digital collections - digitised or born-digital - are now supporting users to create their own contexts and experience several dimensions of information that can change over time.

These web-based collections establish a new information space that reshapes the institution as the Internet gives access to collections anywhere, anytime able to break down and permeate the museums’ brick walls (Giannini and Bowen, 2019; Navarrete, 2013). Without boundaries between the physical and the virtual - formerly being treated as distinct realms of real *or* virtual, authentic *or* faux (Parry, 2007) - museums have arrived in a post-digital space (Berry, 2018), where the digital and the non-digital are not anymore distinct from each other. This was further reinforced by the COVID-19 pandemic and the sharp pivot towards digital technologies to facilitate the online presence of museums (Finnis and Kennedy, 2020) and provide for the increased consumption of cultural content online (Creative Industries Policy and Evidence Centre, 2020). There is hope that this has a long-lasting impact, as the notion of the future after the pandemic is drawn as one noticeably different to the time before, requiring museums to embrace change and adapt to a model where “the physical space of the museum is no longer dominant” (Art Fund, 2021).

From the early 2000s on, coinciding with the rise of the Web 2.0 (O'Reilly, 2005) and more participatory forms of online engagement, institutions had started to appraise the digital as an opportunity and a tool providing a space to explore multi-layered realities and multifaceted narratives (Parry, 2007). Museums are "hybrid spaces" that are "exciting but intensely challenging makers of meaning and facilitators of experience" as digital media and their modes of engagement raise not just questions around how the digital has changed the visitor and their participation and forms of collaboration, but first and foremost questions about "voice, ownership, data and - perhaps most crucially - power" (J. Kidd, 2019, p. 193).

Abolishing the distinction between analogue and digital further raises questions around the materialities forming in and inhabiting these hybrid spaces. Blanchette remarks that "however immaterial it might appear, information cannot exist outside of given instantiations in material form" (Blanchette, 2011, p. 1042), referring to the very core of ICT and, even if digital technologies might be perceived as something intangible, their very functioning relies on hardware and material instances.

There has been a shift in thinking about the materiality of digital media and technologies in museums, as what was formerly seen as excluded of a material culture and a dominant object-centred practice, is now very much at the centre of it. Galani and Kidd call this a shift to "hybrid materialities" as "people assemble the materiality of their heritage encounters through a range of digital, analogue, tangible and intangible resources, their visiting experiences transcend traditional articulations of the physical-digital divide and operate on a continuum of materialities" (Galani and Kidd, 2020, p. 299).

Experiences online form a part of this continuum and museums risk falling behind if these materialities are not considered carefully or are disregarded, potentially leading to a loss of audiences, credibility, and relevance.

Through the establishment of these hybridities, the physical site of the museum has, to some extent, lost its centrality and has been displaced as the focus for decision-making. This shift of power and broadened participation driven by new media enabled heightened peer-to-peer communication rather than one-to-many broadcasts (Kenkins, 2006) and therefore offers a way to circumvent traditional ways of museum communication. This thesis aims to holistically address these hybrid spaces and materialities by giving equal attention to what formerly used to be reserved to solid objects situated in physical locations as opposed to what was deemed the immaterial, digital or virtual. This allows consideration of cultural practices that happen without going through any process of institutionalisation at all, leading to a reattribution of authority and agency through networked digital media that might be more complex, but allow for greater freedom in regard to production, dissemination and consumption (Paul, 2006). Institutionalisation is here understood as institutional processes in and around museums that can lead to the exclusion of certain groups based on political, economic and social factors (Sandell, 1998), which have changed with the arrival of digital media.

Kidd sums up the changing mediascape in six important points which shape questions around recommender systems in a museum sphere:

> *"Firstly, that how we experience a text or cultural artefact is being remodelled in (for example) the rise of 3D technologies, wearable technologies, immersive experiences and non-linear narratives, and that such media threaten our reliance on and ability to consume 'wholes' – whole exhibitions, whole collections, whole narratives. Secondly, that our understanding of identity and community is changing. [...] Thirdly, that the ways we create, distribute, access and assess information are changing, with new ways of managing knowledge creation and information sharing [...]. Fourthly, that the way we organise ourselves, our communities, our politics and our culture is shifting. [...] Fifthly, as has already been noted, the roles of producer, consumer and distributor are changing; the value chain of culture is being fundamentally reworked as various practices become de-institutionalised and dis-intermediated. Sixthly, one of the more problematic claims of the new media is that they democratise access, participation and the right to representation. The ramifications of such a shift could be paradigm-shifting for institutions like museums, but*

> *we must be careful not to overplay their significance at this time.”* (Kidd, 2016, p. 5)

These key points inform the thesis by establishing guidelines for critically evaluating the practice-research, framing understanding of the potential and limits for the RS and its application and helping to interpret and situate the responses of research participants.

Digital technologies have been identified as being useful for museum users (Bearman, 2008) and the ever-increasing use of both digital technology at large, and computing which is becoming ever more pervasive and ubiquitous, is contributing to museums becoming places where culture and technology meet, are mediated through public reception, and are an ideal environment for exploring new concepts in intelligent user research (Stock et al., 2007). Nonetheless, most of the works concerning such systems have been developed to investigate technical aspects. This can lead to hermetic circles of technical experts, who develop and evaluate such systems, excluding the wider sector and the topics so inherently connected to museums, evoking the feeling that the role of stakeholders was not taken into account or was sometimes even mistreated, failing to address organisational or audience needs (Devine, 2015; Pavlidis, 2019) leading to “values and prerogatives that the encoded rules enact are hidden within black boxes” (Pasquale, 2016, p. 5). In turn, this can lead to the development of recommender systems that are focused on either algorithmic efficiency or user satisfaction, whereas it would be necessary to take a more holistic approach, that takes various contexts into consideration and not just parts of the sum (Naudet and Deladiennée, 2019).

Institutions are places of interaction between individuals and nonhuman entities, they already provide immersive experiences which educate and engage the public, and data-driven research provides the means through which public behaviour and response to them can be better understood in relation to its benefit to society (Gilmore et al., 2018). Using a recommender system in a museum setting whilst being beneficial to society means that museums are prompted to address their “history of elitism and exclusion” (Taylor, 2020, p. 7) and issues such as bias, racism, and inequalities that might be prevalent in their collections. This, on the other hand, is also true for the AI community that is slowly picking up pace in considering ethical and fair applications; a practice that led to generalisations, misuse of terminology, and decisions that cause a downstream harm (Mohamed et al., 2020; Suresh and Guttag, 2020), or data being considered ethical just based on their availability (boyd and Crawford, 2012). Data use for museum recommender systems means that museums will therefore have to address the troublesome and challenging parts of their collection (S. Anderson, 2020) and standards of development and data tracing are highly recommended (Gebru et al., 2020). Developing systems that are accessible, inclusive, and openly address issues around bias, privacy, and error is important and sets expectations right from the beginning (Morris, 2020).

It should be further acknowledged that neither museums nor technologies are neutral, they do have agency in our lives and it would be both “undesirable and unhelpful to exclude them [those agents] from a moral discourse” (Coeckelbergh, 2009, p. 181). Thus, it is inevitable to establish an ethical discourse around AI usage in society (Floridi et al., 2018) and particularly around RS as “they shape user preferences and guide choices, both individually and socially. This impact is significant and deserves ethical scrutiny, not least because RS can also be deployed in contexts that are morally loaded” (Milano et al., 2020, p. 957).

Having one all-encompassing space can lead to the establishing of more democratic ways of engagement with museums as the production of culture, whereby access to it and forms of dissemination are not solely channelled at the institutions anymore. However, whilst this can be paradigm-shifting, the notion of digital technologies enabling more democratic ways of engagement and participation need to be caveated, as already pointed out above by Kidd (2016). Situating a recommender system in this space thus means exploring it without situating the focal point of power at the museum *apriori*, but to conceive this space as one with a shared authority between institutions and the constituents who inhabit the very same space. This means for this thesis to understand who these constituents of this space are and consider them as equally important co-producers of the recommender system. The next section therefore explores the literature around contemporary audiences and inhabitants of these now hybrid spaces.

2.2. Constituents

The former section introduced an overview about how data is shaping the institutional structures of museums and the contemporary environments they are operating in. This section will now explore how digital technologies are shaping contemporary constituents of museums in regard to AI technologies and consider what possible userships of the recommender system might look like. Constituents are here understood as an umbrella term to address the different ways the terms such as user, visitor, or audience are used in a museum context as well as cater to the language used by the research partners.

The Audience Agency (2020) reports in their Museums' Digital Visitors Report that most users of museum online provisions visit to learn and be intellectually stimulated, followed by seeking entertainment; they tend to be older, identify as disabled more often compared to other arts and culture sectors, and about 80% of first-time web visitors have never visited the institution in person, with 37% of high-frequency visitors of museum websites stating that they have never visited in person either. Museums are encountering visitors occupying the digital space who are different to their analogue peers (Battro, 2010; Bearman, 2008) and do not fit traditional frameworks of audience segmentation, but are identified "as key agents in the production of digitally mediated material encounters" (Galani and Kidd, 2020, p. 300).

Introducing digital media can make institutions more open and reactive to the needs of their audiences, bring them together and establish practices that address globalisation and multiculturalism (Witcomb, 2007), giving visitors "greater opportunities to engage with museum collections; they also increase the museum professional's ability to document and track interests and needs of museum visitors" (Marty, 2008, p. 135).

To understand the visitor's digital experience and journey, Devine highlights that visitors do not distinguish between digital experiences and physical museum visits as they tend to experience the museum as *one entity*, and the "visitor's journey with the Museum transitions between channels, as they move from website to app to social to website. The visitor's journey exists on many devices [...] and their expectations are influenced by standards that exist in the general world, not just what is the norm in the Museum world; [a concept known as] 'omni-channel thinking'" (Devine, 2015). Remote visits are not inferior to their physical equivalents and "developments in ubiquitous technologies and telecommunications encourage us to think of digital technologies not only as information tools, but also as experiential processes that fit with one's everyday life and interactions" (Galani and Chalmers, 2008, p. 158), resulting in "interactives so transparent that visitors may not be aware of a separate physical museum identity, or that certain physical barriers between artifact and access even exist" (Marty, 2008a, p. 132). Evolving networks and distributed computing enable people to join the cultural discourse and explore frameworks of human understanding (Benkler, 2006), leading to users perceiving "the networked computer as their natural environment, and thus as their main context for any kind of experience" (Quaranta, 2018, p. 55). This means that the computer is not a mere means to reach a goal, but the computer with its networks has become an institution itself. This development is traceable back to early Netart movements in the 1990s and now clearly visible with social media applications of the 21st century, where cultural production does not just happen digitally, but is also meant to be consumed digitally.

Simon (2010) remarks that visitors further need entry points that are familiar to them, such as social media platforms, and offer levels of personalisation and content that resonates with the audience rather than with the museum professional; an audience-centred approach (Simon, 2010; Villaespesa and

Stack, 2015) that caters to the shift to a more user-centric evaluation of recommender systems too (Knijnenburg and Willemsen, 2015).

Visitors to museum online collections are not submissible to a one offer fits all approach, they are a “new public” that “[museum professionals] need to take good care of” (Battro, 2010, p. 145) and can be segmented, e.g., by their different needs in their information seeking behaviour (Skov & Ingwersen, 2008) or their different reasons for visiting museum online provisions and their domain expertise (Villaespesa, 2019; Villaespesa and Stack, 2015).

Digital visitors can further differ from traditional bricks and mortar attendees (J. Falk, 2016; Walsh et al., 2020) and museum professionals, if they want to deliver successful content, need to understand how they can “offer information resources uniquely designed for the needs, attitudes, and expectations of online visitors making digital museums part of their everyday lives” (Marty, 2008b, p. 97). Museum professionals understanding digital visitors is one thing, as system development also needs to include museum professionals not just in their function as content providers, but as users of the systems themselves (Gilliland-Swetland and White, 2005). Designing, implementing, and evaluating a successful sociotechnical system therefore requires collaborative and continuous inputs from both professionals and users (Carayon, 2006).

Catering to various userships means that institutions need to incorporate appropriate principles of collection documentation, being aware of whom they should cater to, guided by the principles of (i) a polysemic role of objects, (ii) an acknowledgement of the meaning of narratives and classifications systems as products of cultural, disciplinary, museum, and curatorial opinion, and (iii) that the current knowledge context needs to acknowledge the role of users in the cycle of knowledge making (Cameron, 2010, p. 86). Cameron further identifies that, although certain audiences will still demand an authoritative scholarly interpretation, opening up collection data to post-structuralist and postmodernist ways of interpretations can foster a “more open and inclusive approach [that] will give greater power to the user to create their own knowledge pathways and to make and “put up” their own interpretations in a kind of shared authorship” (Cameron, 2010, p. 90). This is a demand that was not just issued in regards to documentation of collections, but in general for data management plans (Fresa et al., 2015) and accessibility and useability (Hansen, 2019; Pisoni et al., 2021) as digital museum media can help to facilitate access for constituents with disabilities, special needs and impairments, leading to improved social inclusion of marginalised groups (Economou, 2008). From an audience development perspective, the shift to digital-only participation during COVID-19 brought the potential for new and diversified audiences to encounter museums. This is supported by UK and US surveys (Samaroudi et al., 2020; the audience agency, 2020b), which saw a rise in engagement with less traditional and more vulnerable audiences. However, not all audiences have equal access to digital technologies and there is evidence that COVID-19 has contributed to the digital divide (Holmes and Burgess, 2020), exacerbating inequalities in a society ever more reliant on data infrastructures (Baker et al., 2020). The presentation of museum activities via digital means has been found to provide other societal benefits and public services, opening up access to new resources for research and education (Agostino et al., 2020; Samaroudi et al., 2020) and providing activities which promote wellbeing and combat anxiety, mental health issues, and loneliness (Creative Industries Policy and Evidence Centre, 2020).

All of this further contributes towards the conundrum of narrowing, but at the same time widening, of the digital divide and the “data divide” (Ada Lovelace Institute, 2021) and introduces new complex challenges around ethics, ensuring participants’ welfare, and the need for awareness that “entries into the participatory media space are not inconsequential, and that the tech and platforms that underpin their practices are not neutral” (Kidd, 2019, p. 202). Museums nowadays operate in an environment where power is more and more expressed algorithmically and these algorithms have far-reaching and profound social implications (Beer, 2009).

The abundance of content held by museum online collections can lead to audiences struggling to find meaningful ways of engaging with them, possibly resulting in high abandonment rates, short website visits, and few page views (Villaespesa, 2019; Walsh et al., 2020). The usual access point to online collections still is the search box which leads to several limitations of information retrieval (Speakman et al., 2018; Stack, 2018), providing interfaces that fall short of giving access to the abundance of information held by institutions (Whitelaw, 2015). Issues which prevent access include entries not tagged with words known to a broad usership (e.g., botanical collections using scientific names) (Bearman, 2008), systems being designed for (or replicating) domain expert users (Skov and Ingwersen, 2008) or too many different systems, not-standardised terminologies and a lack of description or images (Beaudoin, 2020). Systems need to be designed to be "compatible with the user's existing browsing logic and user experience" (Pruulmann-Vengerfeldt and Aljas, 2014, p. 172), a personalised web access to museums can help to tailor content to users and facilitate search; this requires understanding of who the users are and what they want through gathering implicit or explicit feedback, serve the personalised content and evaluate the recommendations (Pechenizkiy and Calders, 2007).

Personalisation of content is one form of how interactives based on museum online resources can help to tailor content to each user's needs and preferences, and can therefore create unique experiences and acts of co-creation (Darzentas et al., 2022; Marty, 2008b). The possibility to adapt content and integrate visitor's input was identified to "offer a more active type of learning or general visitor experience compared to more traditional means of interpretation" (Economou, 2008, p. 149) and a "tremendously personal [...] and enjoyable learning experience" (Rodney and Stein, 2020, p. 34). New media and the web contribute to a "personal museum" (Parry, 2007, p. 109), where forms of personalisation shift authority and authorship away from the institution towards an audience that is "increasingly self-directing and self-managing" (Bennett, 1998, as cited in Parry, 2007, p. 109). Sharing and collecting personally relevant and meaningful objects can further provide "an emotional and aesthetic counterpoint to authoritative interpretations in museum exhibitions" (Pierroux, 2019, p. 134).

Offering customisation therefore is not just about offering a choice, but about gaining value and personal meaning for museums and their audiences (Munley et al., 2007). However, the option of personalising content in museums might, on the downside, risk a polarisation of how museums are viewed by visitors (Rodney and Stein, 2020) and possibly contribute to visitors building their own walls "draw[ing] upon their likes and dislikes to create their own personalized set of museum artifacts place artificial restrictions on visitors" (Marty, 2008a, p. 132) which can contribute to the establishment of new barriers with visitors just focussing on their own particular needs and interests, losing sight of the bigger picture.

This section shows that the audience landscape has changed over-time and that traditional forms of visitor research and audience segmentation are not applicable to not just a digital audience anymore, but more generally to all forms of constituents as engagement happens streamlined with users often making no difference between various analogue and digital forms of interaction. This research suggests that hybrid spaces can further lead to a shared authority of interpretation and a more active and rewarding experience through digital technologies, as participation can be personalised and meaningful, opening up forms of accessing museums to groups that either have not been attracted by traditional museum offers or were not capable to engage with such offers due to disabilities or other barriers and thresholds. Reviewing the literature, personalisation is rendered as a rewarding form of engaging constituents, giving access to collections in a possibly more enriching experience. However, as with other forms of participation and engagement there are potential downfalls and risks if not evaluated holistically in terms of siloing information or reinforcing unwanted power structures within the collection data.

2.3. Museum Recommender Systems and Evaluation

This section of the literature review introduces museum and cultural heritage recommender systems relevant to this PhD. The focus lies on the papers of importance for this thesis and will therefore have proportionally more space attributed to them. These include web-based recommender engines and models that are similar to the ones described in this thesis.

Precursors of museum recommender systems were early forms of content personalisation coinciding with the advent of Web 2.0 and a more participatory form of the Internet and the availability of portable devices to enhance the user experience during the museum visit, such as *Sotto Voce* (Aoki et al., 2002) and *ARCHIE* (Luyten et al., 2006). Both of these projects are focused on enhancing the on-site experience, rendering museum visits as something physical and visits are seen as social experiences, mainly for groups.

These early projects, were mainly proof of concepts or, if tested with small groups, were soon followed by bigger projects, such as *PEACH: intelligent interfaces for museum visits* (Stock et al., 2007; Stock and Zancanaro, 2007), which researches intelligent interfaces through a multimedia guide, including mobile and stationary devices, cinematographic techniques and an extensive usability study and evaluation. *PEACH* has laid one of the cornerstones in the field of RS for cultural heritage institutions and was later extended by Kuflik et al. (2011) for the Hecht Museum based at the University of Haifa in Israel, aiming to provide visitors with personalised information during their visits. Whereas both of the above approaches seem to be outdated nowadays (e.g., development of an interface to exchange messages during visits instead of using smartphones), they both contributed substantially to methodologies to prepare museum data for multimedia applications and have shown that the use of a device during visits has significantly increased the time spent at the galleries and can contribute to new forms of visitor experiences.

Recommender systems are further noted as tools to establish social connections (Perugini et al., 2004) through the establishment of relations between users through their implicit or explicit feedback, identifying inherently social components of such systems and therefore the necessity for more HCI-focussed research and assessment, an upward trend from the early 2000s on.

The *CHIP-Project* (Wang et al., 2007, 2008) takes a new approach using an RDF/OWL semantically-driven tour guide at the Rijksmuseum in Amsterdam, serving content-based recommendations in the form of artefacts and art historical topics, described as the *ArtRecommender* throughout their papers, and subsequently delivers recommendations for tours at the museum through the *Museum Tour Wizard*. It specifically looks at personalising access to museum online collections, however, the models to serve recommendations are based on different approaches than the ones described in this thesis as *CHIP* focuses on connecting professional data (Semantic Web) to user data mainly sourced from social networks. In aiming to "bridge the gap between physical and virtual museum experiences" (Stash, 2010) and Wang et al. (2007) identify that the demands to such a system are different for each of the envisioned user groups, such as experts and novices, making the case for this thesis' research to test domain knowledge ahead of the user study and to also consider levels of expertise developing and testing the system. Although the two conducted user studies were at a smaller scale compared to the user study presented in this thesis, their outcome presents valuable considerations for future system development and shows promising results in serving recommendations based on semantic similarity, which, although different, supports the thesis' approach in using word2vec. Unfortunately, the project's website is not available anymore, making it difficult to track any further outcomes or identify reasons for the project's abandonment.

Another project using a content-based recommender based on semantic similarities is the *ITem Recommender (ITR)* (Basile et al., 2008) which aims to engage users by letting them annotate artworks in the form of tags, also known as *folksonomies* (classification systems created by end users), and therefore create semantic-based recommendation profiles that infer user interests not just from the static (e.g., artwork description by curators), but also dynamic UGC. This system is also web-based and uses 45 paintings out of the Vatican picture-gallery's website and was tested with 30 participants. This prompted the authors to suggest further investigation of collaborative annotation with folksonomies for the future as the recommendations containing static and dynamic tags were proven to outperform models using either/or. Semeraro et al. (2012) introduce a more recent system founded on the same working principles of integrating folksonomies into a content-based RS when presenting *FIRSt* (a *F*olksonomy-based *I*tem *R*ecommender sy*St*em). *FIRSt* references WordNet ontologies to establish semantic relationships between words and implementation of it returned more promising results than a pure string match search as this was shown to sometimes retrieve unwanted or wrong results due to the polysemy and synonymy of words (Semeraro et al., 2012). *FIRSt* uses the same evaluation dataset as used for *ITR* and the results are congruent with the ones of Basile et al. in confirming that dynamic and static approaches are combinations that render promising results in their evaluation.

Albanese et al. (2013) develop a very general recommender system with the goal being to browse and retrieve information from multimedia databases, using a social choice problem theory, intrinsic multimedia features, and user behaviours to serve recommendations. The system was trialled with paintings of the Gallerie degli Uffizi in Florence, aiming to serve recommendations to visitors of a 'virtual Uffizi Gallery' by bringing semantics and personalisation approaches together. Whilst implementation was successful, the system was trialled with just 474 paintings and tested by 30 users, which makes it hard to infer wider ecological validity due to very low numbers for both the dataset and the participants. However, Albanese et al. highlight an important issue concerning recommendations based on importance rankings - a part of their model is based on the famous *PageRank* algorithm, which is used by Google to rank the importance of websites based on their links - and that such approaches might not be replicable in a museum multimedia setting as the users' browsing behaviour might defy traditional browsing logics of pure importance rankings without consideration of semantics or other features. The conducted experiments further show that experts can sometimes find suggestions or recommendations as mentally more demanding as it interferes with their browsing logic; this is based on the assumption that domain experts generally know what they are looking for as opposed to novices or non-experts.
A whole *SMARTMUSEUM*, with its foundation built on an RS for the Web of Data, was constructed by Ruotsalo et al. (2013) with the aim to serve site-specific semantic and context-based suggestions. The system is designed as a mobile application that serves recommendations not just in a museum, but also other structures of interest, for example relevant buildings, statues or heritage sites.

Using ontologies, the *SMARTMUSEUM* presents the user with descriptive information about the chosen objects (some choices are RFID-enabled) and related multimedia files. Being implemented for the Web of Data, the recommender profits from semantic technologies, such as RDF integration, with content, location, and context stored as triples. The RS was tested in lab experiments and users reported positive feedback, making the case for an application that successfully incorporates Semantic Web frameworks. However, as with other surveyed projects, sample sizes were small, in this case just 24 participants took part making generalisations hard. Reviewing this project for the thesis, it is further evident that context can play a substantial role in serving recommendations and that users might have different demands when using the recommender, for example educational or recreational purposes and if they are accompanied or not. Another outcome in regard to interface design is that users preferred to have a selection of recommendations rather than an arbitrary list, as this might lead to information overflow and possibly a feeling of too much choice which is the opposite to a working recommender engine.

Keller and Viennet (2015) present an evaluation of several recommendation models as part of their smart audio-guide for the *AMMICO* project, that provides visitors with graphs to recommended artworks in order to break the often linear curatorial narrative and engage the visitor in a more personalised way. The audio-guide mainly focuses on the user’s behaviour and takes into account what is ‘liked’ during a visit. The data feeds into a generic formalism developed by Viennet and Keller, *Social Filtering (SF)*, which enables testing of several recommender strategies. Results show certain trade-offs when comparing the different methods. However, the used SF IB (Social Filtering Item-Based) approach shows an overall better performance compared to traditional methods, such as Bigraphs or Popularity. Social Filtering therefore can serve as an implicit tool to reflect social interactions amongst visitors with objects as the two graphs (*visitor-graph and PoI-graph*) either matching one PoI (*object*) that was liked by at least two visitors or vice versa. For in-house visits this could pose several issues, for example the number of shared preferences between visitors is too low to make accurate predictions, or the number of objects on display is insufficient and the visitor would have been engaged with them anyway as they had to indicate their liking beforehand, or would have passed them anyway through following the suggested movement through the exhibition or spatial constraints. Evaluation of the methods was performed through offline experiments on a dataset that was collected using real life visitor interaction with PoIs, but authors remark that further testing needs to be done in terms of user satisfaction, suggesting future research, such as presented in this thesis, should include how recommendations are perceived by museum visitors and if they are accepted or not, agreeing that offline testing might not fully account for psychological effects. Further, this study is limited to actual exhibits in museums, and whilst this small scale approach might render interesting results in terms of accuracy and prediction, this might not be the case when tested on large collection datasets.

Kislyuk et al. (2015) introduce a hybrid recommender system relying on content-based ranking and collaborative user curation data to serve item-item recommendations on *Pinterest,* an online platform to curate *Pins. Pins* are images mostly containing some form of annotations. The authors' system therefore combines content-recommendations with visual feature representations derived from Convolutional Neural Networks (CNN). Even though the architecture is not specifically built for a museum environment, the model contains methods that could be usefully replicated on the Web with museum collection data. Users curate boards on Pinterest where they can pin (similar to bookmark) content onto a virtual board. These pins can contain image data and metadata, and whilst it is possible to just save an image to the board, most users annotate pins with short descriptions. As various users can pin the same content, the pins get enriched with more and more annotations, and as this content is curated by the user's interest, probabilities are significant that these users will like pins on someone else's board, therefore *'aggregated image co-occurence'* could be a handy tool for the exploration of heritage data. Whilst the use of annotations reminds of the earlier introduced folksonomies, Kislyuk et al.'s approach adds more complexity to textual features and introduces a combination of them with image similarity features. Using CNN together with collaborative methods rendered positive results in terms of engagement and accuracy of recommendations and makes a promising case to consider Deep Learning methods for this thesis and the inclusion of image-based features.

Elahi et al. (2017) bridge the gap between low-level and high-level semantics in recommendations for movies. Whereas Elahi et al.'s research was not undertaken to be applied by museums, it highlights a rarely respected problem, or assumption, in recommending cultural content and these research findings have implications not just for movie audiences but also for museum collections. Elahi et al. assume that consumers of movies are not just interested in high-level features, such as actors, genre, certain eras, producers and directors etc., but also in low-level features (also known as mise-en-scène features), such as lighting, sound, motion, brightness, colour etc. Were these to be remodelled for a museum collection, one could assume that some visitors are not just interested in artworks according to high level semantics, e.g., certain artists, genres or depictions, but also aesthetically-stylistically motivators such as colour, brightness, size and texture. Their user study shows significantly better results throughout by combining high with low-level features, whereas the use of low-level features alone marginally outperforms high-level features in terms of user satisfaction ratings. The use of an Autoencoder might cater to this fact as the features after a dimensionality reduction might be what can be considered as low-level.

A project developed and tested online in a museum setting is the *CrossCult App* developed by Kontiza et al. (2018) for the National Gallery in London. The app works by pre-assessing the visitor through a couple of questions to generate a profile. In accordance to this profile, the visitor is then presented with a stack of images where they can indicate their liking or disliking by swiping left or right, a mechanism adopted from a contemporary dating app. After having swiped through the stack of artworks, the visitor is then served with grouped recommendations of art works purely based on the artworks' metadata associated with it. The app is further enhanced by a map visibility feature to show the actual location of the artworks. With this approach, Kontiza et al. hope to stimulate participants to reflect and engage in a deeper sense with art, which received positive feedback from 35 students who piloted the app imagining a visit to the National Gallery; just 4 took part in follow-up interviews and an actual on-site visit. Whereas there is room for improvement in terms of evaluation and sample size, the app shows positive results in terms of user engagement. Unfortunately, the paper does not include actual model architecture and the algorithms that are used to serve recommendations.

More recent studies present immersive recommender systems in a cultural environment, not just because of developments in computing power and the availability of easily scalable cloud services at affordable costs, but primarily due to an emerging shift from visitors merely being entertained towards personal, more meaningful and knowledgeable visits. David and Kamerling (2019) use a knowledge-based RS to model more complex art historical information, for example iconographic descriptions, to facilitate access to collections without having domain-specific knowledge. The system maps two taxonomies, *ICONCLASS* and *AAT*, to each other via the interlinking of labels as a starting point to calculate relevancy scores of the matched content. In this specific case, *ICONCLASS* concepts - being a high-level classification scheme aimed at a scientific usership - are mapped to the broader and easier to understand taxonomies of *AAT* in a vertical hierarchy. To test the system, the authors used the *Rijksmuseum* Linked Open Data (LOD) collection of objects, which is already indexed in *ICONCLASS*. Evaluation of the web-based prototype returned some false-positive interlinks, but overall valid matches outweighed. Translating higher-level to lower-level concepts can be a promising way to facilitate access to collections for visitors to museums with non-expert or no knowledge about iconography or other domain-relevant fields. A further possible development, although very time-intensive to realise, would be the mapping of expert descriptions to user tags created by visitors. Loboda et al. (2019) developed a content-based recommender system for UCL's Grant Museum that focuses on the visitor experience, offering an app to enhance the on-site museum experience through recommending personalised tours. The system was tested on-site in a pilot with twelve participants recruited from staff and students and it highlights the advantages of early real-world testing to further develop systems that have a clear UX focus. Loboda et al. further conducted “front-end field evaluations” to initially assess visitor needs at an early design stage.

A different form of visitor experience, spanning across several cultural heritage sites in a contemporary approach, is presented in the following two papers. Koukoulis et al. (2019) present a system for travelling to points of cultural interest via a travel guide linked to a museum app. The authors’ aim is to extend recommendations to the outside space surrounding the museum to other PoIs that may be of interest to visitors, complementing their journey. This system runs on a hybrid model, serving recommendations based after the user enters some personal details, collaborative data, and contextual information in the form of constraints (i.e. time, mobility) and object relevance. Whilst this system is not focused on state-of-the-art RS, the authors are more concerned with usefully connecting the museum and its surroundings. However, evaluation shows mostly better performance than other comparable systems, i.e. the *SMARTMUSEUM*.

The presented work introduced by Su et al. (2019) provides one of the first investigations into how Big Data and Edge Intelligence technologies can help to explore new ways through recommender systems in the cultural heritage sector. The system introduces a big data architecture for heterogeneous data sets, accessible via cloud services or an edge-intelligent *Smart Search Museum* mobile app serving holistic recommendations for the museum space and beyond. The RS uses a user-centred model, based on relevant related work, such as the before discussed RS by Albanese et al. (2013). The hybrid model pre-filters users according to social network analysis and their location to generate a list of items to then be ranked. First evaluation of the pilot shows better performance levels than when compared to IPCC and UPCC methods, as the inference engine receives apriori information through the social network analysis, reducing cold start problems. Although the system is designed to handle big data, it seems that testing was limited by small data sets and a low number of users took part in the used *NASA TLX* evaluation. Using Edge Intelligence for cultural heritage recommendations and distributed computing being able to handle big data will be a promising way for future systems to come.

Investigating the history of recommender system development for museums, it is evident that most systems were designed for in-house user experiences and on-site guided tours. The literature furthers suggests that museum recommender systems are inherently driven by semantics, evident throughout analysis of the publications that conveying and reflecting meaning is very important for models.

2.3.1. Evaluation of recommender systems

Traditionally, recommender systems were mostly evaluated from a technical perspective in terms of how accurate returned lists of recommendations were, leading to the fallacy that a higher accuracy or precision of recommender algorithms leads to a better user experience and satisfying results (Knijnenburg and Willemsen, 2015). Over the last two decades, a shift towards a more user-centric evaluation approach was visible, establishing frameworks of *Human-Recommender Interaction (HRI)* (McNee et al., 2006) that focus on the generation of useful recommendations for the user, acknowledging that accuracy and precision might not always be the defining terms of a functioning system serving satisfying recommendations from a user perspective.

In principle, recommender systems can be tested offline or online, depending on the scale of the testing, the focus of interest (e.g., algorithms outperforming current SOTA ones, conversion rates, user satisfaction etc.) and the means available.

Offline methods focus on the evaluation of the algorithms used in the model and their outputs. Methods of offline evaluation therefore include methods such as verification of algorithms, regression testing, and offline experiments based on predictions and historical data (K. Falk, 2019). Offline methods are still amongst the most often used methods today, but as explained above, fall short in actually including real-world users and often lack consistent use of algorithms and rigorous methods of reporting results, making it hard to reproduce or actually compare them (Ekstrand et al., 2011). Offline methods alone therefore should not be used to evaluate a system as they exclude the user, their changing needs and expectations as well as their overall experience using the system (Ekstrand et al., 2015), risking a system that performs well on paper but not in real world situations.

Online methods include users to evaluate the system, and range from classic A/B testing towards more larger scale evaluations, including continuous live testing or loop methods (Aggarwal, 2016; K. Falk, 2019). Knijnenburg and Willemsen state that “[p]roper evaluation of the user experience of a recommender system requires conducting a user experiment, either in form of a lab experiment or a randomized field trial (which includes - but also extends beyond - conventional A/B tests)” (Knijnenburg and Willemsen, 2015).

Controlled lab experiments have the disadvantage of observing the user outside the natural environment where the system would be usually deployed, but give the option to directly observe the user behaviour and ask questions in person (K. Falk, 2019), whereas in-the-wild studies have shown to reduce bias and recruitment errors, enabling testing of the system in a real world setting (Aggarwal, 2016). Loepp and Ziegler (2019) further highlight that evaluation of RS is context-specific and caution is to be taken in terms of if the actual item can be consumed (e.g., e-commerce) or if proxies are used to represent the recommendation, meaning that questionnaires actually ask users about the proxies and not the actual item causing wrong inference.

Museums, seen as learning environments, pose further demands to a holistic evaluation process as technology-enhanced learning environments require different approaches than purchase-driven systems and real life testing under realistic online conditions has proven to yield the best results in terms of evaluating user experience and satisfaction (Erdt et al., 2015).

Finally, the goals of successful evaluations are therefore manifold, and the most common ones include accuracy (metrics-based), coverage (the higher the accuracy, the lower the coverage, leading to some items being excluded), confidence and trust, novelty (recommendation that the user was not aware of before), serendipity (the level of positive surprises in the recommendations), diversity (the recommendations within a single list of items should be as diverse as possible), robustness and stability, and scalability (Aggarwal, 2016).

2.4. Summary

Museum environments and visitor experiences are not just technologically determined, but involve a web of networks defined by social, cultural, and power relations, and constituents that draw on and contribute to hybrid spaces which merge digital and physical attributes.

Data and related data-intensive methods have become of major importance for institutions, not just to create, store, and share information and knowledge, but also to create new visitor experiences and means to engage constituents. Whilst visitor experiences focus on outward data practices, where institutions produce digital media to be consumed, it is further evident that data is used internally for operational and strategic purposes, such as visitor analysis and metrics used for economic justification.

Audiences, especially those being digital natives, are using technologies more and more seamlessly to interact with their environment, not separating the digital between the analogue, incorporating the above identified hybridity into their everyday life. Traditional frameworks of audience segmentation being those focused on physical visitors or trying to replicate offline audiences online, are not applicable anymore as the digital now also engages audiences that have never been to a physical museum or constituents interacting with institutions for the first time through digital means. Digital media further enables users to take part in cultural life and establish cultural practices that circumvent institutional channels at all, which can lead to a democratisation of cultural spaces and practices.
Introducing a recommender system into this environment means first and foremost to understand the landscape of museums and artificial intelligence applications, their uptake and potential use, and the possible barriers as the review made evident that museums lag behind other GLAM sectors in their uptake, but that applications are generally seen as promising in the future **(RQ-S1/2)**. Whereas museums were identified as being slower in their uptake of AI

technologies, there is a gap in academic literature in drawing a comprehensive picture to what those barriers might be. Gaining knowledge around AI technologies specifically is important to understand the environment without making premature conclusions and generalisations drawn from literature around digital technologies and projects found online that might not be applicable to AI and museums.

Concerning the potential role and use of recommender systems **(R1)**, the literature review shows that such systems have been deployed in museums and cultural heritage institutions over the last two decades, but their focus was mainly on in-house applications, tour-like narratives or small pilots. The majority were developed without the inclusion of professional users of the systems from a design and development stage on **(RQ-F2)** or discussions around how they would possibly use such a system in their job roles **(RQ-F1)**. It further seems that a lot of the discussed systems were, if even implemented, not sustainable over the projects' lifespan, which raises questions about the factors leading to this and the reasons behind abandonment of those systems.

Evaluation of recommender systems can be successfully undertaken offline, online or through a combination of both. However, online studies were identified as the main contributors towards user understanding as purely offline methods risk developing systems bypassing constituents and their needs. Some of the publications acknowledge this fact, however they fall short in addressing user-centric evaluations beyond pilots and small-scale user studies, reducing ecological validity of the conducted studies **(RQ-F3; RQ-UXI1/2/3)**.
Using a recommender system with museum online collections **(R2)** therefore requires a holistic approach considering all possible constituents of the system from an early stage, investigating the data used for the system, but also the data output by it. Thinking about how data shapes and is shaped by the institutions and userships requires ethical considerations and discussions around agency, two topics that were not the focus of any reviewed publications **(R3)**.

Thus, having reviewed relevant literature, it is evident that there is a need to understand not just the wider AI and museums environment, but also how those environments will accommodate future technologies sustainably. Whilst it is evident that museum environments are changing, digital technologies are gaining importance and are seen as new ways to meaningfully engage constituents. There exists a gap between theory and practice, and this thesis addresses this through the establishment of a methodology that incorporates empirical accounts and theoretical frameworks to establish a novel and more holistic approach of researching the hybridities around spaces, AI, and materialities that emerged through this literature review. Exactly this hybridity and fluidity of those environments demand a flexible research approach that it can react to and accommodate to the challenges that may arise throughout this PhD - personally, with constituents, and the system.

3. Methodology

The thesis draws upon multiple methods to investigate and answer the research questions outlined before. With practice-based research at its core, the development of the museum recommender system *MuseREC* under a *Practice paradigm* not only scrutinises the technological artefact, but conceives the "world as a network of performances that are durable, because ways of doing things are coded in minds, bodies, artifacts, objects, and texts, and connected together so that the result of performing one activity serves as a resource for another" (Kuutti and Bannon, 2014, p. 3545). The practice of developing, implementing, using, and evaluating the recommender system and the interactions it fosters with its constituents therefore serves as the gravitational centre of research interventions and "the *basis* of the contribution to knowledge" (Candy, 2006, p. 3). Practice and research, thus, are understood as a co-evolutionary process.

As discussed in the *Portfolio* (p. 7), *Five-step recommender life cycle* serves as the scaffolding of the practice-based contribution (see *Portfolio,* p. 7), the development of *MuseREC* with its aim to serve as both a working RS for museum online collections and a research methodology for this thesis. *MuseREC* was set up to develop and evolve alongside the other research interventions and to inform them whilst also being informed by them – a constant iterative cycle of feedback. Considerations around the practical component were feasibility and usefulness as well as availability of technologies. This meant designing *MuseREC* with the goal of being a functioning MVP and an artefact at hand to support and cater to the other research studies. *Doing* practice-based research further means mitigation of risks and challenges and to find (creative) workarounds throughout the process. Developing *MuseREC* therefore meant micro-managing a technical project alongside the PhD, ranging from handling the partnerships that are further outlined later in this chapter and juggling real-life museum practice to overcoming technical difficulties and computational setbacks.

Whilst practice creates new forms and artefacts, research strives to generate new forms of knowledge and understanding (Edmonds et al., 2005). During this process, the recommender system (see *Portfolio*, p. 19) is not perceived as definite, but as an *unfolding* object that is transient, changing, and question-generating, subjected to an epistemic practice of relational dynamics (Knorr Cetina, 2005) and to a rationalising framework to avoid falling into the *technology trap* of digital heritage research, but also to apply rigour, although conditionally (Parry, 2005). Parry describes the technology trap, a term originally defined by Tomislav Šola, as the risk of using technologies in museums just for the sake of it, “allowing technology to become self-serving and [let] ourselves be guided by it” (Parry, 2005, p. 333). Critical analysis and professional experience can be a remedy to avoid falling into such a trap where technologies are used with no regards for any implications on museological practices (Parry, 2005). To avoid the pitfalls highlighted by Parry, and to circumscribe the scope of the thesis theoretically, the next section introduces a framework that anchors the research in a postphenomenological approach to frame the system and open it up to socio-cultural investigations.

3.1. A postphenomenological framework for recommender systems

Rigour is prescribed by an empirical-philosophical approach forged out of a postphenomenological vocabulary which enables investigation of the social and cultural roles of the recommender system and its constituents through the relations humans have with technological artefacts, rendering such systems not as “merely functional and instrumental objects, but as mediators of human experiences and practices” (Verbeek, 2016, p. 190).

The characteristics of a postphenomenological approach can be summarised as follows:

- It focuses on understanding the roles that an AI application, specifically a museum recommender system, plays in the relations between humans and the museum sphere
- It includes empirical accounts of the role the recommender system plays in human experiences and practices as a basis for philosophical reflection

- It investigates how, in relations that arise around the recommender system and its usage, a specific world as well as specific subjects are constructed
- It has the possibility to open up the recommender system and its constituents to conceptual analysis of various dimensions of human-technology relations - they can be epistemological, aesthetic, ethical, political, etc.

Adapted from Rosenberger and Verbeek (2015, p. 31)

The next paragraphs establish the postphenomenological framework for this thesis, touching upon its key concepts in brief. The term *framework* was purposely chosen as postphenomenology is not understood as a methodology that offers a strict step-by-step sequence scholars can follow. Rather, it embodies - which makes it even more attractive - “a specific way of investigating technologies, an approach to technology that combines an empirical openness for the details of human-technology relations with phenomenological conceptualization” (Rosenberger and Verbeek, 2015, p. 32). It investigates the various purposes of the museum recommender system, how it is perceived and experienced by humans, and how it mediates the constituent/human-recommender/technology-museum/world relationship. Further, it acknowledges that what it does *in specific* to this relationship can only be answered by the empirical studies conducted in the course of the PhD. The framework supports practice research since “postphenomenological claims should be understood as posed from within a practical, perspectival, and situated context” (Rosenberger, 2017, p. 472), therefore offering the thesis the necessary space to build a theoretical scaffold *on-the-go* rather than casting it into a mould of theory *apriori*. This openness further enables combinations of postphenomenology with other concepts, where it helps to contextualise the research further or where postphenomenology alone would have its limitations. Such mixed approaches have been postulated (Ritter, 2021) and successfully applied (Hauser et al., 2018; Moens, 2018) in relation to other *materiality-focused* theories, such as actor-network theory (ANT) or Donna Haraway’s writings about cyborgs (Aagaard, 2017) - rendering a more holistic picture that is permeable enough to be open for different perspectives, second opinions, and healthy speculations.

About postphenomenology

Postphenomenology, rooted in Don Ihde's key concepts of human-technology relations (Ihde, 1979) and his relational ontology, has gained popularity as a philosophy of technology and an approach to analyse the networks of relations around technologies (Rosenberger and Verbeek, 2015). It can be understood as a hybrid that draws upon classical phenomenological theories and American pragmatism to create a philosophy that is non-foundational, anti-essentialist and focused on how human beings perceive and experience a world that is mediated through and co-shaped by technological artefacts.

Human-technology relations describe the ways how human beings can relate to technology (bodily-perceptual relationships) and are based on Ihde's four basic forms of relations: embodiment relations, hermeneutic relations, alterity relations, and background relations (Ihde, 1990), however, this list is not exhaustive and has been extended over time, e.g., cyborg relations (Rosenberger and Verbeek, 2015). In embodiment relations, for example, humans perceive *through* technology and cyborg relations are defined through their merger with either the human being itself (e.g., a pacemaker) or its environment (e.g., augmented reality). Applying these various forms helps to define and carve out the relationships constituents can have in regard to a device and the recommender itself, defining the bodily-perceptual accounts of user experience.

The relations are described through a *relational ontology,* where technologies are "understood in terms of the relations human beings have with them, not as entities 'in themselves'" (Rosenberger and Verbeek, 2015, p. 19). Postphenomenology accepts that technologies mediate relationships between humans and their world, but their relations are asymmetrical. Thus, technologies "help to shape the 'subjectivity' of their users and the 'objectivity' of their world" (Rosenberger and Verbeek, 2015, p. 19), but human beings and nonhumans are not indifferent. It is a sole human attribute of *subjectively being in the world*. This does not mean that postphenomenology works on Cartesian assumptions of dualisms, on the contrary, it rejects a dualist world view, but postulates that subject and object are "mutually constituted in their interrelatedness" (Verbeek, 2005, p. 130) and it is exactly this interrelatedness that prescribes the active role technologies play in enabling and shaping relationships between humans and their realities.

Postphenomenology further builds upon Husserl's phenomenological concept of *intentionality* which postulates that "[h]uman consciousness never exists in itself, but only as consciousness-of-something" (Verbeek, 2005, p. 109). In the same way as consciousness does not exist in itself, neither do technologies exist as technologies-in-themselves. The ability to co-shape human-world relations is not an intrinsic property of a technology as "[a]rtifacts can only be understood in terms of their relation that human beings have with them" (Verbeek, 2005, p. 117), they are technologies "in-order-to" and "always and only function in concrete, practical contexts and cannot be technologies apart from such contexts" (Verbeek, 2005, p. 116). Thus, technologies are context-dependent and can have different meanings and functions depending on their application - they are *multistable* and not neutral. Postphenomenologists understand the concept of multistability as "a technology's potential to support multiple relations; a single technology can be understood in multiple ways, taken up in many contexts, and employed for various purposes" (Rosenberger, 2014, p. 376). A laptop, for example, can be used to run various programmes, it can be used at work or to play games, but it can also, theoretically, be used as an expensive door stopper or serving tray. Thus, museum environments themselves and also the constituents and technologies situated in them can be perceived as multistable and can therefore function as spaces of postphenomenological investigation.

3.2. Partner museums

To situate the system in the wild, partnerships with three museums, which are introduced below, were established, rendering them as *civic laboratories*, where social relations between entities and civic programmes are produced, studied, interpreted, mobilised, and enacted upon (Bennett, 2005). However, this research goes further than the original idea and extends the laboratory beyond objects and visitors and their relation between and to each other, by dissolving physical as well as virtual boundaries, redefining it as a hybrid space. This therefore establishes a new hybrid environment to foster a *computer-supported* *practice* (Kuutti and Bannon, 2014, p. 3543) entangled in the relations of the entities museum, technology and constituents and their myriad of socio-cultural and technological implications rendering the museum as a place to “raise philosophical questions and issues while at the same time taking seriously empirical investigation into technology” (Verbeek, 2005, p. 9).

At the beginning of this PhD, partnerships with three institutions - the Smithsonian Institution (SI), the National Gallery (NG) in London, and Manchester Art Gallery (MAG) - were established to investigate their collection data and related practices, work collaboratively with their professionals and to situate the recommender system in-the-wild. These institutions were convenience-sampled based on established contacts or networks and represent similar founding histories, but differences in their digital strategies and varying institutional size and funding structures.

The nature of these partnerships was to gain access to various datasets of collections from institutions with different set-ups in order to get an overview about what digital museum data looks like and investigate their usefulness for AI-applications. Such datasets were contributed by the SI, NG and MAG and were subsequently analysed at the beginning of the PhD to establish both, a field survey and a plan to further inform the practical development of *MuseREC.* Over the course of the PhD, the relationships to these three institutions intensified and professionals at these institutions contributed invaluable insight into their work, mainly through focus groups and talks as part of the co-evolutionary practice-based research methodology. In course of the PhD, further collaborations with the Badisches Landesmuseum in Karlsruhe, Germany, and Art UK were established. These two institutions played different roles than the before mentioned three core institutions: the Badisches contacted the thesis’ author after having given a talk about AI in Museums for AI4LAM and invited the author to hold a focus group in Germany, Art UK provided data for the online user study to avoid institutional siloed data and to provide a broad spectrum of artworks sourced from UK's national collections. Art UK further offered its audience panel to test the recommender system with. Data management plans (DMP) were put in place, where institutional regulations regarding licensing and data sharing required it, namely NG and MAG. The institutions and their accounts about their digital data practices during the COVID-19 pandemic were first introduced in Noehrer et al. (2021b).

Manchester Art Gallery

MAG, a local authority owned museum operated by Manchester City Council, is part of the Manchester Museum Partnership consortium, which includes two University museums, the Manchester Museum and The Whitworth art gallery, and offers shared administrative and research capacity and a network of support and practice for its three member organisations. The main gallery is based in the centre of Manchester, with a former costume gallery and restoration centre in two branch sites in city parks. In 2019 it recorded 731,002 visitors, making it the most visited museum in the city, and also reached high numbers of people across Manchester city wards through participatory programmes (Manchester City Council, 2019a). It makes a GVA (gross value added) contribution to the local economy of £13.7m (2019b).

The history of the gallery dates back to 1823 when the Royal Manchester Institution for the Promotion of Literature, Science and the Arts was founded by a group of artists, supported by local gentry and businessmen to boost the local arts market through exhibition and patronage, providing science and arts education through lecture programmes and honorary professorships (Wolff, 2013). It became the Manchester Art Gallery when it was handed to the City Corporation in 1882, with a budget to extend its collections. Dedicated as a museum for all people in Manchester, it shared similar origins, objectives and champions to its peer, the Manchester Art Museum, which was established in 1886 by Thomas Horsfall as an arts education provider for the city's working classes (Harrison, 1985).

The ambition to “diffuse useful knowledge” corresponds to the interests of the eponymous 19th century Society for the Diffusion of Useful Knowledge, which influenced the foundation of Mechanics Institutes and Lyceums and led to similar movements in the US (Portolano, 1999). MAG’s current vision statement, as articulated on the gallery’s webpages, reflects these historical foundations and outlines its contemporary civic and social mission to position the museum as a “‘Civic Think Tank’; creating a convening space for voices across the city, providing creative education for all classes and cultures” (Manchester Art Gallery, n.d.). It also reflects the interests of the gallery’s current leadership in Ruskinian theories of ‘useful art’, connected to the broader international ‘Arte Util’ network, and, as the website acknowledges, corresponds to the strategic priorities of local government (Manchester City Council, 2020) and those of key funder Arts Council England, to produce social impact through investment in arts and culture, to promote inclusion and education for the residents of Manchester, and to attract visitor economies to the city centre.

The mission of the gallery focuses on attracting and welcoming diverse audiences into the gallery spaces for the purposes of civic dialogue and education, with an emphasis on arts for health and well-being. There is little if any reference to digital or technological strategies as a means to “diffuse useful knowledge”, however, and although the ‘Learn’ section of the gallery’s website signposts engagement activities and offers curated digital content for self-guided exploration for schools and colleges, the majority of content is oriented towards encouraging visitors to enter the physical gallery spaces. There is a collections search interface which allows for simple term searches of text and images relating to over 25,000 objects, although many of these have not been digitised. Personal use of the images is permitted for browsing and viewing; for further use of more than a single copy user need to enquire via a licensing enquiry form. During lockdown there was some further content developed to allow users to access ‘The Gallery at Home’, including online talks and creative activities, however the emphasis remains on temporary activities which are stop-gaps for the time period of COVID restrictions, rather than new strategies to be integrated into the gallery’s future.

The National Gallery

Directly grant-in-aid funded by the Department for Digital, Culture, Media and Sport (DCMS), the NG in London has charitable status and is constituted as a Non-Departmental Public Body. The Gallery dates to the same year as the foundation of MAG, although its origins were ostensibly supply rather than demand-led, when in 1823 first the collector John Julius Angerstein and then landscape painter and collector Sir George Beaumont gifted their collections of paintings to the nation, necessitating a new national institution with suitable buildings for conservation and display of these collections.

The national collection now has over 2,300 paintings, representing classical western traditions from the 13th to the 20th century, acquired via a mixed economy of donation, fundraising via public appeal, grant-in-aid, trusts and foundations and private donors, and primarily displayed at the iconic building in Trafalgar Square. There is a discrete access policy which articulates the Gallery's commitment to "free access for all" (National Gallery, n.d.), although some temporary and special exhibitions have admissions charges and income is also generated through loans, touring exhibitions, licensing of image rights, publications, and merchandise. The Gallery has the power to raise capital via investment under the Museums and Galleries Act 1992 and maintains a carefully worded financial policy which stipulates the principles through which works of art are capitalised and appear on their balance sheet, to the concern of the Trustees, as an arbitrary valuation on their date of acquisition (HMSO, 2020), reflecting the tensions between the fluctuations of the art market, the governance of a Non-Departmental Public Body, and the need for transparency over public funds.

Following a Strategic Review of DCMS-sponsored museums (2017), the Gallery launched a new Strategic Plan that outlined significant ambitions for its business model and its use of digital technology. These included seven strategic objectives which, alongside continuation of the acquisition and conservation of major paintings, signalled an ambition to engage further within their programmes with contemporary artists and museum learning, and notably to "create a National Gallery with digital at its heart, to reflect a more digital world" (National Gallery, 2018). This pre-pandemic shift in strategy aims to embed digital technology and data capabilities across programmes to support visitor experience and audience research within the Gallery, as well as to present new opportunities for public engagement and digital display and consumption. Furthermore, digital is also noted as a key lever for the strategic objective to raise levels of self-generated income to match the Grant-in-aid and become 50% self-funded.

As the Annual Report for 2019/20 states, the pandemic has put a strain on the ambition to achieve sustained self-funding within the next few years. However, investment in digital capacity has furthered the Gallery's mission to provide public access to its collections, not least in supporting digital engagement during lockdown. A first stage in this was the Collections Information Project which required the complete rewriting of collection data entries, as well as investment into ticketing and customer relation management systems. New partnerships with technology and academic partners, including the Alan Turing Institute, King's College London, and Google Arts and Culture are supporting experimentation with virtual and augmented technology and an innovation lab, National Gallery X, a move which will presumably further research ambitions in technical art history as well as provide the means to take the collection out of the Gallery, and may mitigate the plans for an expanded national programme which have been curtailed by loss of funds due to the pandemic (HMSO, 2020).

The Smithsonian Institution

The bequest of James Smithson, in 1835 via his heir-less nephew, to "the United States of America, to found at Washington, under the name of the Smithsonian institution, an establishment for the increase & diffusion of Knowledge among men (sic)" (Portolano, 1999, p. 65) led to a protracted political debate within Congress about how to interpret this vision. One group held that the institution should pursue discovery of new knowledge through the funding of empirical scientific research for the benefit and progress of society, following the Baconian philosophical traditions established by the British Royal Society popular amongst nineteenth century US scientists and their supporters. The other group favoured the governmental reform of public education via 'common-school' educationalism in useful arts and sciences, which echoed the moral improvement and settlement agendas found elsewhere, including in Victorian Manchester. This debate was eventually resolved through the founding of the Smithsonian Library and Museum in 1846, to exhibit and promote the products of scientific enquiry, including entire government collections of art, material sciences and natural history research. These were housed in the National Mall, maintained by resident scientists, complemented by research grants programmes and extremely popular public lectures. The latter were carefully regulated in an attempt on the part of the Organizing Committee to maintain the 'neutral ground of science' and keep the Institution's knowledge diffusion free from politics, a task near impossible to achieve during the rampant partisanship of mass democratic politics in mid-19th century, civil war-torn America, until a fire put an end to the public lecture hall in 1865 (Adcock, 2014).

The SI currently operates nineteen museums, eight research centres, gardens, and the National Zoo. Writing well before the advance of digitisation and digital museology, Portolano (1999) points out that by establishing a museum-dominated complex, the Smithsonian Institution retained its mission to diffuse knowledge to the common man, although it "does so primarily through the medium of exhibition of material objects" (Portolano, 1999, p. 79). Fulfilling the remit to advance scientific enquiry, digital technology and strategy now have a notably central place in the Smithsonian's mission and receive significant attention and investment. Its current strategic plan (2017-2022) identifies seven goals to achieve the vision to "build on its unique strengths to engage and to inspire more people, where they are, with greater impact, while catalyzing critical conversation on issues affecting our nation and the world" (Smithsonian Institution, 2017). Two of these goals - to reach 1 billion people a year, and to drive "large, visionary, interdisciplinary research and scholarly projects" - combine to articulate the Institution's continuing commitment to increasing and diffusing knowledge, and are clearly predicated on digital technology, innovation and data science. The "digital-first strategy" is supported by a Digital Access Agenda which was incorporated into its strategic plans as early as 2014, and which emphasises technology use for enhancing visitor experience within the museums, as well as full digitization of the collections with easy, accessible, and low/no cost access to extend engagement and participation “among learners everywhere” (Smithsonian Institution, 2014). This has led to metadata, 3D objects, datasets, and a huge stock of images being released into the public domain as Creative Commons Zero (CC0) that can be used, manipulated, transformed, and shared without the need for institutional permissions (Smithsonian Institution, n.d.).

3.3. Qualitative and quantitative studies

This section introduces the three studies that were conducted over the course of the PhD, aiming to establish a series of research interventions that contribute the empirical and experimental backbone to answer this thesis’ research questions, to render a picture of the status-quo, and future directions alike.

3.3.1. Survey: Museums and AI Applications (MAIA) Survey

At time of writing this thesis, academic literature lacked a comprehensive picture of the actual usage of AI applications in museums, making it difficult to anchor the recommender system and its possible future use in the wider museums and AI landscape. Therefore, an online survey was conducted in 2021 to investigate the adoption and use cases of AI in the sector. Especially timely as data-led activities in museums are further declining successively since 2013 (Arts Council England and Nesta, 2020), possibly risking museums not being able to keep pace with the rapid rollout of new technologies or indicating that developed software and products do not cater to the needs of CH institutions. It is therefore important to determine a bigger picture view about the usefulness and application of AI in museums to identify future symbiotic relationships and also barriers to its implementation.

Participants were recruited via email (including all ICOM country committees with addresses listed on ICOM’s website), listserv messages, community forums, and Twitter - with the option to further share the web link with colleagues and other relevant professionals. The survey had clear inclusion and exclusion criteria and addressed the international museum community. The survey was constructed using an established HCI framework, containing open and close-ended questions

(Lazar et al., 2017) and was piloted with five staff members of the University of Manchester with experience in survey design. Participants received a Participant Information Sheet (PIS), had to give consent before being able to access the survey and had to fulfil the inclusion criteria. Responses were collected anonymously and the dataset is publicly available as CC BY 4.0 via Figshare (Hughes-Noehrer et al., 2022a). A Jupyter notebook including Python code to analyse the data was uploaded to Zenodo and is freely available (Hughes-Noehrer, 2022d). The survey was hosted online using Qualtrics (2021), was logged in the University of Manchester's ethical review system and received approval (Ref: 2021-12842-20370).

3.3.2. Focus Groups: UX/UI Museum Professionals

To gauge museum professionals' perception of the recommender system, its potential use for museums and to gather valuable information about how recommender engines can support museum professionals and other users, four focus groups were conducted in February and March 2022. Working together with domain-expert end users is a proven method to not just elicit domain professionals' opinion regarding computational systems, but can further contribute to a variety of other variables, such as desired functionalities and insight around other end-user groups, helping to create and modify systems for the better (Costabile et al., 2003). Participants (please refer to *Table 3.1.* below) were drawn from pre-existing social groups - being work colleagues at the institutions - as this set up shows to provide a stimulating environment for discussion and debate, but also challenges any possibly occurring discrepancies and subjective beliefs due to a shared social context (Bloor et al., 2002). Professionals were either representationally sampled (MAG and Manchester Museum Partnership), representationally and snowball sampled (SI and NG) or invited after a group of professionals showed active interest in the system and contacted the researcher (Badisches Landesmuseum). The Badisches Landesmuseum is, at time of writing, looking into developing a recommender system for creative user empowerment, and was therefore deemed suitable as an additional focus group partner albeit not being a partner institution. For this reason, the museum was included as the shared interests could provide synergies and valuable input to the system development, its implications and use. Whilst all participants apart from three agreed to the use of their names and direct quotes, it was decided to not include any names in the presentation of the results as participants were asked to contribute through the lens of their professional roles at the institutions. However, the author is aware that names could be attributed through matching job roles and institutions and therefore participants who did not agree to the use of direct quotes were subsequently anonymised. The focus groups were composed of a twenty-minute demonstration and introduction to the system by the thesis' author followed by semi-structured group discussion lasting around two hours. Three focus groups were held online using the video conferencing software Zoom (Zoom Video Communications, Inc., 2022) and the group at Badisches Landesmuseum was conducted on-site. Design of the study was informed by the Standards for Reporting Qualitative Research (SRQR) to assure sound and rigorous data collection and reporting (O'Brien et al., 2014). Altogether there were 30 professionals participating, coming from a varied range of professional roles. It is expected that professionals will have different requirements of such a system than the general public has. This can further inform system development for both user groups, giving insights into workflows and data practices of museum institutions.

Focus groups followed a topic guide (Lindlof and Taylor, 2017) based on generative questions (Rubin and Rubin, 2005) to encourage extensive replies in an open format. All focus groups were transcribed in detail and cross-checked by a second researcher, to ensure reliability and validity (Kvale, 2007), and analysis was performed by using inductive thematic analysis (Braun and Clarke, 2006) using the software NVivo (QSR International Pty Ltd, 2020) to analyse, code and re-code the focus group data, matching sentences to the primary themes in the groups with an unweighted Cohen's Kappa ($\kappa = 0.95$) to reach a high level of agreement between coders (Landis and Koch, 1977).

Participants received a PIS and were asked to give written consent before the groups had started (Hughes-Noehrer, 2022b). Due to the nature of the focus groups solely involving individuals who participated in their professional roles, ethics approval was not sought after having consulted the University of Manchester's Ethics Decision Tool.

Participant (P)	Position	Institution	Location
1	Museum Director	MAG and the Whitworth	UK
2	Curator of Egypt and Sudan	Manchester Museum	UK
3	Collection Information Manager	MAG and the Whitworth	UK
4	Digital Communications Manager	Manchester Museum and the Whitworth	UK
5	Senior Curator	MAG	UK
6	Learning Manager: Lifelong Learning and Volunteering	MAG	UK
7	Collection Information Manager	The National Gallery	UK
8	Conservator	The National Gallery	UK
9	Photographer	The National Gallery	UK
10	Research Centre Manager	The National Gallery	UK
11	Web Analyst Developer	The National Gallery	UK
12	Business Analyst	The National Gallery	UK
13	Anonymous staff	The National Gallery	UK
14	Co-Director National Gallery X	The National Gallery	UK
15	Art Handling Coordinator	The National Gallery	UK
16	Principal Scientist	The National Gallery	UK
17	Scientist	The National Gallery	UK
18	Acting Director, Smithsonian Office of Educational Technology	Smithsonian	US
19	Supervisory Museum Specialist, NMAAHC	Smithsonian	US
20	Research Data Scientist, Lead Smithsonian Data Science Lab	Smithsonian	US
21	Anonymous staff	Smithsonian	US
22	Senior Digital Program Officer	Smithsonian	US
23	Program Manager, Smithsonian Open Access Initiative	Smithsonian	US
24	Scientific Leadership Trainee	Badisches Landesmuseum	DE
25	Head of Department Ancient Cultures, Curator of Egypt and Classic-Greek Archaeology	Badisches Landesmuseum	DE
26	Anonymous staff	Badisches Landesmuseum	DE
27	Communications and Controlling Leadership Trainee	Badisches Landesmuseum	DE
28	Curator Middle Ages 1050-1500 and History of Baden	Badisches Landesmuseum	DE
29	Project Lead AI and Museum, Creative User Empowerment	Badisches Landesmuseum	DE
30	Cultural and Folklore History Leadership Trainee	Badisches Landesmuseum	DE

Table 3.1. *Focus group participants*

3.3.3. Online study: User experience/user interaction (UX/UI) study

To discover users' preferences on interaction with a museum online collection recommender engine and gather results about the models' performances in terms of engagement, an online user study was conducted in 2022. Participants took part in a controlled between-subjects/within-subjects online study about novel methods of browsing museum online collections and the study material is available via Figshare (Hughes-Noehrer, 2022f). Online studies have proven useful in HCI research (Horton et al., 2011; Kohavi et al., 2013; Lazar et al., 2017) and in the case of this study the advantages of running the experiment online (e.g., ease of access, better generalisability, reduction of experimenter and demand effects, and greater openness of the research process) clearly outweigh the disadvantages (e.g., multiple submissions - which were controlled for via IP verification - possible drop-outs, or absent interaction) (Reips, 2002). Further, controlled studies provide the "best scientific design for establishing a causal relationship between changes and their influence on user-observable behavior" (Kohavi et al., 2009, p. 140), therefore conditions were randomly assigned and controlled for. The study was designed in accordance with established HCI guidance on controlled experiments (Blandford et al., 2008; Kohavi et al., 2009, 2013; Lazar et al., 2017) and was piloted with five persons at the University of Manchester who have experience in designing and conducting online user studies.

To establish a closer to the real world browsing environment with greater ecological validity that enables a more voluntary engagement - two points highlighted as often missing in lab-based engagement studies (Doherty and Doherty, 2019) - participants took part in their own time and using their own device, simulating a natural browsing experience. The study was hosted on a University of Manchester owned virtual machine. To be able to use a varied set of collection data from multiple institutions and to avoid pitfalls of institutional siloed data, a partnership with Art UK[5] was established and a dataset containing 35,000 objects out of Art UK's database containing 309,456 items (date of export:

[5] The Art UK online platform contains artworks of any public collection in the UK; currently over 3,400 institutions (for more information please visit https://artuk.org/).

June 2021) were randomly selected. Participants were recruited via the Art UK audience panel, online advertisements on university channels, Twitter, and Prolific.co. On landing on the experiment's webpage users were supplied with a PIS and were required to give informed consent before proceeding further. In case a user did not give consent, the process abandoned and took the user to the end page of the study, barring them from participation - no data was collected about these users. Participants were not made test aware, however, they were informed about the nature of the study being held online, the data collected, and the approximate duration of the study. After having completed the post-study questionnaire, participants were presented with a debrief and contact details of the experimenter in case there were further questions or more information desired. Demand effects, even if participants elicited the study's purpose, were not expected (Berinsky et al., 2012; Mummolo and Peterson, 2019). The study received ethical approval from the University of Manchester (Ref: 2022-13746-22392).

Data collection

During the study (see *Portfolio,* p. 38), intrinsic and extrinsic data was collected through a mixed-methods approach to capture pragmatic (i.e., usability) and hedonic (i.e., fun and pleasure using the recommender) characteristics (O'Brien and Lebow, 2013). Intrinsic data includes application-specific user interaction events (i.e., logging whenever a user interacts on an interface element), such as clicks on artworks or buttons, and page views, to then calculate interaction metrics post-study. Extrinsic data was gathered through users' explicit selection of five to ten artworks to proceed and a post-study questionnaire. For exemplary screenshots please see *Figure 3.1.* (artwork selection overview) and *Figure 3.2.* (detailed artwork view).

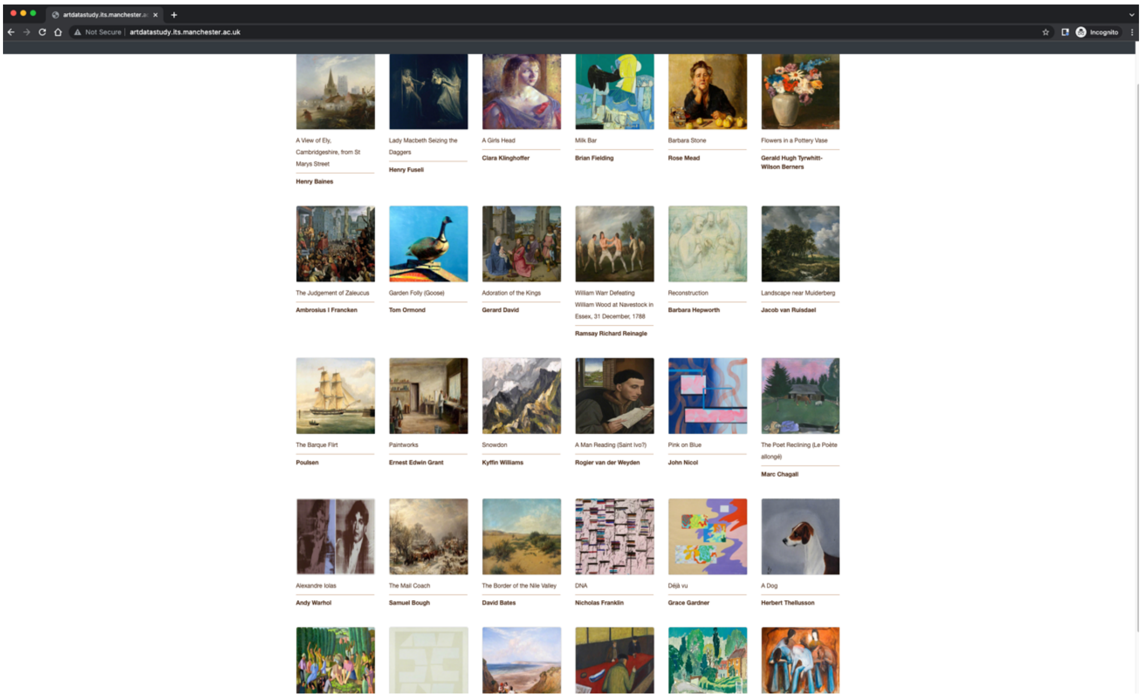


Figure 3.1. *Artwork selection start page*

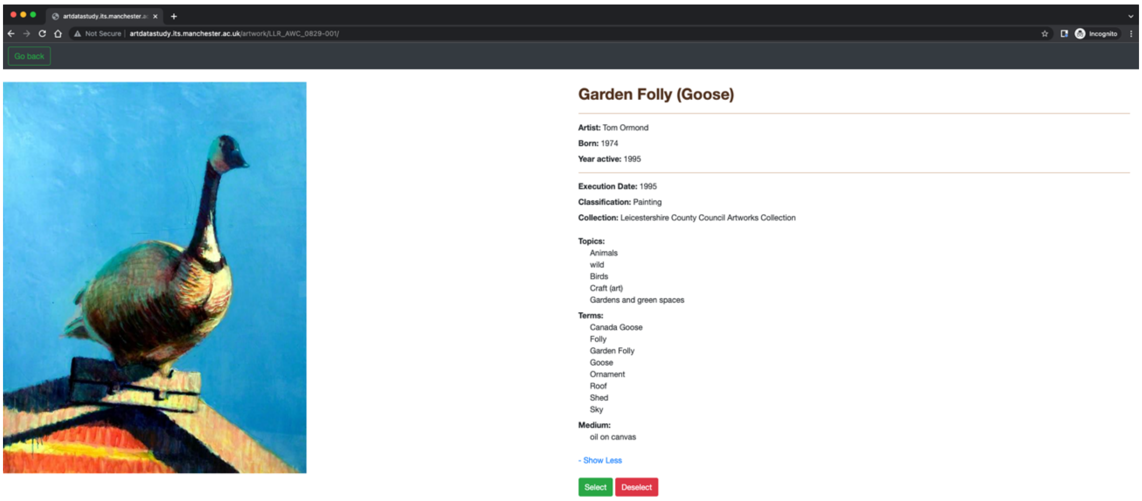


Figure 3.2. *Detailed artwork view*

Post-study questionnaire

The questionnaire was informed by the user-centric framework of Knijnenburg et al. (2012) to capture the users' experience of interacting with the museum recommender system. Participants were asked questions about how they perceive the system and further questions to evaluate it for each part of the study (*Part I* and *Part II*). Whilst perception and evaluation of the system might seem very similar, there is, however, a subtle difference as "[p]erception denotes whether certain objective system aspects register with the user at all, while evaluation denotes whether the perceived aspect has any personal relevance to the user" (Knijnenburg et al., 2012, p. 445). To give a short overview, the evaluation framework enables investigation of the following components already adapted to suit the thesis' remit:

Component	Description	Applied in thesis
Objective System Aspects (**OSA**)	Aspects of the system or a subset of it being evaluated, such as algorithms, input or output mechanisms	Not directly measured as SSA used as proxies
Subjective System Aspects (**SSA**)	User's perception of OSAs; mediates effects of OSAs on EXP and INT	Perceived recommendation quality
User Experience (**EXP**)	Users' self-relevant evaluation of the quality of the RS	- Perceived system effectiveness and fun - Choice satisfaction
Interaction (**INT**)	Evaluation through users' interaction with the system	- Logging intrinsic user data, e.g., clicks and dwell times (INT_log) - Subjective component in questionnaire: Intention to provide feedback (INT_q)
Personal and Situational Characteristics (**PCs** and **SCs**)	Gathers RS-independent user characteristics that can add valuable insight to evaluation	- General trust in technology - Demographics - Domain knowledge

Table 3.2. *The Knijnenburg et al. Framework adapted to the thesis' user study*

Interaction metrics

Whilst participants were interacting with the study, interaction data was anonymously logged in the background and securely stored on a university-owned server. Events were logged whenever a user performed an action on the web app's interface or the context of the web page view changed. Such data includes *go-back*, *click*, *artwork-selected*, *show-more*, and *artwork-deselected.* Data was stored in the following format: *id* (ID of the event), *timestamp* (timed to milliseconds), *content_id* (the ID of the artwork), *user_id* (an anonymous ID string generated per user) and *page* (the current web page when the event was performed). From the data, 19 interaction metrics were derived, such as time spent in Part I and Part II, time spent looking at artworks, number of artworks visited etc. (see *Appendix C,* p. 217, for a full list of metrics).

Statistical analysis

To discover relationships between the gathered interaction metrics, users, and the conditions, statistical analysis was performed using correlation analysis and regression. To determine if data is parametric or non-parametric Shapiro-Wilk's test for normality was applied to choose the correct method to test for correlations.

For parametric data, Pearson's correlation coefficient (*r)* was calculated and analysis was performed using ANOVA (*F)* to test for statistically significant differences between conditions.

For non-parametric data, Spearman's rank-order correlation coefficient (r_s) was calculated and conditions were analysed using the Mann-Whitney *U* (MWU) test and common language effect size (CLES) to evaluate possible relationships within the metrics.

To analyse the post-study-questionnaire in relation to the four study conditions, a Kruskal-Wallis H-test was performed to test for statistical differences between conditions and the questionnaire components as outlined above.

Findings were deemed statistically significant at $p < .05$ and statistical testing and calculations were performed using Python 3.8 with the Pandas (The pandas development team, 2020), SciPy (The SciPy Dev Team, 2020), and Pingouin (Vallat, 2018) libraries.

The above presented three research interventions are each presented over the following three chapters with in-depths summaries. Chapter 7, the discussion, then serves as the anchoring point to thread these strands together and critically evaluate their outcomes through a postphenomenological lens.

4. Museums and AI Applications (MAIA) Survey

The objectives of the MAIA survey were to establish a sense of how many institutions are using AI technologies, what they are used for and if they are not used, to understand the barriers to the technology and opinions thereof. It aims to investigate the breadth of uptake of AI in museums and what the current applications and use cases are. It further identifies the barriers of using AI technologies in museums. The survey questionnaire used to gather insight about the three questions can be found in *Appendix A,* p. 207. The survey further aims to draw a picture of the environment the RS will operate in and to comprehend the implications around the application of AI technologies in museums, observed through a wide lens as the survey did not aim to elicit in-depth accounts of specific museum professionals, but rather to develop a high-level understanding, a zoomed-out perspective of the current museum landscape in relation to AI technologies.

4.1. Results

Who responded? MAIA survey demographics

The MAIA survey received 288 responses, of which 22 had to be excluded due to irregularities, duplicates, or use of abusive language, therefore reducing the final total count to 266 participants, all of whom are included in the following results. Inclusion criteria was to work at a museum (including art galleries) in either a leadership, curatorial, collection management, technical/digital, conservation/restoration, educational or research role. Participants came from 27 different countries, with the majority of the population residing in the UK (60) and the US (24). Participants were asked about the location of the museum they are working at, the collection type, the professional role they are holding, and how many years they have been working in the sector (see *Table 4.1.*).

Demographic variables (dominant in bold)	**n (%)**[6]
Museum locations	**N=139**
Urban	**55 (40)**
Capital city	45 (32)
Multi-site (mostly urban)	21 (15)
Rural	15 (11)
Multi-site (mostly rural)	3 (2)
Collection type	**N=136**
Universal	**40 (29)**
Art	32 (24)
History	27 (20)
Other	20 (15)
Ethnography	6 (4)
Science	6 (4)
Natural History	5 (4)
Professional role type	**N=141**
Curatorial	**34 (24)**
Other	29 (21)
Leadership	28 (20)
Digital/technical	26 (18)
Research	11 (8)
Collection management and registry	11 (8)
Conservation and restoration	2 (1)
Time working in sector (years)	**N=141**
>15	**45 (32)**
5-10	39 (28)
11-15	29 (21)
1-4	28 (20)

Table 4.1. *MAIA survey participant demographics*

Institutions' size was measured via three indicators: visitor numbers (*Figure 4.1a*), annual budgets (*Figure 4.1b*), and number of employees (see *Figure 4.1c*).

[6] Percentages are rounded (.1-.4 rounded down, .5-.9 rounded up).

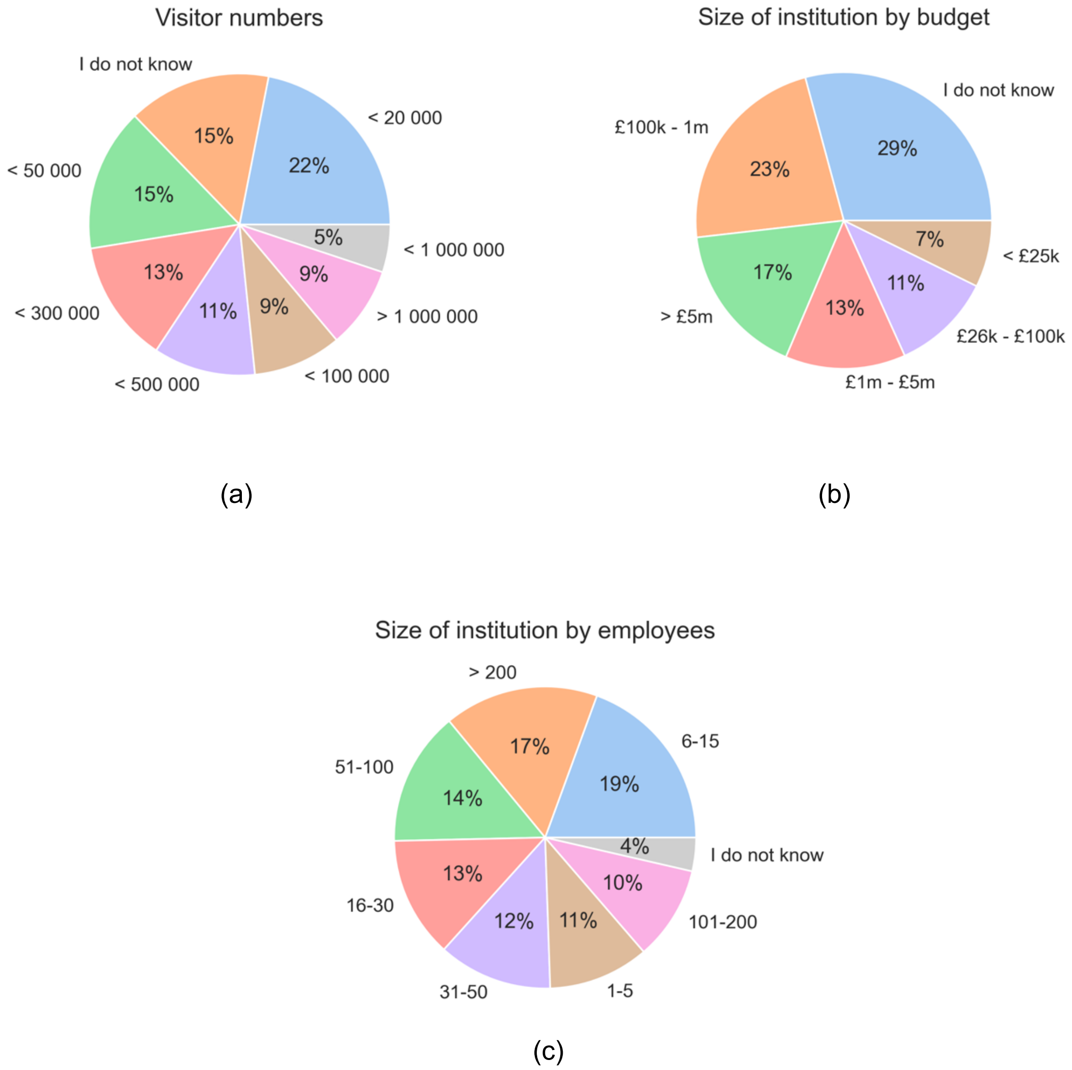


Figure 4.1. *Institution size*

The majority of institutions does not use AI

Asked if professionals are currently using AI technologies at their work (*Figure 4.2.*), a majority of 71%[7] stated that they are not, whilst 18% remarked that they are currently using at least one application of AI. Just about 4% have used AI in

[7] Percentages are rounded for legibility (.1-.4 is rounded down, .5-.9 up).

the past, and about 6% intended to use AI in the next twelve months following this survey.

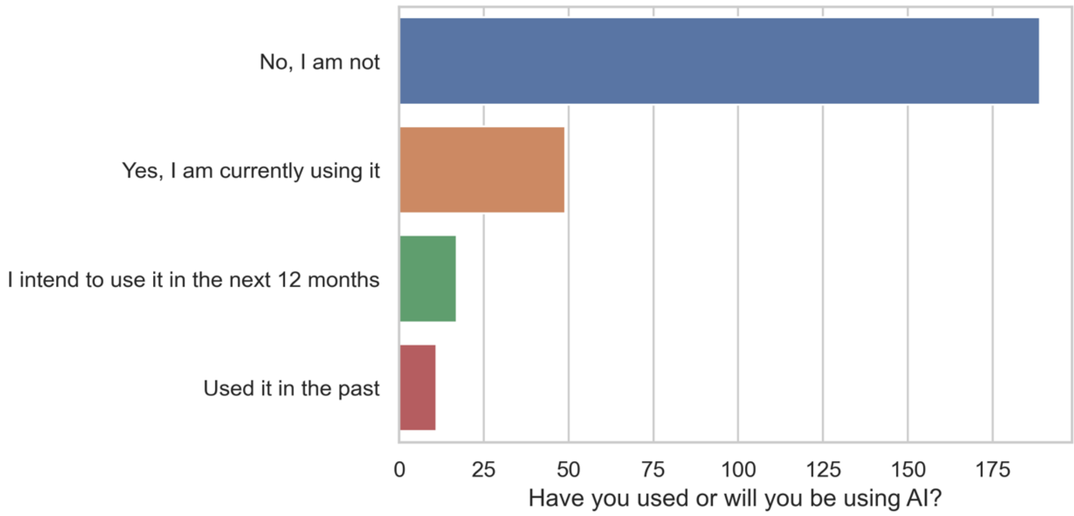


Figure 4.2. *Use of AI technologies*

Of those who have not used AI so far[8], 100 participants stated that they lack the skills or training to use it, whilst 61 said that their institution has not got the technology or that AI technologies are too expensive to deploy (60). A further 55 found that there is no need for AI applications at their museum, followed by 45 who indicated that there is a lack of staff to conduct data-intensive methods. 17 made use of the free text entry under *other* reasons, stating they lack knowledge about the technologies and the potential benefits of using it or issues around ethics and bias in AI. Two participants submitted longer explanations compared to their fellow respondents. One highlighted that there is an institutional ignorance around AI and missing knowledge:

> *"In the State Pushkin Museum of Fine Arts where I used to work for the last 7 years, the topic of AI never was taken up for discussion. I doubt that our IT department of the museum possesses the necessary skills."*

[8] Participants were able to submit multiple answers, the results are therefore reflected as raw counts and not percentages. This will be carried on throughout the survey wherever there were multiple answers allowed.

The other stated a major impact factor being technologies not rolled out yet and too expensive to apply:

> *"In our part of Africa the technology is still in developing stages, which ends up being too expensive for the institution."*

To gather further information not just about the individual responding to the survey, but also about colleagues and possible external partners, the survey also asked participants if they know about someone else using AI applications at their institution. 44% replied that there is no one else working with AI, followed by 33% who were not aware of anyone using it. 10% reported that some of their colleagues are currently using AI applications, whilst 9% stated there are external partners deploying such technologies at the institutions they are working at, and just about 4% knew about cases, where internal and external persons were working with AI at the time of the survey.

If participants selected *no* as their answer, they were further prompted to indicate possible reasons why neither colleagues nor external partners were working with AI at their institutions. They drew a similar picture as for their personal reasons, but judgement about others was more moderate compared to their own. 54 stated that their colleagues or partners might lack the necessary skills or training to use AI, followed by such technologies simply being not available (43) or too expensive (40). A further 33 highlighted a lack of staff and that there is no need for them (28), with 13 noting *other* reasons, such as their institutions being too small, a lack of understanding or that AI is not something that they feel their visitors are missing.

After having elicited information about the general AI usage, participants who indicated that *they are not using AI*, were asked about their intent to use it in the future. 46% were not really sure if they are interested in using it, whilst 39% were aiming to use it in the future, just over 9% issued a clear *no,* followed by 6% who did not know an answer to this question.

AI is mostly used on a project basis and applications were generally successful

Participants who indicated that they or someone else *have used* AI technologies (*Figure 4.3.*), stated that 45% of the applications were used on a project basis, followed by 31% using it in long-term integrations and about 17% just trialling it. The remainder did not know about the duration of usage.

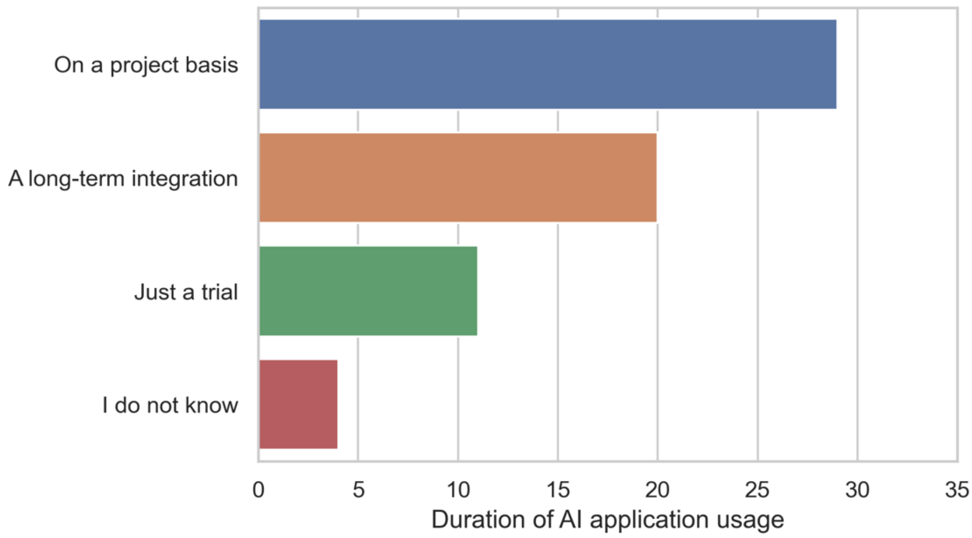


Figure 4.3. *Duration of AI application usage*

Professionals further stated that for however long they have used or experienced someone else using AI, that in 72% of the time the application was successful, 23% did not know about the actual success, and just 6% said that the application was not successful. 58% of the participants who said that the application was successful were also still using it at the time of the survey, and about 21% each stated that they stopped using it or are using it sometimes. Generally, participants who have worked with AI described their experience mostly as positive. Feedback included AI contributing to the creation of new knowledge and the possibility to see collections under a new light:

> *"Very positive. Using ML to infer features allows us to infer features about works in the collection and similarities between them make it more connected, accessible, and interpretable as a whole. With a good technical foundation/platform, it's as easy to integrate as any other new technology."*

> *"It's not a silver bullet, but when it works properly, it can seem like magic."*

The general positive tenor was often interspersed with some caveats about the applications and, especially, the interpretation of the output results, highlighting the often underestimated skill requirements to clean and process data ahead of feeding to a pipeline or challenging learning curves of being able to interpret outputs:

> *"Positive with a caveat: it helped to do work which would otherwise be too much for people to do, but the results do need to be checked and interpreted. Logical, but it's important to keep in mind that AI is not a magic wand, it's a toolbelt in which each tool has flaws."*

Respondents further gave evidence that it enhanced visitor experiences and was welcomed by audiences:

> *"My experience using AI is super positive. I can plan routes according to visitor's profiles."*

> *"It helps to improve the visitors' experience by contextualizing the exhibition."*

Some other participants painted a more neutral picture about their experience with AI, stating that those applications come with positives and negatives. A thread through all of their responses was that very particular, often experimental, trials returned interesting results, but that a broader roll-out is still wishful thinking of the future. Ethical concerns and issues around bias seemed to be further indicators that made participants sway towards a neutral opinion, dampening a purely positive outlook:

> *"The experimental results are exciting, but the challenges of executing any of this work on a large scale are massive and frustrating."*

> *"Neutral. It has been very successful for facial recognition, particularly of our donors. Its assigning of gender however is very problematic. It is also not good at assigning tags to contemporary art, but for a lay person searching it works fine."*

Very few participants (4) commented negatively on their experience with AI technologies, mainly based on disappointing outcomes:

> *"The outcome was more interesting for the technical researchers than for the museum."*

> *"Negative, image recognition did a poor job of suggesting relevant keywords for non literal and non figurative artworks."*

Working on AI projects collaboratively works well

Participants who worked on an application with a partner were given the option of free text entries about positives and negatives of their collaborations. 27 participants reported positive outcomes, such as tapping into new knowledge and the prompt by AI to come up with new and novel research. They also highlighted the knowledge exchange and importance of teamwork between technical experts and museum professionals to unite several skillsets to successfully deploy AI applications. Beyond research-driven objectives, participants also remarked on the usefulness of AI to foster creative projects and collaborations with artists, which has produced “beautiful and very interesting results”.

Just three participants reported negative experiences, mainly around costs of using AI and an ongoing financial justification to leadership teams or being roped into licence fees by external vendors at high costs without any own training capabilities of the models.

The survey further gave participants who remarked that the application was not successful, the option to explain the reasons why they stopped using it. Reasons given were that the end of the project term was reached or that software owners introduced charges that the institutions could not meet. In other cases, the employee working with AI left the institution or they stopped using it as it did not yield any interesting results yet.

Museums use various applications and computer vision is at the forefront

Participants who stated that they were already using AI or have the intention of doing so in the future, indicated a broad spectrum of AI technologies. However, by far the most mentioned applications were computer vision and pattern recognition techniques, including image recognition and classification tasks, as well as movement analysis, facial recognition, optical character recognition (OCR), and sentiment analysis. This was followed by applications around visitor metrics, such as visitor tracking and capacity calculations, and the usage of AI to enhance exhibitions and displays. Respondents also stated that they are using NLP to explore their collection data, e.g., to determine coordinate locations for collections having only descriptive locations together with geographical information systems (GIS) or speech to text to transcribe media for exhibitions.

Not as prevalent as the above mentioned techniques, respondents also deployed clustering and similarity calculation methods, AI to support search, or for transcription and text mining purposes. Lesser used applications included AI for visualisations, AI aided investigations of collections to support decolonisation and bias reduction efforts, entity recognition or read aloud to enhance accessibility of online provisions.

Institutions' digital capacity is not future-proof

Apart from inquiring about AI technologies directly, MAIA also asked participants about the general digital set-up of the institutions they are working at. Ranging from databases and ontologies to software and technological practices and policies.

Museum staff were asked if they think that the museum or gallery that they are currently working at has sufficient digital capacity to tackle the challenges of the future (*Figure 4.4.*). The responses paint a dark picture as just about 6% stated that their institutions are future-proof in terms of digital. 44% said their institutions are not ready to tackle future challenges, followed by 35% who claimed that their workplaces are partially ready, but their systems and set-up need major updates. 11% answered that they are mostly ready, just needing minor updates and 4% felt like not being able to answer this question sufficiently.

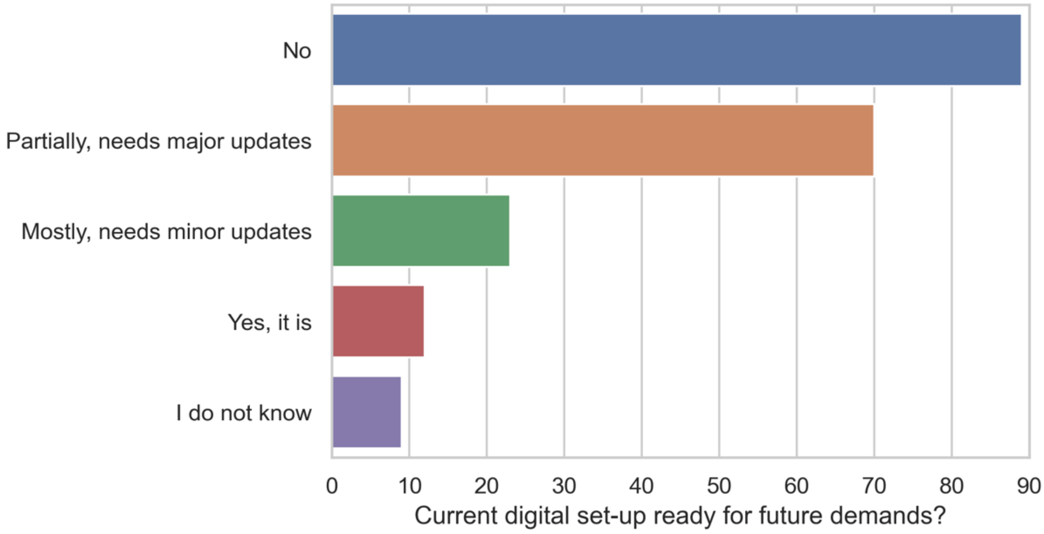


Figure 4.4. *Digital set-up of institutions*

From spreadsheets to dedicated tailor-made software

The software used to fulfil professionals' workflows are diverse and span a trajectory from basic applications, such Excel spreadsheets and Google docs to custom-made software specifically designed to suit the needs of institutions. Most participants reported the use of CMS software provided by vendors, such as Axiell, TMS, mimsy and adlib, followed by products that were not specifically developed for museums, but support data practices at institutions. Amongst others, these are cloud services (e.g., IBM Cloud, Microsoft Azure, and Amazon Web Services), data visualisation products, such as Power BI and specific Python libraries, and custom-made systems as reported by participants from Turkey, who are using MUES, a National museum system, which has integrated AI and Deep Learning features or RUSSOFT's KAMIS, which was specifically developed to offer a unified system to 870 museums based in the Russian Federation. A small remainder of participants stated that they are either using no specific software to fulfil their job roles, but rather basic products offering text editing and calculation and computing features (Microsoft Office, Google Suite) or no software at all.

When asked about if the software provided by their institutions are suitable to fulfil their job roles, 37% answered that they are suitable, however, 26% stated that they are suitable to do their jobs but there are shortfalls and things to improve, or that they are not suitable to fulfil their daily duties (16%).

Participants who answered that the software used is partially suitable were able to further expand on the issues and, once again, high costs were often a barrier, especially around licensing and cloud-based computing. Another theme resonating amongst participants was software not being able to be customisable, often not reflecting the actual needs of the professionals, interfaces not being user-friendly, often being out of date, and the system not suitable for the data held about the collections or not enabling any beyond standard operations. Other participants reported that:

> *"There are shortcomings in nearly all major collection / library / archive management systems and DAMS."*

> *"There is no integration between AI and collections management or digital asset management software, and vendors are reluctant to implement without support."*

> *"Search functions are limited and quirky. Producing tailored outputs is difficult as we are limited to what the application allows. Bespoke outputs can be designed for us, but they have a cost that is beyond our means."*

This shows how deep rooted the issues around museum software are and that integration of contemporary or data-intensive methods is often not feasible due to current institutional provisions. Over 40% of professionals state that there are at least improvements to make which paints a troubled picture ahead as funding is unlikely to be raised concerning the current environment.

Professionals can generally work with the databases they are currently using

Databases holding collection information were also focus of the survey, and 31% replied that the databases they are working with are somehow suitable to represent the data they are working with, but databases would need improvement, followed by 31% who were happy with databases they are using. 26% felt like they are not in a position to answer this question, whilst 13% stated that their databases are not suitable at all. It is evident that some participants used this question to include their general sentiments about the systems they are using.

Participants who found that the databases/systems were either not suitable at all or needed improvement were given the option to comment on their selected answer. Participants generally missed the options to integrate national and international data into their CMS, as well as other systems. They further missed support for cross-domain materials. A lot of participants highlighted that the technologies they are using are not flexible, not user-friendly, often too basic and highly reliant on vendors:

> *"Brittle architecture and simplistic understanding of knowledge makes our tools outdated and reinforce certain cultural hierarchies."*

> *"Programs like TMS are too highly structured and end up encouraging work-arounds that make interoperability almost impossible."*

> *"Bare bones collection information system, staff find it hard to find information and there are no links to public facing websites etc."*

"I think a move to graph-based, object-oriented data storage and away from relational is the future. Also, I strongly believe that narrative text is a better knowledge dissemination format than data."

Ontologies and other forms of knowledge representation have not really gained traction so far

As museum collections represent a wealth of knowledge and information, the survey also asked participants if they are using any formal forms of knowledge representation, such as ontologies (e.g., CIDOC-CRM). 67% replied that they are not using any ontologies or other forms of knowledge representation, whilst 33% said they do. Participants who were not using it remarked that this is mainly down to not knowing about their existence or usefulness, or that to use them would require staff and training as their current data would need to be cleaned and/or migrated to different platforms.

Missing education about AI and high costs are the top contributors negatively impacting AI

The survey asked for participants' opinions around what might negatively impact AI technologies and the perception of them in general (*Figure 4.5.*). Missing education and high costs were rated as the two top reasons having a *very high impact,* this was followed by a lack of explainable/interpretable technologies, and again high costs with *noticeable* impacts. Issues around bias, AI's effects on society, and the environment were the three top mentions attributed with a *modest* impact. Participants who indicated that they are *not at all* impacting AI negatively led with issues around the environment followed by *other* reasons mainly around staffing, ethics, and dystopian futures:

"Lack of specialized staff within Museums and necessity to rely on external suppliers for maintenance."

“Not having the staff to support and introduce it, embed it and make it part of the Museum’s future plan.”

“The potential of destroying civilization.”

“Exploitation of invisible labor.”

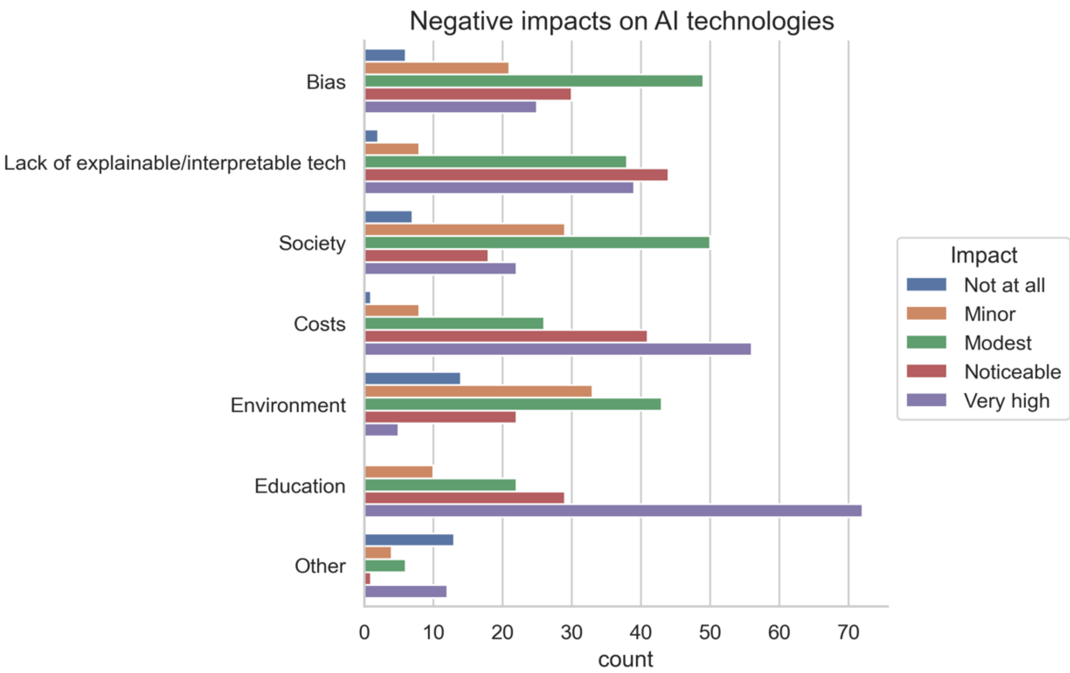


Figure 4.5. *Negative impacts on AI uptake*

The uptake of AI in museums is mainly hindered by a lack of funding, skills shortages and missing technologies

After having asked participants about their general perception of what might negatively affect AI in general, the survey inquired about the impact of certain factors contributing to the uptake of AI in museums (*Figure 4.6.).*

Clearly leading were issues around funding, skills shortages, and missing technologies in museums, followed by a lack of strategies for AI, which all have a *very high* impact. Missing technologies was further identified as the leading factor having a *noticeable* impact on the uptake of AI, as well as a lack of explainable technologies and the organisational culture not being the right one. The latter was also named as the leading *modest* factor, which was followed by a lack of strategy, society not being ready for it, or that there's no need for AI. *Other* impact factors mentioned by participants included "the establishment of AI standard technologies for museums", a "lack of imagination about what it could achieve", the need of "leadership to move museums to new directions", or "becoming ever-reliant on technologies".

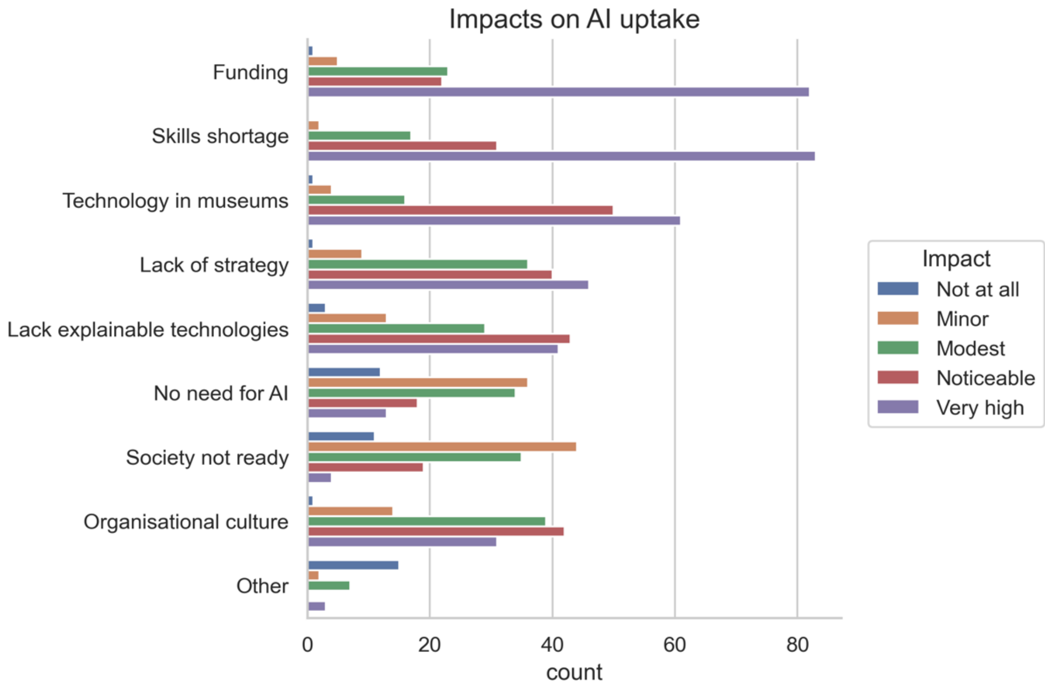


Figure 4.6. *Impacts on AI uptake*

Plug-and-play products and data sharing partnerships could drive AI in museums forward

Professionals were asked what variables would favour an AI uptake in their institutions (*Figure 4.7.*).

Plug-and-play products were mentioned as having the highest impact on contributing towards an AI uptake, closely followed by the establishment of data sharing partnerships and OA software. The implementation of data standards and government and regulatory bodies especially focussed on the heritage sector. *Other* variables stated by participants having an impact were "young professionals aiming to bring their skills to museums", "new dedicated professions in the Museum sector", "training for museum professionals to understand data and its usage (standards alone are not enough)" and "real open science".

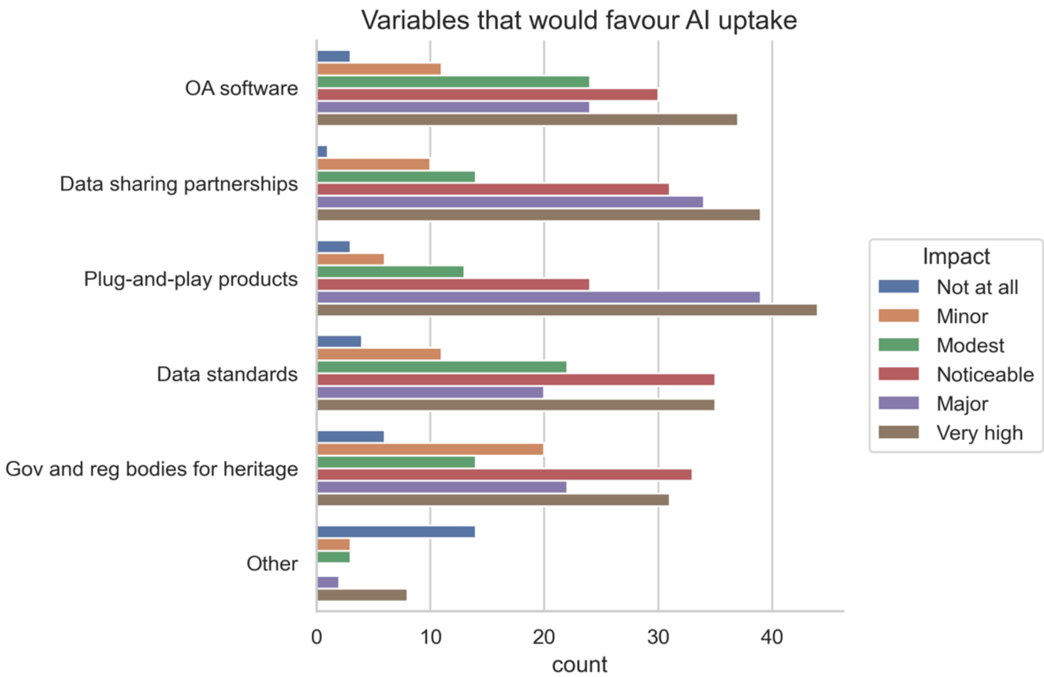


Figure 4.7. *Variables favouring AI uptake*

Leadership teams and those working with technologies on a daily basis are not on the same page

Participants were asked if they think that there is a dissonance between senior leadership teams and those who actually use the technologies on a daily basis. The answers revealed that in most cases (33%) SLT have a different perception of how technologies are used, whilst 26% stated that there sometimes is dissonance, depending on the application. Just about 12% thought that SLT are aligned with professionals who use technologies and that there is no dissonance in how those technologies are perceived. The remaining 26% did not see themselves in a position to answer this question satisfyingly.

Dedicated AI and data science roles are a rarity

To find out if institutions were employing staff specifically dedicated to AI or data science, the survey asked if institutions were employing a data scientist and/or an AI engineer/machine learning specialist. The range of possible answers slightly deviates between these two questions, as the field of DS is generally seen as much broader compared to the remit of purely AI or ML focused roles.

Over three quarters of the respondents (78%) reported that their institution is not employing dedicated DS staff, followed by 12% who stated that their workplace employs at least 1 to 5 people specifically holding DS positions. Just about 1% of the participants said that their museum employs 6 or more data scientists, and the same percentage reported that they used to employ one. The remainder felt not able to answer this question.

Asked about AI/ML specific roles, the lack of such dedicated staff was even more significant compared to DS roles, with 88% responding that there are no AI/ML roles at their institution. About 5% stated that there are at least between 1 and 5 dedicated staff at their workplace, with the remaining 6% not knowing an answer to this question.

The museum as a place to educate about AI, and a place to offer AI enhanced visitor experiences on-site and online

Beyond a professional remit, MAIA asked participants about AI and a public usership. 70% of respondents agreed that museums are suitable places to educate the general public about AI (*Figure 4.8.*), with just 11% saying that this is not their institution’s remit. The remaining 18% felt like they were not in a capacity to answer this question.

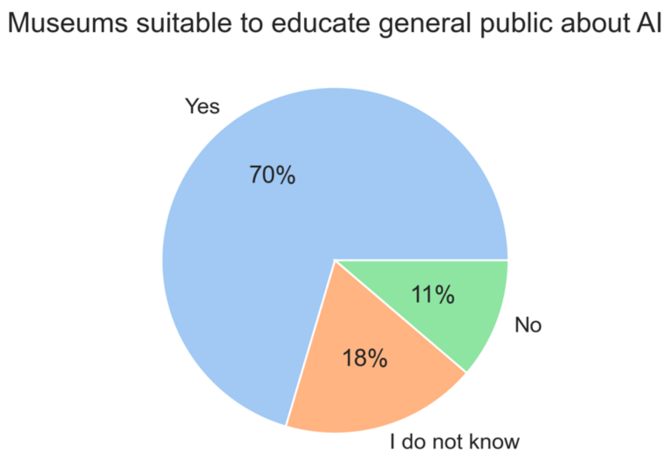


Figure 4.8. *Museums and AI education*

Major agreement across all respondents was found around the two themes of AI being able to enhance visitor experiences on-site as well as online.
Whilst 38% found that AI can enhance visitor experiences on-site (*Figure 4.9.*), another 38% stated that they might do in the future. Just about 2% of respondents answered with no, with the other 22% remarking that they cannot give an answer to this question.

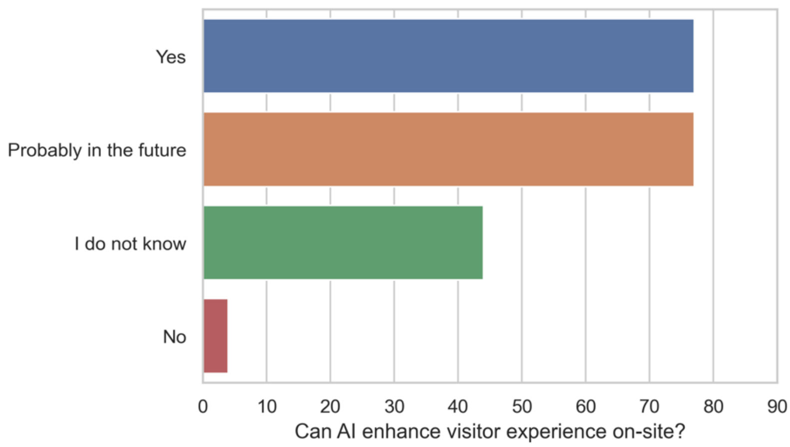


Figure 4.9. *AI and on-site visitor experience*

When inquired about visitor experiences online (*Figure 4.10.*), the picture was even clearer with a strong tendency towards positive answers. 59% felt that AI can enhance the VX online, and another 24% agreed that this might be the case in the future. Again, just about 2% answered with a no, the remainder selected that they do not know an answer to this question.

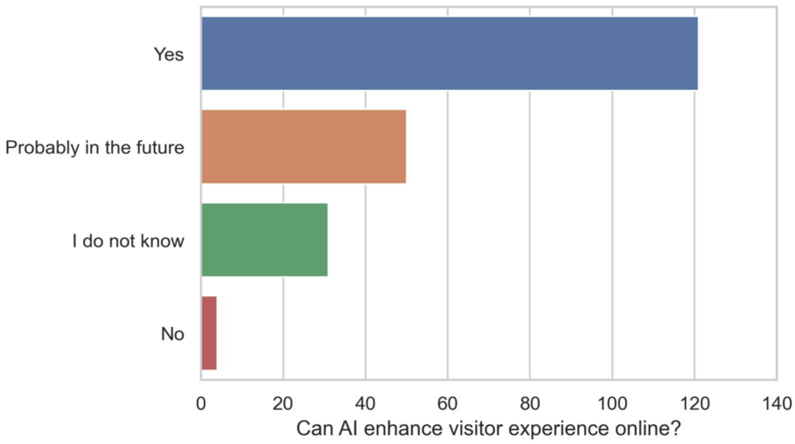


Figure 4.10. *AI and online visitor experience*

4.2. Discussion and summary

The results of the survey indicate that museums still have a long way ahead to fully enable the possible potential of AI applications across the sector. Whilst the literature review has highlighted that these technologies look promising to solve some of the 21st century's challenges museums will be facing and to contribute to new forms of visitor experiences, their roll out is slow, future uptake seems to be low, and institutions are facing barriers that are yet to be overcome. Highlighted as magic by some, but still a hidden craft for most.

Barriers are not the only problem affecting emerging technologies, but they seem to echo issues that the museum sectors in the UK and beyond have been facing now for decades. Generally tight budgets and a lack of funding and skilled staff contribute to a slower roll-out of AI technologies in the sector. Smaller museums, which have to prioritise their spending, may fall behind their bigger counterparts that tend to be the main users of AI and have the capital to attract prestigious project partners and the institutional infrastructure to accommodate for them appropriately.

This research shows that there are manifold reasons to the barriers of taking up AI, not only rooted in problems directly related to AI technologies, but also affecting the wider museum environments they are supposed to be operating in. The survey indicates that a lot of museums are not AI-ready yet and staff report that their current digital set-ups are often far from optimal, which means that the necessary infrastructure is not available and museums need to get the digital basics right first. There is evidence that databases used to store information and represent the data held might not always be suitable for most of the professionals' work or need some form of update and the software used often confines them to a limited space of operations, preventing staff from performing out of the box operations or deploying more in-depth methods. Thus, the museum environments that are able to adopt and adapt to AI technologies are yet to be established; getting the right provision in place to be even able to think of integrating AI is therefore a first pivotal step towards successful future long-term integrations.

The high profile media coverage around AI and its popular appeal may also give the impression that AI is already ubiquitous. Reality looks different though, and reveals itself as a double-edged sword for institutions, where on one hand, professionals are dubious about practices around AI and concerned with their dystopian connotations, and on the other there are staff who are eager to experiment and transition towards a new age of knowledge and meaning-making driven by AI.

Currently, computer vision and related methods seem to be at the forefront of deployed AI applications in museums. This might be due to digitisation of image-based collection data advancing at a much faster rate compared to metadata and descriptive information and whereas museums might be able to take a picture of every item in their holdings quite easily, they struggle with missing metadata to conduct text-based research meaningfully. The availability of computer vision software and cloud ware that allows rendering results rather quickly and conveniently could be a further contributor towards more accessibility for museums. Furthermore, computer vision seems to be aimed at tasks that replicate usually human-based tasks, but at much higher rates, e.g., tagging of images, classification, and clustering. Another reason, although to be consumed with a caveat, might be art museums skewing results due to their often iconographical focus and investigation of their collections and attractiveness to present constituents with visually stimulating and/or aesthetically pleasing outcomes explaining the popularity of artists like Refik Anadol or Trevor Paglen.

The survey suggests some potential opportunities to tackle the challenges ahead that could lead to a wider integration. First and foremost, collaborations tend to be key enablers of currently used AI applications in museums, allowing to combine various skills and knowledge to implement such technologies and create successful outputs. Fostering collaborative projects with reasonable outcomes for museums and engineers alike can therefore be a fruitful symbiosis. On top of that, results and products arising from those collaborations should be made open accessible and reusable by other institutions who do not have the capacity or network to establish such collaborations on their own. Easy to integrate plug-and-play systems would further contribute to a wider roll-out, reducing friction of staff having to scramble everything together just to end up with a patchwork system made from various software, hardware, and added legacy and interoperability issues.

Adding to those institutional barriers are personal thresholds of museum staff, ranging from not having enough knowledge around AI to not seeing any added value of using them either for their work or other constituents.

Professionals not seeing the value of AI technologies caused by a lack of education, knowledge and training seems to be widespread across the sector. A finding that is also resonating with results of the DASH survey, which found that heritage professionals generally engage with technologies if they “could offer tangible gains to them in role” (Newman et al., 2020, p. 6). This might also be related to staff having the feeling that there is no need for AI technologies, caused by a lack of understanding of their potential usefulness or that there literally is no need in applying them.

Issues around vendors are prevalent throughout and may be contributing to holding museums further back from gearing up their digital game. Licensing issues, fees, and products that are not versatile enough are putting a tight cast around what is possible and what is affordable for institutions, making factors around funding and staffing prevalent again. Whereas bigger institutions with the right budget and/or staff seem to be more agile in circumventing restrictions around vendors through the development of custom-made or in-house products, institutions without such a pouvoir seem to be less reactive in terms of digital.

To those external factors come internal struggles as museum staff have the feeling that SLT are often not on the same page with the wider institution. Those dissonances between people making strategic decisions and those that execute operations might have led to the procurement of systems that were not fit for purpose and missing guidance and interest around ramping up digital and future-proof strategies around it.

Museum professionals who have used AI paint a positive picture of the applications and their success adding a much needed positive momentum that promotes the usefulness and variety of use cases of such technologies, as they have been identified in the survey to be (future) contributors towards an enhanced visitor experience on-site and online and museums can be the right place not just to apply AI, but also to educate the public about ethical considerations and bias around it. Museums therefore can become thought-leaders in AI and cater a broad usership, fostering a savvy usership and societal value, whilst highlighting current and future problems of a technology that will become ever-more pervasive and ubiquitous. This can establish the museum as a thought leader around trust, authorship, and authenticity through the creation, presentation, discussion, and mediation of AI applications in museum environments and their constituents, which will be further discussed in chapter seven.

This research makes evident that the future requires museums to work together and pool their resources, skills, and knowledge to face challenges together. Projects, such as *Towards a National Collection* or *Heritage Connector*[9], are important to not just start a discussion, but also test the ground of how such National integrations can contribute to data sharing and a more equal access to computational methods and data for all institutions across the UK. This can be further nurtured by institutions who have the means to conduct computationally-intensive research in making their code available to others and engage in Open Research endeavours to build constructive grassroot AI movements for museums of all collections and sizes.

[9] For more information about the projects please visits: https://www.nationalcollection.org.uk/ and https://www.sciencemuseumgroup.org.uk/project/heritage-connector/

5. Museum Professionals Focus Groups

This chapter presents the results of the focus groups held with museum professionals across four institutions - the Manchester Museum Partnership, the National Gallery in London, the Smithsonian in the US, and the Badisches Landesmuseum in Karlsruhe (Germany).

To briefly reiterate, the main aims of the focus groups were to demonstrate the system to professionals and discuss possible use cases, considerations around it, and development needs. Whilst the RS itself was scrutinised, it further catered as an opportunity for reflection of AI-driven technologies in general, to open up the field and stimulate wider discussions around data practices at the institutions where the groups were held. The analysis is split thematically into three sections: 5.1. *Museum environments*, 5.2. *Constituents*, and 5.3. *Technology*, to reflect the human-technology-world relationship as outlined in the postphenomenological framework and give some structure to the themes that emerged during coding the transcripts. As the focus groups were semi-structured, some themes were discussed in relation to each other and therefore a clear separation of themes is not always possible. Museum professionals were asked to participate in their professional role, contributing their knowledge and experiences of working in the sector. Whilst it should be acknowledged that their different institutional contexts, practices, and histories will have a relevance to their accounts, they were not adopted as the defining lens through which to analyse their contributions. It is important to note that some participants explicitly highlighted that their participation reflects their own opinion and is not representative of the institution they are working at. Thus, where appropriate, institutional context is given throughout this chapter where deemed useful and clarifying.

5.1. Museum environments

This section presents the results of the focus groups that mainly concern museums and their environments. As introduced in chapter 1, these are themes that directly or indirectly shape the fabric of the institutions, their practices and the landscapes they are operating in as well as the materialities they constitute or are subjected to.

Access and barriers

Access and availability were highlighted by participants to cover several different points that share a common vocabulary. Professionals remarked that access to data and technologies can be an issue, touching upon various barriers in accessing and sharing collection data as well as relevant software and hardware. Some of the systems used to manage collections are like “walled gardens” (P4) and “don’t allow for novel or playful ways of engagement with collection data, such as cooperations with Wikimedia, which would massively widen access to collections” (P4). Participants, depending on the institutional set-up and legacy agreements, stated that it is sometimes very difficult to share data with other constituents - although owned by the institutions themselves - yet let them interact with it. This makes it hard to even achieve the first step of acquiring data that can be used for the RS, having to jump several hurdles before being able to export datasets for external, but also internal purposes (see *Portfolio*, p. 8).

Such issues are not just affecting practices on an institutional level but can also lead to state-wide or national endeavours aiming to unify collection systems failing due to animosities and preferences of decision makers. P25 described the struggle with their documentation software, which was initially acquired by the State of Baden and planned to be rolled out across all State museums, including the Badisches, but some institutions showed reluctance from leadership teams right from the beginning to incorporate a common approach, exercising their power over collection systems, which led to legacy issues and non-interoperable system across museums (P25).

Apart from the above mentioned, mostly internal, barriers of accessing data and institutional dissonances, data access issues were further aggravated through licensing regulations and a lack of open access to collection data.

Whereas the Smithsonian offers most of its digitised material under a Creative Commons Zero (CC0) licence, meaning that there are no rights reserved (Creative Commons, 2022), it emerged as a predominant problem at the European institutions who took part in the studies. Although participants were in agreement that OA would contribute to a heightened engagement with collections, in which way though was seen discordantly. P5 stated that even making low-quality collection datasets available can have surprise effects. They described the experience during a hackathon when coders and artists at the event were looking at the gaps rather than the data itself and created some beautiful outcomes, which was also observed by P25 who remarked that once people realise that they can use the content, they put their own creativity to work and create something out of it. Other participants, however, had different experiences when running hackathons or with digitisation programmes in general, stating that just throwing data out there in CC0 is not enough, claiming that museums need to provide data which people can actually do something useful with and this requires sensitivity to the content and careful consideration and insight of professionals (P29).

Participants also highlighted that museums who want to offer their data to the public, let them interact with it, and see it put to use, need to actively raise awareness about their offer, after users have spent decades with museum content where they were told not to photograph, not to touch, and to be as passive as possible in their interactions. Highlighting actively how digital content can be used is a necessity as a lot of users do not understand what the different licensing codes at the bottom of the image or the webpage mean, let alone know that using museum content is even possible (P29). This makes the case to change the mindsets of users guiding them towards playful and enjoyable experiences with museum data. For example, Smithsonian's OA Remix initiative, which actively invites users to “Create. Imagine. Discover.” (Smithsonian Institution, n.d.) the institution's data and be subsequently featured on one of their museums' websites.

Similar lessons were learned at the Rijksmuseum as P26 highlighted that, whilst having worked at the museum in Amsterdam, they soon realised that licensing did not yield great results in terms commercialisation and putting the collection online as CC0 returned more valuable outcomes than paying staff dealing with licences just because there were “like three requests a year” (P26) of actual paid licensing. Participants state that it makes sense to keep certain data, such as personal details of donors or collectors, redacted, but museums need to cater to the fact that most of their collections are effectively owned by the public and “[m]aking this data available is actually our duty and it needs to be accessible externally” (P26). With data in the public domain, risks of misuse, bias, and (un)intended harm increase proportionally, especially with today’s social media channels where viral content can hardly be stopped or mitigated for. This was also one of the participants’ major concerns, which is evolved further in the next section.

Bias, explainability and sensitive data

Bias was one of the participants' major concerns with regards to the RS using their collection data, but also in terms of how these systems can either reinforce or help to reduce various forms of it. Collections were seen as inherently biased through the way they were formed and how objects and information were aggregated. Those biases might be known but omitted, known and accepted or unknown and unquestioned. Bias in collection data can range from the misrepresentation of gender, biases formed at time of collection (non-)entry to colonial bias or may be rooted in personal preferences of people having access and authority over the collections and processes such as de/accessioning, documentation, and selection of displays, to illustrate a few (see *Portfolio*, pp. 8 and 16).

P1 addressed issues around general bias, such as geographic or ethnic bias, but mentioned further the issue around an “extended bias, caused by how we mediate ourselves” (P1). When asked to evolve further about those extended biases, P1 stated that there exist the various forms of bias that are already in the very fabric of the data, inherent to the data so to say, but that there are further forms of extended bias that are internal to every person interacting with collections and such biases can, depending on the collection data be reinforced or weakened, e.g., the reinforcement of presumptions or the push to positively rethink them based on the experienced collection content. The RS was seen as a possible means to lay these bias issues bare and address them. This can happen algorithmically, by engineering the system to mirror or even reinforce the bias in the system and therefore show to users what patterns of bias are inherent in collections. This approach would mean that rather than trying to hide bias in the collection and obscure it, magnifying and highlighting the issues could actually help to articulate and address them. A “bias correction filter button” (P1) for the recommendations served. P6 remarked that the algorithm, in common parlance, is quite often seen as “the villain that keeps us in our bubbles” (P6) issuing concerns around the manipulation of algorithms and how they can be put to use to actually cause harm or reflect unwanted views, with the recommender adding a new level to common discussions around algorithmic bias as participants agreed that the RS probably has the power of cognitive manipulation, tricking users into a false sense of subjectivity or into establishing relations between information that might not be reflecting a true account of events.

These concerns led to further questions about racism and how algorithms are trained. P21 inquired about avoiding falling into the “white dude symptom AI trap” (P21), based on the fact that their institution holds a lot of portraits of predominantly older white men and if the algorithm would reinforce that representation or if it would be able to reflect BIPOC and actually serve useful connections and recommendations that have the power to equalise misrepresentations and that could contribute to a more balanced approach away from the White centrism of many Western institutions.

Discussions about racism further led P22 to think about how gender is modelled in the framework they are currently establishing for the Women’s History Museum, part of the Smithsonian Institution, around how gender is represented in collections in relation to OA endeavours and how new technologies would work with it. They then reported a recent incident, where someone developed a bot for the National Portrait Gallery’s collection, which tweeted a very racist image and it made staff at the Smithsonian realise that there needs to be some form of mediation or control. The post brought several issues around cultural sensitivity to light as, although operating as one institution, it was evident that, for example, art museums work differently in terms of cultural considerations than the NMAAHC or NMAH, making it difficult to address issues around the removal of data (P22).

Participants agreed that it is thanks to OA programmes and systems such as the recommender, that institutions have to address how they are dealing with sensitive data that was never thought to be in the public domain. They further highlighted that it probably was wrong to just exclude such data in the past from public displays as making such data accessible with the right communication around cultural values can help marginalised groups to be heard and seen. Thus, institutions need to find ways to work with such data they are stewards of in the digital age. A possible solution was seen in curatorial oversight, especially important when dealing with sensitive data, as participants doubted that algorithms will have enough knowledge or can be trained sufficiently to avoid possibly hurtful displays or the establishment of connections which are neither explorative nor reflecting true events. A balance between the human and the machine was suggested as a remedy to still make such data accessible whilst understanding that “there are these constituent and cultural and audience sensitivities around that need to be explored” (P19).

Participants also highlighted the need for better explanation of the models used to serve recommendations, especially when they are based on machine learning algorithms that are colloquially known as black box algorithms, causing a “tension between sort of more curated, or the more explicit and understandable form of recommendation, which is very much based on curated metadata, and could almost be hardcoded and the thing that is AI or machine learning based where effectively it is a black box and you actually don’t know why it’s making these connections” (P7).

It is evident that there is an emerging notion of generally assuming that AI or machine learning algorithms lack general explainability, whereas by far not all algorithms are black boxes and there are systems that are well explainable working with clearly defined and transparent algorithms. However, just like black boxes, this also needs to be explained to users and cannot be assumed as common knowledge of constituents.

The above issues suggest the need for greater transparency. P4 suggested to have the system blogging about its current status, what data it is gathering, and what the outputs are based on, would, for example, help users understand what is happening in the backend. A bot-like account that updates regularly, but that people can also interact with in terms of infiltrating the algorithm and feeding the system with content to explore their own interests and see how the algorithm reacts to it:

> *"I think it's a great opportunity to increase transparency and improve ethical frameworks and understanding. I love the idea of a public output on any platform, but then it's also good to be conscious that who sees that is down to, e.g., Instagram's algorithms. So we've got the Russian doll of algorithms working with us sometimes."* (P4)

Having access to a lot of collection data also means addressing issues around sensitive metadata and images. Those issues are plenty and need careful consideration and weighing up between fulfilling missions to share collections and impart knowledge, and protecting vulnerable groups, avoiding the use of objects or information to spread hate and abuse.
The Smithsonian went through an exhaustive review programme before making data available as OA and the standards developed for this process were established through careful consideration and expert insight. This would need to be built into algorithms and their training as well to avoid unwanted scenarios:

> *"I mean, we have lynching photographs, but we also have photographs of the KKK and you would never want those displayed in a way together without explanation or curatorial insight. [...] We display and share them because we have to tell the unvarnished truth, that's what we do. And so we just would want to make sure that those sensitivities are built in so that users can still access them."* (P19)

Besides bias inherent in collections, discussions also arose regarding the completeness of datasets and the pitfalls or opportunities around putting *imperfect* data online. P28 questioned their practices at the Badisches of feeding just perfect data into their systems and suggested the RS or other suitable technologies could actually be used to enrich the data - fact checked and free from any abusive or harmful content - and feed it back to the museum. This would take away pressure from curators who are currently operating the bottlenecks of data decisions and transfer it to automated systems as the crowdsourcing hype failed to materialise for the museum. Both of the partaking curators at the Badisches added to this that, admittedly, curators are also not "Fachwissenschaftler für sämtliche Sachen, die hier im Museum stehen" ["experts of all things at the museum"] (P25) and that there are gaps in documentation and cataloguing of data that was entered to comply with certain regulations and agreements, but which is often outdated and impractical. This has led to enormous workloads that were not accounted for since the database was never supposed to be made publicly available, so staff were required to censor and remove information, such as details about private donors, insurance values, or sensitive internal data. Museums having the courage to just leave gaps in database entries and, for example, mark data as missing rather than removing the whole entry, would have allowed for a more organic approach from the beginning of electronic records (P25).

The NG issued concerns around conservation and restoration data of paintings. The reasons why such data is not being shared were not disclosed, except from one participant who stated that "they [conservators] are very worried about who has access to things like paintings during treatment. And while they are happy to share with their own colleagues and other institutions and scientists, the idea that the general public could tumble upon possibly sensitive images would mean that they would say absolutely can't be included until you give us the reassurance that they can't be found by the wrong people" (P8). These "wrong people" are people not able to interpret images taken during conservation of paintings when even slight damages can look rather shocking to an untrained eye, leading to fallacies about care and treatment of objects in custody of institutions (P8).

Thinking about the above prompted participant P26 to think about their time working at the Technisches Museum in Vienna, which decided to put incomplete datasets online as their main aim was to make the information, even if incomplete, accessible. They had positive experiences in doing so:

> *“Und das coole fand ich halt immer, dass da Forscher*Innen kommen und sagen “ah, ich arbeite im Moment am Pferdesattel und ihr habt ja da Pferdesattel-Modelle” und die kommen dann und wir mussten mit denen dahin und das Anschauen. Und ich habe dann was gelernt über unsere Datensätze und das geht nur dadurch, dass das eben schon online ist und gefunden wird. [...] Wenn es da ein recommender system gibt [...] und Leute dazu ermutigt werden was zu tun, glaube ich, dass da die Interaktion mit unserer ganzen Sammlung, das Interesse stärker wird und aber auch zu sagen, dass der digitale Katalog nicht abgeschlossen ist, sondern Leute können damit arbeiten.” [And I always thought it’s cool that there are researchers visiting, saying “ah, I’m currently investigating horse saddles and you have got some models” and they came in and we had to show them around and it was actually us who learned something about our collection data. This was just possible because we put it online and made it findable. If there would be a recommender system and people would be encouraged to do something with it, I believe, that interaction with our collection and interest in it would be strengthened, but we also need to say that the digital catalogue is a work in progress, but people can work with it].* (P26)

Whereas the above sections consider how data is reflected in the database as it stands, further consideration was given to the data practices in general and whether the ways of reflecting and using data are appropriate.

There might be the need to think of data more from afar, not just from the narrow perspective of the museum, but more in terms of generally advancing the nature of data and its use. Adding geo-data to collection data was one of the proposed ideas to track local history over time, being able to research how places change, or let people add ancestral stories to collections and bringing the city to life through a mix of collection data and data of the environment surrounding it (P1). Breaking out of the confinement of using just the collection database was seen as a possible means to reference data points lying outside of the collection, adding other feeds into the RS to create playful content that is more likely to be shared than object-related data. Such integrations could be social media feeds, but also news feeds cross-referencing other forms of data (P4).

P3 however highlighted that most CMS, in this case EMu,[10] are not made to include such experimental data and that they are still “at the challenge of getting basic data into EMu and after all these years, we are still inputting basic data just to make our normal search work” (P3). This prompted further discussion around what really constitutes *basic* data and when talking about collection data, *basic* data was seen as always defined in terms of Spectrum standards and museological point of views, which constitute basic data by fields such as accession numbers and the current location of objects. However, basic data might look different in terms of usability and therefore institutions should reflect about what is deemed essential information and what is desirable information - “data perhaps needs a bit more nuance” (P5).

Future proofing the museum

Participants were also keen on thinking about the time ahead, envisioning what museum futures could look like, and what forms engagement might take. They thought of the enhancement of museum online profiles through multi-museum online offerings, that aggregate content of several institutions, which might attract more clicks. This can help museums, especially smaller and regional ones, to rank higher in search engines, making them easier to find. Another possibility was seen in seamlessly integrating social media channels rather than just hosting profiles on several platforms as this fosters marketing and outreach, but hardly attracts people to online collections or catalogues. Participants felt that a lot of the things that attract audiences happen on social media platforms, do not actually trickle through to the online collections or are not making enough use of

[10] A CMS by Axiell.

them, building a wall with all the socials on one, and the museum website or dedicated online collection offerings on the other.

Whilst most participants appreciated the possibilities that emerging technologies might and are already bringing to museums, some highlighted that the emergence and roll-out of new technologies also causes a conundrum. Whilst more people are digitally connected and *online* than ever before, the Smithsonian identified that the demand for low-tech products - where a device is used to access the material, but it does not necessarily require a high bandwidth or content to be printer-friendly - was increasing at the same time. One participant highlighted that some people have either no device at all to access museums online or not enough bandwidth to use content involving significant data. This prompted the Smithsonian to think of ways to reach people *beyond device*, handing out educational materials at school lunch drop-off locations where families could still come to get food although schools were closed or to print collection-based information for children on the inside of Amazon delivery cardboard boxes (P18). Museums will have to think of how to bridge the digital, but also the data divide.

Having explored museum professionals' discussions around the museum environment, its practices, and related themes that emerged from the analysis, the next section presents their accounts with a focus on the constituents interacting with the environments and the technologies.

5.2. Constituents

This section elicits participants' opinions around constituents and various userships. It touches upon various audiences and different forms of engagement, and considers the synergies and tensions between objects, meaningful content, and constituents themselves.

Audience landscape

Participants eagerly shared their thoughts about the various types and profiles of users and the related motivations to interact with the system. Professionals agreed that technologies, further reinforced through the COVID-19 pandemic, constitute a space which challenges the idea that simply opening the doors of the museum is enough and institutions will have to accommodate to this changing situation. Nonetheless, the emergence of these new environments was overall acknowledged as positive and participants welcomed the heightened engagement with museums, especially online. Museums also realised the need to adapt and research these newly constituted userships as:

> *“gerade im digitalen Bereich [...] müssen wir uns auch ein bisschen wegbewegen von dieser Idee wir haben DEN User. Wie machen wir das? Es gibt nicht DEN User, es gibt zigtausend User mit zigtausend Interessen, die wir teilweise gar nicht kennen. Und wenn man dann noch so ein Standardpaket liefert, dann denke ich bedienen wir die nicht wirklich.” [especially in the digital space, we have to move away from the notion that THE user exists. How are we managing that? There isn’t THE user, there are thousands of users with thousands of interests, which we probably don’t even know about. When we just deliver standard packages, then, I think, we can’t really serve them.]* (P28)

There was overall agreement that the stereotypical image of *the* visitor is a redundant remnant of the past and not applicable anymore, whether to the analogue, the digital or hybrid variations. Participants also showed an overall agreement to the notion of new audiences emerging and digital-first visitors. However, museums have been working with established patterns and user profiles for decades, and research was built on those personas. Museum professionals warned that the slow implementation of new audience research methods clashes with an urge of institutions to quantify visitors, a trend visible throughout the thesis’ research interventions. This is fostered by the ways of how institutions are supposed to report their success, mainly expressed through numbers (P18), but also highly motivated by AI and other technologies making it possible to track audiences on and offsite and infer understanding about them from data.

It further emerged in the focus groups that how audiences are understood varies not just from institution to institution, but seems to depend on each professional's interest and personal focus. The personal and professional experiences of the participants played a greater role in how users were understood than institutional definitions as found in their missions and visions. Again, this echoes the possible discrepancies between strategic decision makers and those directly engaging with constituents, especially audiences.

Professionals saw a solution to the problem of unsuitable audience metrics and difficulties to find an apt, generalisable definition in user profiles. Constituents self-defining their profiles was seen as a promising way to engage people and create attachment and belonging, but also to gather valuable data about who their audiences are. The notion of user types was considered by the NG and the Smithsonian, suggesting a more tailored approach in terms of user profiling rather than content-focussed approaches.

Specific to the RS, P16 remarked that users should be able to set the scope of the recommendations before using the system, such as having scholarly endeavours or casual browsing. However, as setting up user profiles requires not just time, but also the right means to process and store such data, P16 suggested, for example, a five-scale slider of initiating the recommender between very open and very specific recommendations to give users some form of agency over the content beforehand, whilst enabling the algorithm to serve tailored content. Whereas the idea of user profiles resonated especially with the bigger institutions, participants further discussed user groups that cannot be targeted by user profiles alone, especially communities.

Communities

Communities and how to engage them was at the heart of most participants when discussing user engagement. Before delving into participants' accounts, some clarification on the terminology shall help to frame the term for better understanding. Crooke proposes that “[r]ather than attempting to define the word, understanding how the term “community” is used is likely to prove more productive” (Crooke, 2011, p. 173). Therefore, the next paragraph aims to establish a sense of how community was understood in the focus groups. Listening to professionals, “communities" can be formed out of constituents, but they are distinct in the sense that they have some form of shallow or deeper level of connectivity to each other, which can range from pursuing a common task or goal (such as crowdsourcing or tagging), living, working, studying, interacting, or being in similar environments (e.g., neighbourhoods, schools, care homes, the same online channels and many more) to - and this is by no means exhaustive - sharing the same beliefs, values, sexuality, or backgrounds. Members of communities further do not have to share all of the same characteristics, but their common denominator is that they interact with the museum environment in a specific practice mainly through sharing, belonging to, or identifying with one of the above. Whilst there can be a notion of marginalisation, often with an aim to give communities that are underrepresented a voice in museums, it is not a defining factor of a community.

The RS was identified as a valuable tool to engage neighbourhoods and help to situate the collection for them:

> *“Thinking about how to situate the physical collection that lives at Platt [Hall], part of Manchester Art Gallery’s collection that forms the Platt Hall collection, in the wider neighbourhood as one part of the material culture of the neighbourhood and finding ways of dissolving boundaries between the museum collection inside Platt Hall and the wider material culture collective of the neighbourhood [...] bringing all of that together and looking and finding connections, this [the RS] seems to have potential here”.* (P5)

Giving communities - and the wider public - a voice in terms of attaching alternative narratives was seen as another form of fostering engagement and meaningful interaction. Ways of attaching community input to collections was considered such as free-text entries, but also in the forms of video and audio files acknowledging that standard pre-formed text boxes might limit users' input and restrict the forms of expression to yet another institutionally mediated form of interaction (P3). Those added narratives and the RS as a means to give communities a way to add to curatorial content mixing perspectives was seen as useful for the Smithsonian to offer "audience engagement at a new level [...] being a museum very much focused on the audience and community that it builds" (P19). With another participant highlighting that opening up collections to various narratives should be part of any institution's postmodern understanding of collecting (P26).

Delegating responsibilities and tasks to communities was also described as a jump into cold water in terms of having people engaging with the collection, but risking criticism or having no one engaging at all:

> *"Ich glaube, es tut auch mal ganz gut sich die Frage zu stellen, was besser ist: ein Shitstorm oder dass einfach kein Mensch in die Digitalkataloge schaut? Das ist nämlich was passiert." [I think, it's important to ask the question: what is better, a shitstorm or just no one looking at digital catalogues? Because that's exactly what's happening].* (P26)

Technologies can further tackle issues that had been overlooked by museums in the past and help people who cannot so easily engage with museum collections. One participant reflected on their engagement with care home residents and that

> *"such thoughts were provoked through experimenting with social media and digital interfaces, [...] enabling us to have discussions and conversations that may have not been possible without technology."* (P2)

Beyond engaging communities, museums also depend on them. The SI relies on crowdsourcing and their communities around it to cope with the sheer amount of information and data they are holding and those communities that "love being part of our mission" (P22). Communities are engaged in a range of activities, from digital volunteering at the Smithsonian Transcription Center, where handwritten historical texts and biodiversity data are transcribed to digital formats, to user groups around testing digital accessibility of content on various platforms across all of the institution’s museums, zoo, and research centres.

Another participant suggested off-loading tasks, such as using the recommender as an outsourcing platform, opening the system up to, for example, crowdsourced image tagging. The combination of tags and pathways created by the system could further establish new relations between users, objects, and information, however, participants questioned the value of such data and P8 asked if it would need professional oversight to check new links and deem them as good or bad in terms of capturing longevity of useful data that is worth preserving, which would counteract useless data collection and unnecessary storage without a clear goal. Tagging and crowdsourcing activities using the RS as a platform for community-based activities, led further to professionals engaging in thoughts of deeper and more meaningful ways and of investment and educational benefits.

Educational and meaningful engagement

Participants thought about the educational value of technologies as experienced through the RS and forms of engagement that do not just entertain but add educational or personal value for constituents. Digital learning and education were seen as possible ways to engage everyone, everywhere and museums were perceived as trusted sources to go to:

> *“They [audiences] didn’t want just anything they could find online, they didn’t necessarily want something from more of an entertainment source, they really were looking for things that they assumed would have some educational value, because they were from a museum institution.”* (P20)

Activities, such as an Antiques Roadshow where visitors bring objects from their homes to the NG, upload information to the recommender and get potentially interesting outputs from it, such as attributions, similar objects in the collections, or potentially information around periods and eras were suggested (P13).

P14 distilled the essence of the RS rendering it as an interface to the gallery, catering to the “very serious person who comes regularly anyway, giving them a new way of navigating and curating or collecting their own set of experiences”, but he also specified this idea of a roadshow further and suggested turning this approach around and actually let groups that are not physically coming to the NG interact with the collection through the RS. Pupils could bring objects to school and through a web application or an API people could then upload their objects to establish connections with objects of the NG’s collection, using the RS as a “distribution channel” (P14).

MAG’s curator for Egypt and Sudan imagined something similar when issuing the need to relate personal stories to the collection, adding attachment and meaning when “someone at home turns up at grandad’s attic and finds something that looks like an Egyptian shabti [...] they take a picture and submit it” (P2).

Discussions about uploading personal photos let ideas around connecting such options of the RS to social media channels arise. Suggestions were made to upload Instagram-like images of daily items that are not necessarily museum-related, but people tend to engage with in millions every day, such as food pictures. They could then be fed to the RS to return similar images. This was rather sarcastically met and shut-down by P18:

> *“The National Gallery collection is kind of beige. So I think you’d get a lot of connections to cat food.”* (P16)

Educational use of the system was also positively noted by P18, who can see both educators and students using it to engage with content beyond traditional means stating that current interaction with museum databases is very much like “searching academic databases [...], which is an impediment to connecting objects or *relevant* objects and resources with sort of non-specialists” (P18). Especially by staff of the Smithsonian Learning Lab, the RS was seen as a tool to connect with kids and teachers and break down barriers between the institution identifying wide potential. The system was seen as a means to overcome such barriers, also through its ability to track user activity on the platform which helps to incorporate such data into objects establishing “new learning objects” (P18) and paths, making it a system of search and discovery.

The reality of new audiences and surge in demand for online educational content prompted interviewees to advocate for more nuanced sets of data around audience research that ask specific questions around reach, engagement, and accessibility. They asserted that to answer those questions requires different methodologies and forms of interpretation compared with on-site metrics.

Professionals saw the RS as a means to help to equalise the relationship between institutions and constituents. P6 stated that they “think that people are used to having more equal relationships with other interfaces like Wikipedia, like Instagram, like social media. They are used to being able to have their voice” (P6) and the RS could possibly shift power relations towards inclusion of their data that was collected through, for example, learning sessions, school visits, community projects. Giving users the option to do what would not be possible through Emu, but through engaging with the system and giving them something to interact with would be really interesting to see (P6). However, to realise engagement similar to those on social media channels, museums need to change user behaviour from just accessing museum online provisions for opening times, or other information, to accessing them because they want to engage with the content for entertainment and content that is as valuable to them as the one shown on social media (P26).

The notion of a possible meaningful personalisation of content through the recommender system and deeper engagement with the collections was caveated by some participants, who urged not to be fooled into ideas of replicating social media on museum websites as what might work well on their platforms, might not work for institutions or might not be wanted. P18 therefore remarked that museums need to think about the experiences they want to offer carefully and talk to users beforehand before going down a path that leads to nowhere:

> *“I think as educators or educational institutions, we want them to sort of think really critically about what they’re about to see. Something that might*

> *be really fun to us [the ten participants at the Smithsonian focus group] might not be really connected with most of what users want."* (P18)

Participants also dared to expand their thought into utopian modes of engagement as envisioned by some of MAG's participants, envisioning novel forms of personalisation beyond the standard interaction through intrinsic and/or extrinsic feedback. P1 asked if there would be the possibility to do something "more sophisticated" to a degree where you actually try to stimulate certain brain areas through the content shown to users, recommendations based on facial recognition or interaction with the collection based on haptic feedback.
Embodied relations to collections that are established, or at least supported by the RS, led participants' thoughts around what the introduction of an RS means in relation to the object held by institutions. The next section therefore analyses what the introduction of such a technology might constitute.

Away from object-centrism towards emotions and relations

Talking to participants made evident that the RS could enable different ways of engaging with objects away from an object-centric view stating that the RS could be a good starting point to create a system that actually defines objects not through their "thingness" (Dudley, 2012), but, amongst others, their relations or emotions.

> *"We think of objects as being like nouns, like kind of discrete entities. And if we think of objects more as adjectives, then they become sort of more qualities. They don't make sense without linking to other things and you know the database positions objects as nouns as discrete entities and actually breaks them down into even smaller entities. So [the RS] is maybe not even object-centred, maybe it's emotion-centred?"* (P5)

Introducing the RS to participants, P1 gave thought to "reverse engineer" (P1) the system, where it is not the collection database that dictates the content, but users through entering their data. This concept is based on the idea to open up the usually closed, or as described before *walled* collection systems to constituents through letting them upload various kinds of data, e.g., textual (jotted down words or data captured through conversations and notes in the gallery), directly through digitally terminals that can also scan handwritten notes or directly from personal devices. This means giving users agency over the content that goes into the database, mediated by a system that is accessible from basically every device that has internet access:

> *"So actually, you could shift how people understand the collection, and actually how curators understand the collection shifts from being object-based to object plus relationship, which is kind of more reality and it seems like there's the means to do that in a professional way."* (P1)

Apart from users adding the data, professionals highlighted that not every collection has the same amount of data and it would be a false assumption that there is data in certain formats at all. Thus, collections with not much or any concrete data can profit from the RS allowing a new permeability of information and shifting nature of objects.

Platt Hall's Mary Greg collection is one of such collections that do not have much concrete data (i.e. collection data as it is usually stored with artists' names, title, date of creation etc.) and P5 noted that there exists a lot of "soft data", which would be valuable to safeguard and preserve for further use, but current systems do not account for it. Traditional systems make it hard to navigate collections that do not fit the standard pattern of concrete data, and interaction between constituents and the RS was seen as having real potential.

Therefore, the RS could cater to a multivalency of objects (Gaskell, 2003), as participants felt enabled to make connections between objects and see them in a light that they would have never thought of before.

This helps to establish objects as non-static entities and the RS was seen as a means to open up the field of discussion around what happens when they come together as objects encountered by humans, but also as means to lead thoughts off the beaten track giving momentum for new thoughts.

During the system demonstrations at the focus groups, participants were able to interact with the system, inspect objects, and choose artworks they want to *see more of*. After the demonstration of the RS and having seen a series of artworks that were recommended based on a prior selection of objects, P5 made a very interesting observation in relation to the sequence of images (see *Figure 5.1.*):

> *"It did raise a smile, seeing the image of the deceased body part, and then the image of the wood cutters. It made me think about that picture completely differently and I'm suddenly thinking about the bodily nature of labour and you know what kind of strength and health somebody who is a wood cutter has doing hard physical labour. So it totally changes the position from which I encounter that object. And that's fascinating. I think that's really, really fascinating."* (P5)

(a)

(b)

(c)

Figure 5.1. *Series of images recommended during focus groups*
(Artwork metadata in *Appendix B,* p. 216)

The same sequence of images also prompted P28 to reflect deeper about it:

> *"Wenn ich jetzt als Wissenschaftler tatsächlich recherchiere zu einem Thema, dann kann ich auch Inspiration vielleicht brauchen, Denkanstöße, die ich so nicht habe und ich hätte dann auch gern mal so eine Suche, die mich auf solche Abwege führt, wie zu dieser Zeit mit den Körperteilen. Das*

kann für mich als Wissenschaftler ein Anstoß sein." [Actually, if I were a researcher investigating a topic, then I might need some inspiration, food for thought, which I'd usually wouldn't get and it would be nice to have a search function that leads me astray, like the one with the body parts. That can be an impetus for me as a researcher]. (P28)

Beyond direct encounters with objects and the establishment of new connections and relations between them and human constituents, the system could enable new citizen science projects, providing a platform of serendipitous coming together, which could especially contribute to collections that are not well documented. Blockchain technologies were further identified as possibly adding to community projects as means to have a traceable interaction, storing activities in ledgers, which could be a way of adding a sense of authorship and accountability (P4).

Questions were also raised around how data is captured and if those methods are still appropriate or need rethinking, having been introduced to the Autoencoder and a short introduction about how the machine processes images at different dimensions than humans process them (see *Portfolio,* p. 29). Participants thought of breaking with standard ways of visual object documentation, inspired by the RS to look at and document images differently. Classic professional photographs of objects were deemed as probably isolating objects and their representation in front of neutral backgrounds risks taking them out of context presenting them as "discrete, isolated, timeless entities" (P5). Such ways could include taking pictures of objects whilst they are being used, rendering them "much more animated and at different levels of proximity" seen through "a different kind of intelligence" (P5).

Participants reflected deeper means of engagement with objects and a meaning-making that goes beyond their physicality. Object-centric practices that are often still prevalent in institutions were questioned, so were the forms of how data is represented in systems used by museums. Not only the authority of objects was questioned though, so was the authority of museum experts themselves and implications a heightened public involvement could mean.

Experts versus public

The notion of professionals talking about other constituents as inherently external to the institutions was emerging throughout the studies, at some institutions far more than at others. The *us versus the others*, not just in regard to views and demands, but also in terms of system usage and implications clearly drawing a line between experts and *the* public, but also between professionals not working with digital technologies and professionals who are working with them as one participant working on digital strategies cynically self-described his team as the in-house “digitale Belästigungstruppe” [“digital harassment troop”] (P24).

Some participants had the feeling that the RS is too basic for experts and they struggled to see the benefits of the system beyond provision of user entertainment, stating that “professionals who are coming to look to engage with collections from a traditionally academic point of view, this is not going to be so useful” (P2), giving consideration to the system possibly not being suitable for an expert usership. This culminated in P28 arguing that museum technologies are generally far too often just developed to suit one type of user, who is mostly assumed to be some form of standard visitor and certainly not an expert, advocating that systems should also cater to the demands of researchers and professionals, emphasising that museums are research institutions and have to cater to scholarly endeavours:

> *“Also ich will das jetzt nicht banalisieren, aber es geht halt schon immer ein bisschen um mehr [...], dass man eben sagt, es gibt auch den Interessierten und den Wissenschaftler und den müssen wir auch bedienen als Wissenschaftseinrichtungen. Eben nicht nur was für großes Publikum.” [So, I don’t want to trivialise this, but there is always more at stake, and one needs to acknowledge that there’s also the interested one or the researcher, whom we need to serve as research institutions. Not just something for the general public].* (P28)

Deeper engagement with the system and further explanation of it managed to eradicate some of the above prejudices and other professionals tried to help to foster more understanding as P16 tried to counter his colleagues' doubts in highlighting that the system was trained on mostly academically curated descriptions and texts anyway, so, depending on the machine learning model used, the system can be very well tweaked to reflect an expert approach deepening contextual information. P16 added that such descriptions are available with the NG's endpoint access in form of around one hundred word short curatorial descriptions (see *Figure 5.2.*) that get played out with the metadata held about paintings and that they had positive experiences with publishing short, but very concise pieces accompanying every painting in the collection.


```
"description": [
    {
        "long_text": "<p>In 1880 financial hardship forced Sisley to leave
        Sèvres, in the suburbs of Paris, for the Seine-et-Marne region,
        south-east of the city. He lived in this area for the rest of his life.
        The move marked a turning point in the artist's career, and he painted
        the river landscape with a new vitality and freshness of vision.<br><br>
        This painting depicts 'le Chemin des Petits Prés', the wooded path which
        ran along the left bank of the Seine, connecting the villages of Veneux
        and By. It has now been replaced by a paved walkway. The village visible
        on the opposite bank is Champagne.<br><br> The young girl, who has been
        identified as the artist's daughter, Jeanne, appears as an embodiment of
        spring.</p>"
    }
],
```


Figure 5.2. *Example of NG's artwork description (here stored as JSON queried via end-point-access)*

The impression of the system not being made for experts caused tension, which some professionals sought to ease through further exploration of the root causes to such presumptions. One factor was the discrepancy between the interface and the backend of the system, where non-tech savvy experts tended to judge the system based on aesthetic factors, such as design of the interface, without realising its power in the background, others highlighted that they would prefer various entry gates to the system depending on domain knowledge and expertise, whilst remarking that what might be deemed interesting enough for a broad audience (i.e., the RS presented to them as an MVP), was not necessarily seen as attractive for knowledgeable users. P30 thus suggested settings that tell the system roughly what users are looking for, such as artists' biographies, topics etc. and added that such functions would be great to implement in the RS as their current browsing options of their digital catalogues are not able to support this (P30).

However, there was not just the question of the system not being suitable for experts as P11 issued doubt that it generally might lack expertise that was sought by audiences engaging with institutions, claiming that

> *"when people turn to the National Gallery they expect a certain expertise, knowledge, in what's being delivered to them. So if I'm looking at an exhibition, there was a curator who put those objects together for a specific reason and kind of guided that. So with this [the RS] the audience that would like to explore things in, I'm not going to say randomised, but very unguided in terms of expert knowledge and a guided way, would they even come to a platform such as the National Gallery to be doing that?"* (P11)

It was not all doom and gloom though, as other professionals also discovered positives of the system related to their workflows. Tasks that would otherwise take away valuable time from professionals or volunteers could be supported by the RS, especially if those tasks demand quantity over quality and rather shallow than in-depth professional attention:

> *"If someone comes to me and wants a very specific Egyptological knowledge, because I'm the curator of Egypt and Sudan, I have subject specialist knowledge that can tell them that, but if you've got, like we have, 50 photography students that are exploring depression, this might be better to connect a Turner to an ancient Egyptian statue of a goddess and the pose of mourning. So in a way, in a good way, this [the RS] takes over my scrambling around trying to find it."* (P2)

Curators, as well as other professionals, were eager to talk about the relationship of curatorial practice to technologies. It was clearly the job role that stood out the most in terms of possible fields of tension, sentiments, and implications around technologies. The grain of conversation though laid mostly not in curatorial practice itself, but in exercising power and a feared loss of expert authority. One curator's comment, when asked about digital endeavours at their institution, summed the curatorial opinion up quite poignantly:

> *"Und deswegen hat er [der Direktor] diese digitale Gruppe gegründet, die uns Kuratoren manchmal vor etwas Herausforderungen stellt, weil das nicht so unser eigenes ist. Aber wir lernen ja gerne und wir lassen uns gerne mitziehen und sind da natürlich immer ganz gespannt, was die andere Seite, in Anführungsstrichen, wie auch unsere Seite mitbringt." [And that is why he [the director] has founded this digital group, which causes challenges for us curators, because that's not really ours. However, we are eager to learn and being roped in and are curious what their side, quotation marks, and our side can contribute].* (P25)

Giving away authority to the *other* side over hitherto expertly curated content, for example opening up the digital catalogues to user comments, was leaving curators in fear of a flood of wrong and unnecessary information. However, after initial trials they soon realised that these comments were not occurring, admitting that the articles or comments posted were "professionelle Mitteilungen [...] und ein echter Gewinn" ["professional messages and a great addition"] (P25). An indicator where experts acknowledged the contribution of the public and the existence of domain knowledge outside the institution and their collections.

Most participants however appreciated the RS as a means to break with institutional hierarchies, although one participant cynically already planned for a hostile take-over, hacking the machine. A domestication of the algorithm was suggested (P16) in the sense that the institution could skew it in terms of what was important to the curators, circumventing the algorithm.

The discussion of curatorial matters concerned hierarchy, authorship and power relations.

"The notion of taking the sort of curatorial presentation of the collection and just doing it digitally rather than physically is probably not the best use of resources" (P7). This sentiment was further shared by other participants who felt that their digital curatorial work was mainly driven by writing the same digital content as they would write for analogue catalogues and apply a copy and paste approach rather than tailor content specifically for a consumption online:

> *"Wir sind da alle in unserem Denken und Wissen zu statisch einfach auch. Also unser digitaler Katalog ist von der Herangehensweise - ich bin jetzt mal ein bisschen böse - nicht viel mehr als ein gedruckter Katalog." [We are just too static in our thinking and knowledge. So - I am a bit mean now - but our online catalogue is nothing more than the print version].* (P28)

Relying on curatorial knowledge is either contested or questioned by participants, as the RS should be open to various interpretations that are accessible without a curatorial background and not necessarily show "another white industrialist in a painting [...] but something that challenges the prerequisite curatorial knowledge" (P2). P2 referred here to the long history of Western museums of showcasing predominantly white men in portraits.

Having to contest curatorial practices was seconded by P6, who found it fascinating that the RS can create pathways that are undetermined by either curators or generally other humans:

> *"I would also say it's a really good lesson for a curator, because you see things that you think you know everything about from a totally fresh perspective and it's a bit humbling isn't it? Taking a step back."* (P6)

Another participant drew parallels between the curator and the algorithm, stating that actually both have the ability to look at objects in depth with a very fine granularity, but through different ways of seeing (P4).

AI technologies were identified as possible useful tools to support curatorial practices, mainly to facilitate research endeavours in building an algorithmic scaffold around objects, based on information on the web and not just in the collection database itself. The RS was seen to have potential for reversing common hierarchies of access to objects, on and offline, highlighting labelling and standard descriptions as “the tombstone that always comes first and that sets and frames the space. But if you take that away, then the space is shaped differently, and it’s shaped by the encounter”. (P5)

Curatorial documentation led to the introduction of a hierarchy of value into the database, which is traceable historically through the levels of documentation of objects. Participants therefore saw technologies as a possible way to circumvent these documentary hierarchies as algorithms do not judge between high-priority and low-priority objects or what is regarded as important by curators. Accessing online collections further means breaking off the limitations of the physical space and its “marquee objects” (P19), which can generate interest and discovery of objects not on view, minimising possible “common curatorial refrains about technology” (P22).

Thinking about the RS further fostered thoughts in participants around technologies being able to equalise relationships between institutions and the world they are situated in, especially participatory forms of social media or crowdsourced websites, such as Wikipedia, where users have a voice.

> *“I think we just seem so backwards in a way that we don’t have that in the museum [...] and I think the next stage in power dynamics between the objects and people is that we become listeners. We all have that data, we’ve had it for years and years through learning sessions. Every learning session produces different kinds of data that we never get on EMu and the conversations didn’t start with an object, it started with the person and what was of interest to them that day in that moment.”* (P6)

This picks up the thought of the RS as a mediator of power relationships towards a more person- or data-centred museum. “Mediation” itself is a contested term as it emerged throughout the studies that this is mainly a question of attributing agency to the RS or not. Some participants clearly attributed agency to the machine, and whilst others acknowledged that the RS has an equalising effect, as in actively changing power relations, the access to data was described as "unmediated" in terms of getting access to the "real" data as compared to having been engineered and tampered with by a curator or a person. Mediation of content and new ways of exploring collections with the RS beyond a curated approach of content led to the discovery of new and often unexpected relations, objects and narratives which led participants to discuss the power of an unexpected surprise.

Serendipity

The concept of serendipity, an often used term in relation to RS and their functioning, elicited interesting viewpoints of participants and how users might experience the collection through interaction with the RS.

Upon interacting with the demonstrator, P5 was reminded of their low-key engagement programme at Platt Hall in the form of a very serendipitous coming together of groups of people to chat about objects. The whole interaction was based "purely on how you react, respond to these objects in their dialogue to each other [...] and what emerges is a set of narratives that is built around the relationship between objects, and people, and place, and time" and the RS was identified as means to replicate such serendipitous encounters, probably enabling the creation of "a completely different set of stories that intersect, but also take you somewhere new. And that's really powerful" (P5).

Other participants found that possible serendipitous effects of the recommender could break with consistencies of search results, making it hard to find things again as they have been moved around, which disrupts things, reminding P18 of Netflix about which he said "irritates me, because I can't find the same thing, they keep on shifting the order of stuff" (P16).
P13 picked up that thought of getting lost on pathways adding to possible feelings of irritation:

> *“I think the thing that worries me is that idea that you could get incredibly lost. You know, that you start off with something you were interested in and very quickly realise you’re going down a rabbit hole and have no idea where you were, how you got there, what it was you were originally doing and after that you know at that point, I think I would certainly abandon it in horror and try something different.”* (P13).

Which was countered by P7 who stated: “You are saying it like it’s a bad thing. I mean, serendipity may be good.” (P7)

Consensus was found in declaring that serendipity per se is not a bad thing, however, it might contribute to getting side-tracked when looking for specific information and “whether you can go back to your original route or whether, effectively, each time you look at it, it offers you a different set of things, therefore you can never reconstruct your past” (P13).

Another concern around serendipitous exploration of collections was sensitive content and exclusion of it from “algorithmic interventions, because if there’s too much serendipity in the system that [...] could cause unintentional hurt to someone” (P22). It was interesting to follow these discussions as it was more and more evident that the system was seen as something with a rather high entropy, disorderly presenting one object after the other that needs some form of containment before people end up in places they do not want to be - again with the slight undertone of agency of the system *doing something* that users cannot necessarily control.

With further explanation and the possibility of safeguarding, serendipity was overall seen as a way to establish new connections in collections for users. Further reassurance was found in the system being able to actually replicate journeys and trace back visited paths, a digital Hansel and Gretel “master breadcrumb trail” (P14), which can take users back to the homepage they had started on.

This section touched upon the human factors in relation to the RS with considerations around a new understanding of objects, the power to create alternative narratives, and novel ways of exploring collections, however, it was also the section that inherited the most tension regarding power plays and authority, mainly driven by a possible threat to break with traditional curatorial practices and a shift of authorship to other constituents. The following section investigates the technological factors and give evidence of professionals' accounts that were mainly related to the RS itself.

5.3. Machine/Technology

This section specifically focuses on themes where participants either talked about technologies directly, such as algorithms and their evaluation or where they spoke about system functionality in relation to the RS. Whereas bias around data was mentioned in the first section, another section in this part of the results addresses bias and explainability further down the pipeline directly related to the system.

Search

Search was a common theme amongst discussions with focus group participants and identified as a fundamental issue for museums; professional or public users alike.

Defining search results and translating them into successful search results was seen as often tedious and cumbersome. "If you don't know the words, you don't know what you're searching for" (P7) is one issue, but then also having to know the words how they were entered into the database is another, leading to search fatigue and frustration across constituents when retrieving information:

> *"Ich suche Schwerter, also suche ich Schwert. Dann kommt aber nicht nur Schwert, sondern alles mögliche. Schwert, der Heilige Georg hat ein Schwert in der Hand, dann kommt die Schwertlilie, Lilie und Schwert, also alles was irgendwie mit Schwert zu tun hat. Das ist absoluter Blödsinn. Das braucht man nicht. Das heißt man sucht sich dann ewig durch und*

probiert die dann aus zu x-en. [...] Der arme Mensch muss dann durch 7350 Objekte gehen wo Schwert drin ist, das will der ja gar nicht." [I'm looking for a sword, I'm searching for it. However, it's not just the sword that appears, but also everything else. "A" sword, Saint George holding a sword, an iris [Iris means Schwertlilie in German and therefore shares the same stemmed words as Schwert, sword], iris and sword. That's total rubbish. No one needs that. So this means that you're searching for ages trying to cross out the unnecessary objects. Poor human beings going through 7350 objects containing the word sword, no one wants that.] (P25)

Frustration about search issues was also vented by P28, who saw in the recommender a "means to just collect likes'' (P28) at first, which he deemed not enough to support his curatorial work wishing for an algorithm that can actually take over some of the tedious tasks of object research as those would be the tasks where he had the feeling that "das System intelligent wird" ["the system becomes intelligent"] (P28). After further explanation of the word2vec algorithm (see *Portfolio,* p. 26), his opinion changed and deploying the algorithm in search scenarios was identified as supportive. Searching metadata either via text strings or filtering, the amount of wrong objects returned is a real hindrance to working effectively as it either returns all related words to the string, e.g., cluttering a search about armoury with lots of botanical objects, or if the search is restricted to the exact word which means knowing the exact word before the search, automatically excluding other maybe relevant things (P28).

Particular interest was shown towards image-based search and how the Autoencoder encodes images to be used for image similarity calculations. Visual search was seen as a new way to establish relations between artworks, such as schools and artists who picked up similar motifs or colours and a way to break away from the reliance on metadata and all its related problems, such as missing entries, spelling issues or undocumented objects. P3 therefore saw in the visual image and related search an "almost like more pure version [of the object] whereas the textual thing is then purely subjective, even if you're trying to just give an objective account of what it is" (P3). Participants saw in image upload functionality also a possible new way for a variety of users to engage with collections. It was further highlighted as a tool to enhance accessibility to online collections for those not native to the museums' languages (P22).

Explainability and transparency of systems

Participants showed eager interest in how machine learning models were trained and how they function to actually serve recommendations (see *Portfolio,* p. 20). There was some form of understanding around what a recommender system does, but none of the participants could identify the algorithms used or what they are based on.
P25 described the word2vec algorithm, when explained that the system is able to recognise various forms of blue even when just having typed "blue" into the keyword search as "spooky" (P25) or were generally seen as rather irritating at points and hard to comprehend (P13).

Explainability and transparency of journeys (see *Portfolio,* p. 32) created by the RS were identified as important when talking to participants. It was evident that there was an urge to grasp either how and why the technology is producing certain outputs or to exercise a certain level of control. It was interesting to observe the use of terminology as explainability often meant to participants to retrace the pathways set by the algorithm rather than being able to explain the functioning and why something might be happening. A seemingly random selection of objects was mostly deemed as unexplainable whilst a coherent or "makes sense to humans" selection was seen as explainable.

P30 felt irritated by the artworks selected by the algorithm, stating that this combination felt weird to her, asking for a certain amount of control and the feeling that she wants to "turn it off when I want to" (P30).

Others were asking for more honesty around the use of algorithms demanding explanations to users when they receive search results that were algorithmically supported. This was mainly driven by P34 and his impression that users get subjected to algorithms basically every day without being told about it:

> *“Von daher finde ich es schon gut, dass man dem User erklärt, dass da ein Algorithmus im Hintergrund arbeitet und das ist ja auch ne Ehrlichkeit. Ich meine, wir arbeiten ständig mit Algorithmen ohne dass uns das gesagt wird und da fände ich es gut, wenn die Ehrlichkeit da wäre. [I think it would be good to explain to the user that there’s an algorithm working in the background, that’s a certain honesty. I mean, we constantly work with algorithms without anyone telling us, therefore I think it would be great if there would be some honesty.]* (P28)

Transparency in regard to findability was a concern issued by participants at the NG. They questioned if the system needs to be more transparent in terms of tracing search results and user paths back to their origins; retracing movements of users in the hope that this helps to offer context and some sort of explanation in a visual or in another form. There was a sense that implicit gathering of data, such as user interaction data, should be made explicit to users.

Algorithms

Beyond the functioning of the recommender as it was presented to participants, there were a lot of comments around the possibilities to use or integrate different algorithms. This was expressed through functions that participants would find useful or interesting rather than mentioning certain models or algorithms specifically. There was also a misconception around how data is processed by algorithms and their capacities.

What distilled out of the focus groups was the aim to establish relationships between data, internally, but also on the web. Participants wanted to search beyond the standard means, imagining algorithms that, amongst other tasks, can support curatorial work or establish semantic relationships between objects. Technologies helping to cope with serial collections, such as large numismatic collections, targeted to grasp quantitative details rather than in-depth qualitative information was seen as interesting.

Also, P25 saw the power of algorithms in aiding search and clustering similar data in the database, but also to make research more transparent, admitting that some of the catalogue texts and entries in the database might be outdated by now, so having a system that links that to external sources and flags similarities was seen as beneficial.

Apart from how algorithms could support professionals in their workflows, more public-facing approaches were topics of the focus groups too. Inspired by Refik Anadol's work, participants imagined deep learning algorithms that use MAG's collection to play with the data held by the institution, giving users the ability to define "levels of fluidity and search flexibility" (P1) or "a slider, from narrow to wide" (P3) where narrow would return very similar objects to wide enabling serendipitous finds.

Algorithms making sense of hidden forms of data such as X-ray fluorescence images of underpaintings or other data that shows objects by means beyond a standard camera image, could add to the establishment of new connections. Experimenting with algorithms and thinking beyond prefabricated goals and defined outcomes "may be completely nonsense, but it may also give us some associations that trigger some associations that we hadn't made before" (P14).

The word2vec algorithm sparked interest around the topic of how algorithms are trained on the metadata, especially around using information from platforms such as Wikipedia and Gigaword (see *Portfolio,* p. 26). Having models trained on data created outside the institutions was seen as "more real world" (P2) in terms of avoiding siloing effects and too narrow knowledge representations that just reflect peculiarities of institutions, which can be rather subjective at points. However, adaptability of algorithms to be trained on preferred terms and taxonomies rather than having to adhere to the pre-trained model (P23), was also mentioned. Professionals also appreciated the idea of training a language model purely based on institutional text corpora, such as catalogues, artwork dossiers, or any other textual material related to the collections. However, one caveat with this approach is that the textual data has to have some richness to avoid the above mentioned silo effects and such approaches need to be further scrutinised in regards to bias issues as closed, institutional systems risk to reinforce possible problems inherent in the data.

The above discussed algorithmic operations highlight some potentially novel ways of looking at collections, however, as playful and promising they might sound, thorough evaluation is key to ensure optimal performance, but first and foremost a pleasant user experience without pitfalls and unwanted outputs.

Evaluation

Evaluation of algorithms (see *Portfolio*, p. 35) was a further concern of participants, made evident by several considerations. There was doubt that quantitative methods of evaluation are not suitable for museum collection recommenders, such as clicks and dwell times. P18 remembered that Netflix collects a lot of quantitative data, but highlighted that the Smithsonian's system would be "based on connecting individual teachers or their students with objects, and so we would be asking questions about did you find what you were looking for?" (P18), suggesting a more qualitative and user-centric approach.

Model-centric evaluation approaches were seen critically and participants had the feeling that, when explaining how model-centric evaluation parameters, such as accuracy or prediction work, could risk developing systems that work algorithmically, but not for the museum usership.

Successful engagement was also seen as very difficult to define as benchmarks and metrics can differ from institution to institution and it is very hard to extract such data and infer if the engagement was successful in the eyes of the user or not:

> *"A lot of users come from a Google search result, they're finding an object, they're visiting just one page and they're leaving again. I have no way of knowing if that's successful. [...] They might just grab an image and leave, and that's highly successful to them, but we don't see it as such as we want people to register and create an account and start making their own*

> *things. And I guess this is actually just a small user group that wants to go that deeply."* (P18)

Surveys were seen as a way to elicit further information and find out about the initial motivations of engagement to then check if users successfully completed their endeavour or not (P25) and then, after having had this form of initial pre-screening, to then

> *"try and figure out how this AI algorithm can contribute to that success or maybe it doesn't fit their profile at all. And we always talk about the user as a monolith, but they are not. There are many different profiles and we need to determine what success means for those profiles."* (P21)

Evaluation was further discussed from an internal perspective. Participants had the feeling that most tech projects are lacking diversity in terms of people and professionals working in positions responsible for the projects, but also, wider stakeholders circles were mostly deemed as actually unfit to reflect a diverse and inclusive usership. This can lead to systems that are unsuitable already at a development stage, but worse, deemed fit for purpose and get deployed without the right people having evaluated it. Hence, professionals stated that it is paramount to include a diverse and representational group of people from early design on right through to the deployment and evaluation stages.

5.4. Discussion and summary

Demonstrating the RS to the professionals taking part in the focus groups and the following discussions that were had elicited concrete and practical in-depth accounts about the RS and the wider museums and AI field. The results present a kaleidoscope of promising futures, deep concerns, and a lot of work ahead, anchored in a system that can possibly change practices and the encounters constituents will have with museums shaped by technologies. Whilst the system itself was scrutinised, it functioned as a placeholder for far-reaching discussions about the museum environments of the four institutions it was situated at.

Kidd's six points introduced in the literature review can be applied as a framework to help shape this discussion and carve out the points made by professionals. Kidd states that the "ways we create, distribute, access and assess information are changing" (Kidd, 2016, p.5) and data used in museums is one of those forms of information subjected to those changes. The focus groups show that data is an integral part of today's networked society and profoundly shapes cultural practices. As evident through the MAIA survey, the steps ahead of actually *doing* AI are pivotal. The systems museums currently use are often the first hindrance to manage and process the data held by institutions appropriately and it is pivotal for institutions to address shortcomings caused by legacy systems, licensing of their digital content, and a lack of OA strategies. Using and sharing the data held by institutions means to address issues around bias and possibly harmful content being published. A possible way forward, rather than completely refraining from unlocking the digital stores, is installing the right risk mitigations and to share content with suitable safeguarding measures in place to enable rewarding and meaningful engagement in the future. Professionals rethought the very fabric of data, questioning if current forms of data representations are appropriate for their use in the RS or if documentation standards generally need to be updated to cater to richer formats beyond what was described as *concrete* data to allow new forms of knowledge creation, diverse interpretation of objects, and narratives that reflect a changing society with digital at its core.

The RS changes how objects are used and perceived, the *thingness* of objects was questioned, and it seems that the RS, or AI-driven technologies in general, will raise further questions around the use and perception of them held in collections and the narratives and interpretations about and around them as argued in chapter seven. Both Kidd and earlier Manovich postulated the end of "whole narratives" (Kidd, 2016, p. 5) or the establishment of "anti-narratives" (Manovich, 1999, p. 82) through digital media and the web respectively. Considering their points from the perspective of the RS it looks like that it can foster various, more personal narratives and forms of exploration that can be more emotionally driven and subjectively meaningful rather than trying to establish a dogmatic linearity of one valid narrative. The system seems to have been a conduit of thoughts about authority and authorship over collections, and a means to unfold practices and museological discourse - widening contribution and allowing for a more equal participation of varied groups of constituents.

It emerged that constituents are more and more perceived as fluid, changing, and not able to be pinned down to a few selected profiles. Profiles can vary and constituents, often enabled by technologies, can take on different personas and roles in society depending on their reasons of engagement (see *Portfolio,* p. 13). As the literature review highlighted, the one user or visitor does not exist and constituents might use many channels with no clear boundaries and 24/7 availability perceiving their environments as well as the museum environments they interact with as a *panenvironmental* structure where the notions of physical and digital realities have started to dissolve, and probably will so ever more with the emerging technologies of AR/VR and the metaverse. Metrics formally used to segment physical visitors do not suffice (see *Portfolio,* p. 37) and the move towards environments that seamlessly accommodate the already emerging interactives of the 21st century and those that are yet to come, probably means for the museum sector a complete rethinking of old structures and to open up the local environments to form bigger structures that encompass co-operative, collaborative, and open practices from how data is generated, used, and shared within and between institutions and the systems they are using.

Education about those systems is key and it was evident that a mix of a lack of knowledge, doubt, and false assumptions could lead to opinions about RS that do not reflect reality or attribute more powers to it than currently algorithmically possible. This makes the case for involvement of a variety of stakeholders from the beginning of the development stage on, not just to set expectations straight, but to meet on an eye-level and approach new technologies with caution and be aware of the level of significance they might or might not play and claims around democracy, all-encompassing access, and participation (Kidd, 2016).

Some participants were prompted by the system to think further ahead into the future, where the boundaries between human and machine are ever more blurring and integrating into each other. Ideas that might seem rather utopian at present might probably push thinking forward to define the museums of the future.

A future where there is a lot of work ahead to drive the promising future of AI in museums and to enable its usage for museums of all sizes as summarised later on in the discussion chapter. Whilst frameworks and guidance about AI in museums specifically, such as toolkits (see Murphy and Villaespesa, 2020), can be very helpful for those institutions that are AI-ready, there is the need for support earlier on in the pipeline helping museums to address issues with their collection data and the means to even start creating, processing, exchanging, and storing data appropriately.

6. Online user study: RS user interaction and evaluation

To evaluate the RS, collect user interaction data and elicit users' perception of the system, a controlled online user study was conducted over three weeks in the summer of 2022. The main objectives of the study were to see if there are any subjective differences between different models of recommendations and if recommendations enhance the UX compared to a random presentation of artworks. Users were further asked about their satisfaction with the recommended artworks.

To create a common understanding of some of the terminology specifically prevalent in this chapter, the next paragraph elucidates the usage of terms user experience, user engagement, and user satisfaction.
User experience can be defined as a "person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service" (International Organization for Standardization [ISO], 2010). In the case of the thesis, this is browsing museum online collections with the RS and the user experience is assumed to be depending on the "user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.)" (Hassenzahl and Tractinsky, 2006, p. 95).

User engagement describes a period of engagement defined as users being "able to focus on their task and the application, the novelty of the experience, their level of interest, and their perceptions of challenge, feedback, and user control inherent in the interaction" (O'Brien and Toms, 2008, p. 943). User satisfaction defines the quality of the user experience and can be measured via interaction metrics that are used as proxies for engagement (Lehmann et al., 2012), such as through the collected interaction data during the user study.

This chapter is structured as follows: first an introduction to the apparatus and the study design, followed by the demographics and domain knowledge of the study participants and, finally the results are presented and synthesised in a chapter summary.

6.1. Apparatus and study design

A web app was developed to resemble museum online collection interfaces of medium to large sized institutions across Europe and the US. Surveying ten online collection websites of renowned museums, the number of displayed objects ranged from six to 55, with some options to include up to 100 items on the first page. Therefore, 30 was deemed a reasonable number of objects to be displayed at the beginning of the study.
A random sample was drawn out of Art UK's 350,000 items, which were received via a data export from Art UK in July 2021, to obtain a final set of 35,000 artworks to be included in the study.

Study design

On landing on the experiment's webpage users were supplied with a Participant Information Sheet and were required to give informed consent before proceeding further (for a 'click through the study' please see *Portfolio,* pp. 38). They then had the option to enter their email address should they want to be kept updated about the study and its results. If a user did not give consent, the process abandoned and took the user to the end page of the study, barring them from participation.

Participants were not made test aware, however, they were informed about the nature of the study being held online, the data collected, and the approximate duration of the study. Demand effects, even if participants elicited the study's purpose, were not expected (Mummolo and Peterson, 2019).

Part I:
Participants were presented with a web page (*web_page_1,* see Portfolio p. 40) containing 30 objects (displayed as *image, artist's name* and *title*; for a full list of

artworks please refer to the catalogue in *Appendix C*, p. 218) representative of the set of 35,000 objects that were sampled earlier to be included in the study.

Participants could interact with the objects and were able to obtain more information through clicking on the artwork image, which opened up a single view page containing the artist's name, artwork title, year of creation and a 'show me more' option displaying further information, e.g., themes, topics, and notes (see *Portfolio,* pp. 41 and 42).

The single view page could be closed again and the subject returned to webpage_1, where they had the possibility to either inspect more elements or start to complete task 1 (*task_1*), which prompted the user to select up to 10, but at least 5, artworks that they wanted to see more of. After having selected up to 10 artworks participants proceeded via a button to webpage_2.

Depending on the randomly allocated condition, participants either saw a set of 30 recommended (*RM, RI, RMI*) or 30 randomly (*R*-) selected artworks on webpage_2 (see *Portfolio*, p. 33).

The study included 4 different conditions, split into two groups:

<u>Recommendations:</u>

- *RM: Recommendations based on metadata model*
- *RI: Recommendations based on image model*
- *RMI: Recommendations based on both, meta and image data (concatenated model)*

<u>Random:</u>

- *R-: Random selection of objects without recommendations*

The conditions were randomised and participants who were presented with either RM, RI, or RMI in Part I, were shown randomised artworks in Part II and vice versa. To ensure further randomisation, the recommender conditions were split between subjects and randomly assigned accordingly.

This means that every participant got to see one of the recommendation models in one part and a random selection in the other as the baseline to compare to. Thus, one recommender model and a random selection were tested within-subjects, whereas the recommender models themselves were tested between-subjects.

After having repeated the task of selecting 5 to 10 artworks out of the newly generated sets of 30 (webpage_2 to webpage_5) for five times (task_2 to task_5), participants' attention was checked through a short, one question task as outlined in the study material (Hughes-Noehrer, 2022f).

Part II:

As explained above, Part II consisted of either one of the recommender conditions (RM, RI, RMI) or the random model (R-) depending on what conditions were assigned in Part I. Participants then went through the same procedure as outlined in Part I, being displayed with webpage_6 to webpage_10, completing task_6 to task_10.

Post-study questionnaire:

On completion of Parts I and II, participants were asked to complete a post-study questionnaire for each part (see *Appendix C*, p. 228). The questionnaire used adapted questions from the Knijnenburg et al. (2012) Framework to understand user satisfaction of RS. Participants were asked to respond to questions around the perceived quality of recommendations, system effectiveness and fun, as well as choice satisfaction.

Answers were entered via 5-point Likert scales ranging from 1 (completely disagree) to 5 (completely agree). Each questionnaire further contained two questions to ask about users' test awareness at the end (see *Portfolio,* pp. 45).

General questions

General questions were asked after participants had completed the questionnaire (see *Portfolio,* p. 47) to gather information about their intention to give feedback, trust in technologies, and questions specific to online collections of museums to elicit a more general understanding about the users.

Debrief

At the end of the study participants received debrief information presenting the study aims and design. Contact information was also presented for users to ask questions about the study and to raise ethical concerns, as detailed in the PIS.

Study participants' demographics and domain knowledge

Altogether, 167 participants completed the study of which six had to be excluded due to failing the attention check task, resulting in a final total of 161 participants to be included in the study analysis. Demographics of the study participants are summarised in *Table 6.1*.

Demographic variables (dominant in bold)	**n (%)**[11] **(N=161)**
Age	
18-20	5 (3)
21-29	**38 (24)**
30-39	**38 (24)**
40-49	23 (14)
50-59	24 (15)
60 or older	33 (20)
Gender	
Female	**104 (65)**
Male	50 (31)
Prefer not to describe	5 (3)
Non-binary	2 (1)
Education	
High school	14 (9)
College	31 (19)
Bachelors	**56 (35)**
Masters	48 (30)
PhD	12 (7)
Employment	
Employed	**101 (63)**
Retired	27 (17)
Student	23 (14)
Unemployed	10 (6)
Disability	
No disability	**140 (87)**
Disability	17 (11)
Prefer not to say	4 (2)

Table 6.1. *Demographics online user study*

Pre-study, users were asked about their domain knowledge (*Figure 6.1a*) in regard to art collections and how often they view collections online (*Figure 6.1b*). They were also asked how often they tend to visit museum and art gallery websites (*Figure 6.1c*) and physical museums (*Figure 6.1d*). The questions to measure visit and viewing frequencies were taken from ICOM surveys (ICOM, 2020a, 2020b) to ensure comparability to the sector.

[11] Percentages are rounded (.1-.4 rounded down, .5-.9 rounded up).

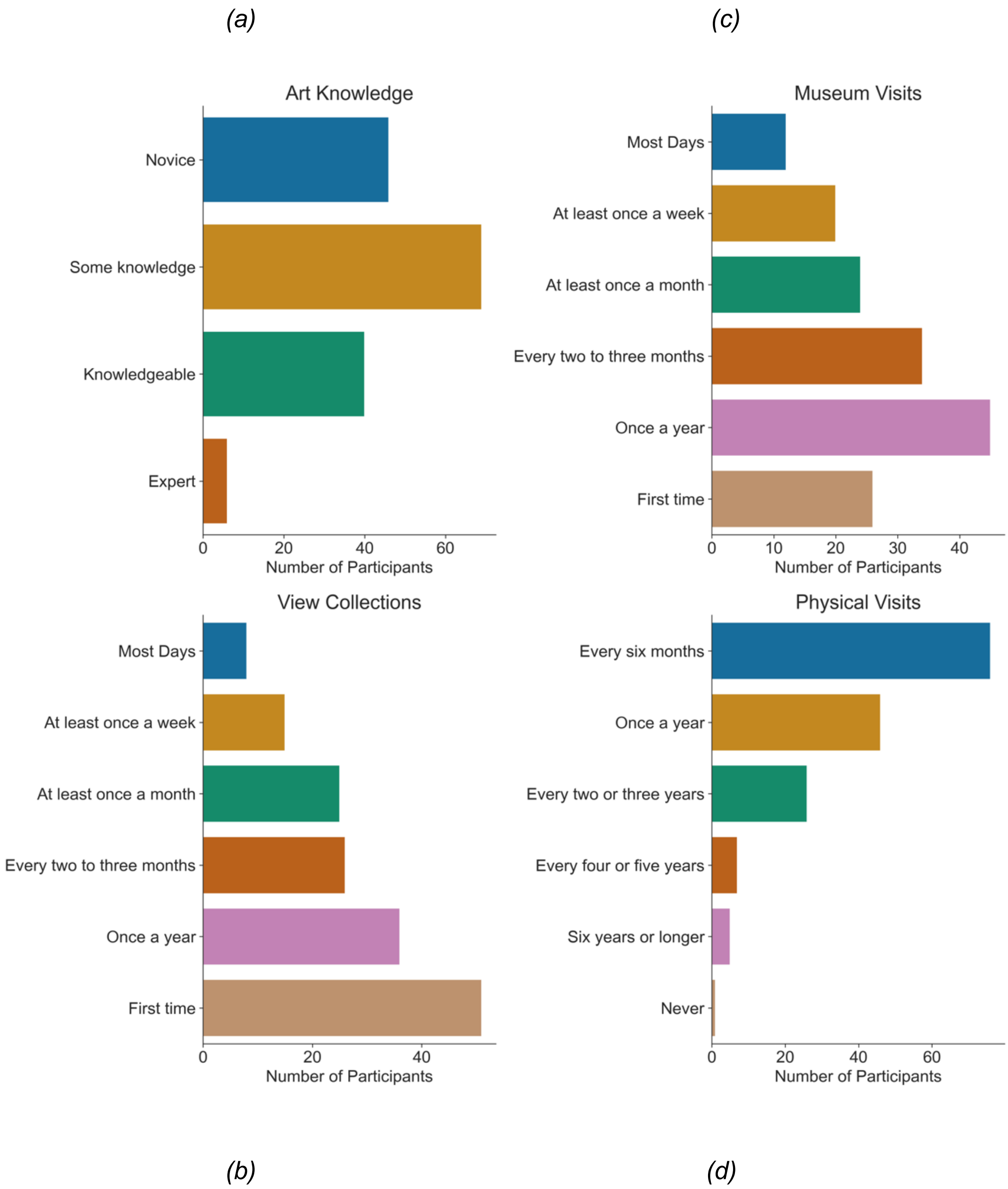


Figure 6.1. *User study participants' domain knowledge*

6.2. Results

Results of the user study are presented in two separate sections, split into extrinsic and intrinsic user feedback. To briefly refresh, extrinsic feedback is acquired through direct user feedback in the post-study questionnaires and intrinsic feedback is gathered through user interaction events in the background, i.e. without conscious, direct feedback from the user. Data about the logged events, questionnaires, and mappings as well as the Jupyter notebooks for statistical analysis are open-sourced at the following: Hughes-Noehrer (2022g).

6.2.1. Post-study questionnaires

The results of the post-study questionnaires are presented according to the grouping that was used during the study and are shown in direct comparison of recommendation conditions to random. Participants completed the same questionnaire for each part of the study, this means that every user submitted two post-study questionnaires and one general questionnaire at the end. There were no statistically significant differences[12] found in the data between the recommender models and the random condition in all of the tested categories (SSA, INT_q, EXP, PS, and SC). Thus, results are presented as direct comparisons of recommendation models (aggregated) and the random condition.

Perceived system effectiveness recommended vs. random

Most of the participants had fun interacting with the system and the majority of users agreed that they would recommend the system to others. By far most users agreed that using the system was a pleasant experience, they indicated that they found interesting artworks with it and artworks they would usually not come across. Overall, the system was deemed useful and users stated that it made them aware of their choice options. Those options also led to more informed choices of users and they stated that they would not have been able to find better items without the system.

[12] For statistical data table please see Jupyter notebook *post-study.ipynb* in Hughes-Noehrer et al. (2022g) as table exceeds size to be presented in the thesis.

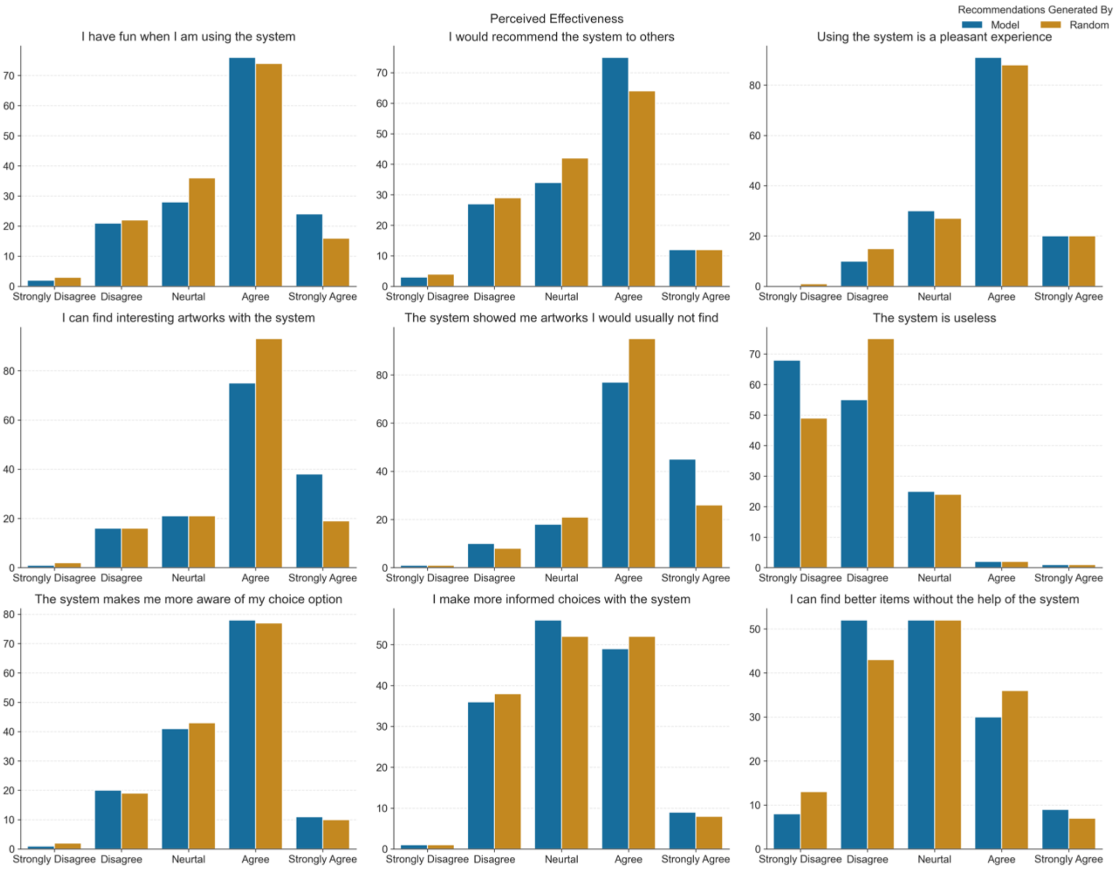


Figure 6.2. *Perceived effectiveness by comparison recommender and random*

Perceived quality recommended vs. random

The vast majority of users liked the artworks that were shown to them and they fitted the users' preferences. Although not significantly, users indicated that the artworks randomly suggested suited their preferences slightly more. Participants further deemed artworks as well-chosen by the system, favouring again the random selection, however, not significantly and this time the recommender models received more counts than random in the strong agreement.

Relevancy of artworks was mostly seen as neutral, closely followed by agreement though. In terms of taste, most of the users thought that there were not too many bad artworks shown to them and there was significant disagreement when asked about if users did not like any of the artworks shown.

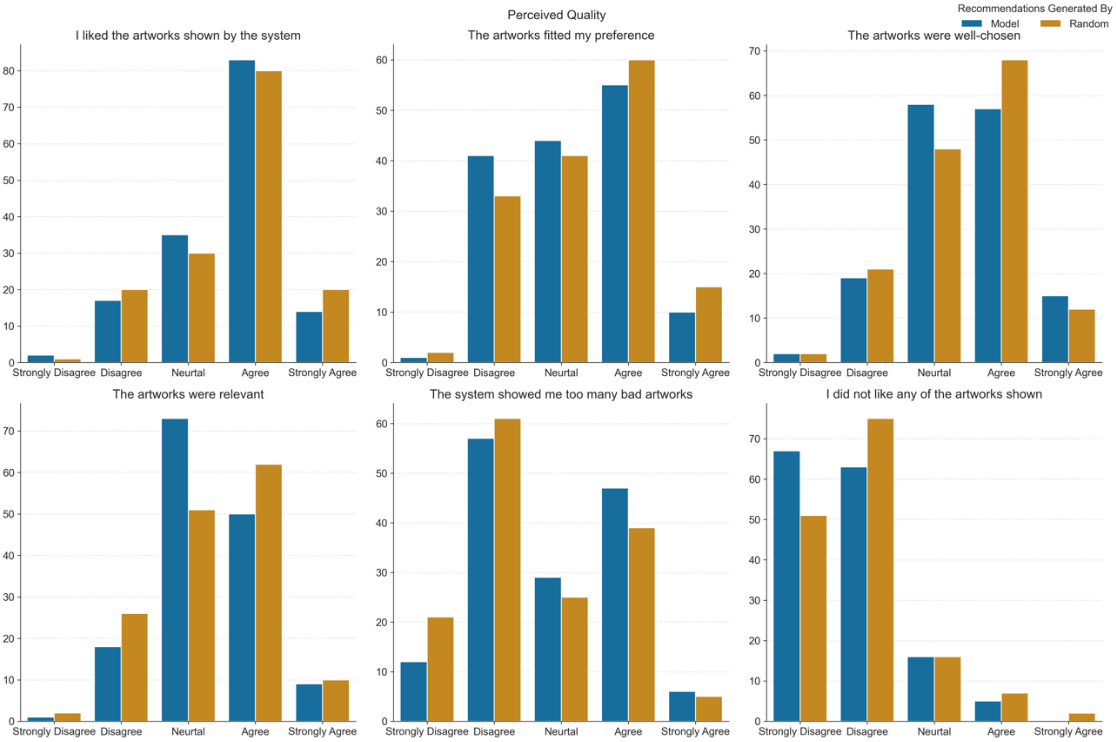


Figure 6.3. *Perceived quality by comparison recommender and random*

Choice satisfaction recommended vs. random

Users liked the artworks they have seen and most of them were excited about the objects they were presented with, with a slight majority preferring the ones chosen by the model, however, not statistically significant. They further overall agreed that they enjoyed the artworks they were looking at. Artworks presented were seen as diverse by users, with users agreeing more so in terms of random presentations, however, they did strongly agree to artworks that have been recommended as being more diverse. Most also deemed the selections as novel, but a lot of users also were not sure about their novelty and preferred to opt for neutral. When asked if artworks shown are serendipitous, users indicated that they mostly cannot agree or disagree with this statement, the only case where this was actually a majority selecting this option. A small number of users thought that the objects they were looking at were a waste of time as the chosen items did not suit their preferences most of the time, with the majority reporting the opposite. Participants also agreed that they would recommend some of the shown artworks to their family or friends.

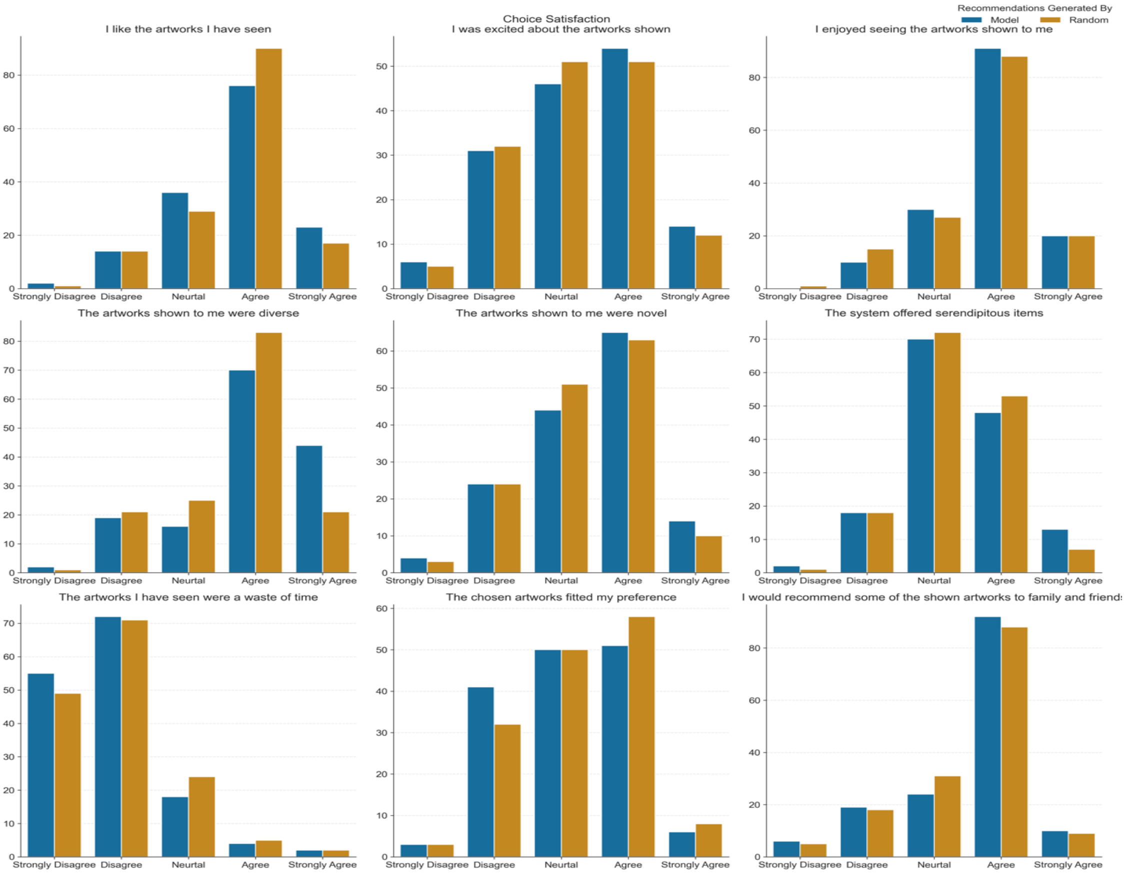


Figure 6.4. *Perceived choice satisfaction by comparison recommender and random*

Test awareness

The questionnaire asked participants if they were aware of the system showing them recommended items and artworks specifically suiting their choice of artworks (see *Figure 6.5.*).

This has yielded interesting results, as users tended to agree that they saw artworks that were recommended although they were randomly chosen, with a slightly higher number strongly agreeing that the presented artworks actually were recommendations, drawing an inconclusive picture. Users further had the feeling that the artworks were not suiting their choice when they were recommended to them, however, these differences are not statistically significant.

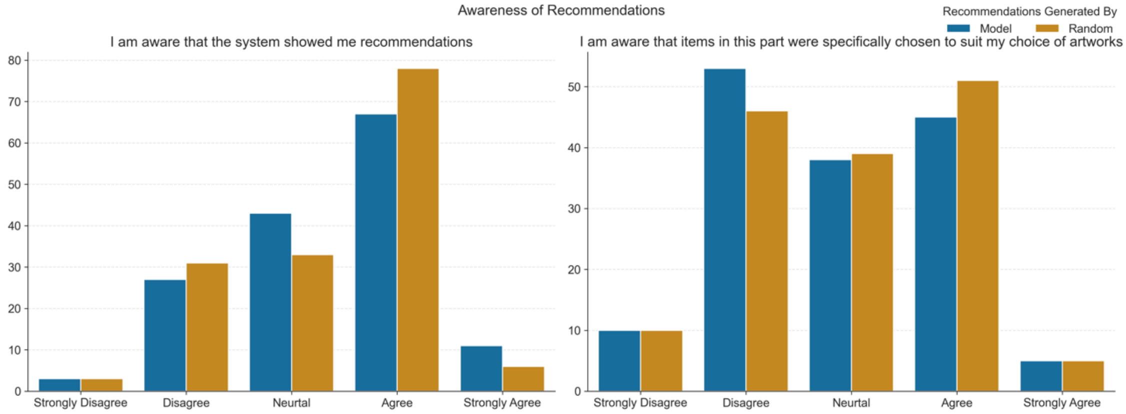


Figure 6.5. *Test awareness by comparison recommender and random*

After having presented the results comparing recommendations versus random, the following sections have a closer look at the recommendation conditions (RM, RI, and RMI) themselves.

Perceived system effectiveness by type of recommendation model

Users had the most fun using the concatenated and image-based recommendations, and there was no significant indication that users did not enjoy any of the models. When asked if they would recommend the system to others, they mostly agreed with image-based recommendations being the most popular choice. Users who were assigned to the image-based model had the most pleasant experience using the system, followed by those who received image-based ones. Users agreed that the image-based model returned the most interesting artworks, however, more users strongly agreed that the concatenated model helped them to find interesting artworks. There was further consensus that the system showed them artworks they would usually not find, with a slightly overall preference for the concatenated model. The question of whether users find the system useless was answered with strong disagreement. Recommendations served by the concatenated model made users more aware of their choice option compared to the other two, whilst users also agreed that this model helps them to make more informed choices, however, the overall most frequently chosen response was that metadata models neither supported nor hindered users in making an informed choice. For the most part, users did not agree that better artworks could be found without the system, however a large proportion of users remained neutral.

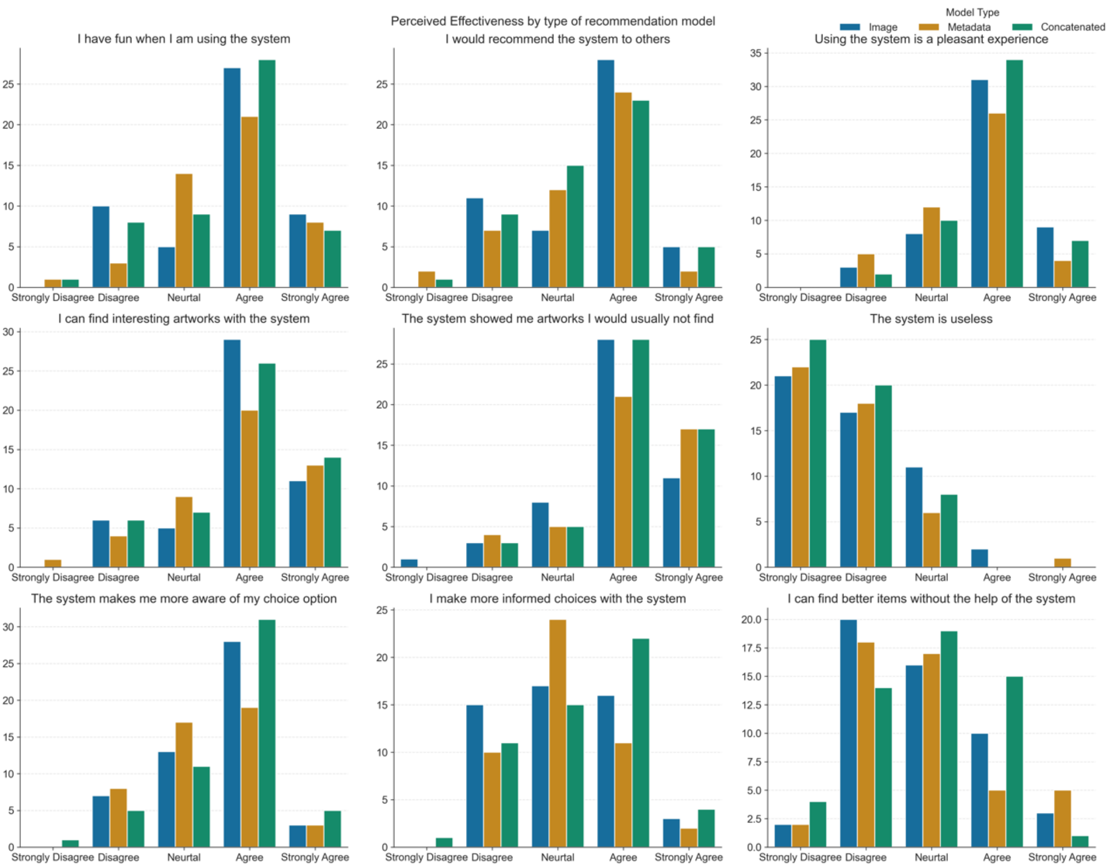


Figure 6.6. *Perceived effectiveness by type of recommendation*

Perceived quality by type of recommendation model

Users agreed that they like the artworks that were shown to them by the system. Artworks mostly fitted the users' preferences, with image-based recommendations leading, which also ranked highest in how well-chosen users thought the artworks were. When asked about relevancy of the artworks, users did not give a clear indication, answering that they neither agree nor disagree with the statement. Inverse statements about the models showing too many bad, or just bad artworks were met with disagreement.

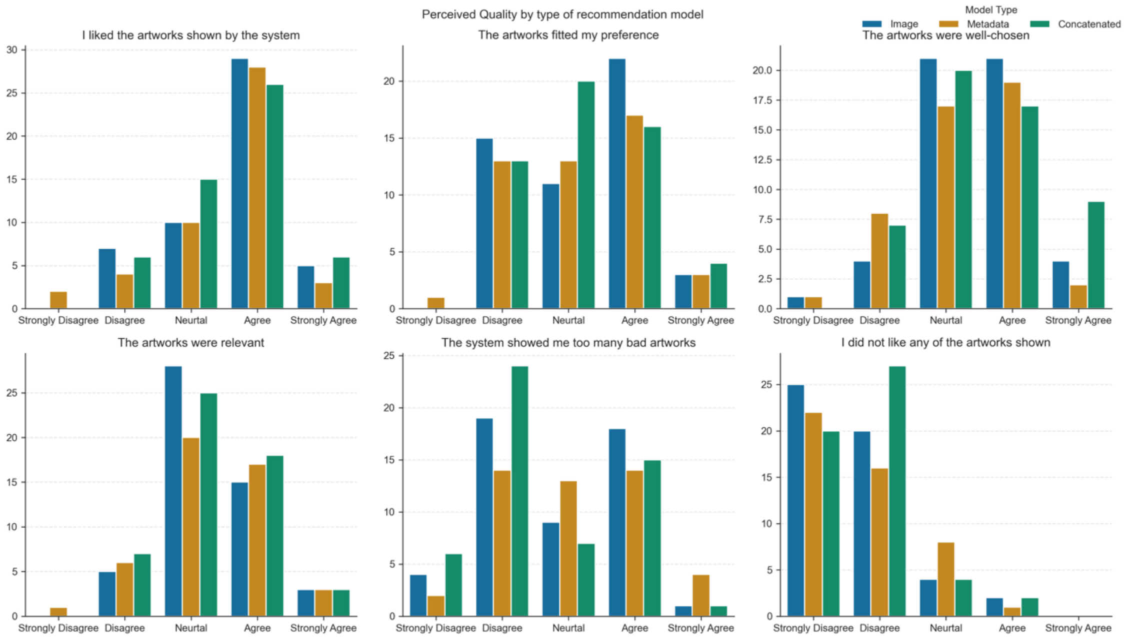


Figure 6.7. *Perceived quality by type of recommendation*

Choice satisfaction by type of recommender model

Models successfully showed artworks to users that they liked, with a slight preference for image-based recommendations. Users were most excited about those artworks that were suggested to them by concatenated and image-based models, whilst the metadata-based one was met with neutral sentiments. Overall, users enjoyed artworks recommended to them, with concatenated and image-based ones in the lead. The artworks shown to participants were deemed as diverse, again with an overall preference for those artworks that have been chosen based on their concatenated or image data. The concatenated model was the clear favourite (add Figure or number) when it comes to novelty of the artworks, and when asked about serendipity, users were neutral. There was strong disagreement that the artworks shown were a waste of time. However, when asked about if they fitted the users' preferences the picture looked dispersed. Most participants agreed that they would show artworks to family and friends and those recommendations would be mainly based on artworks chosen by the concatenated or image-based models.

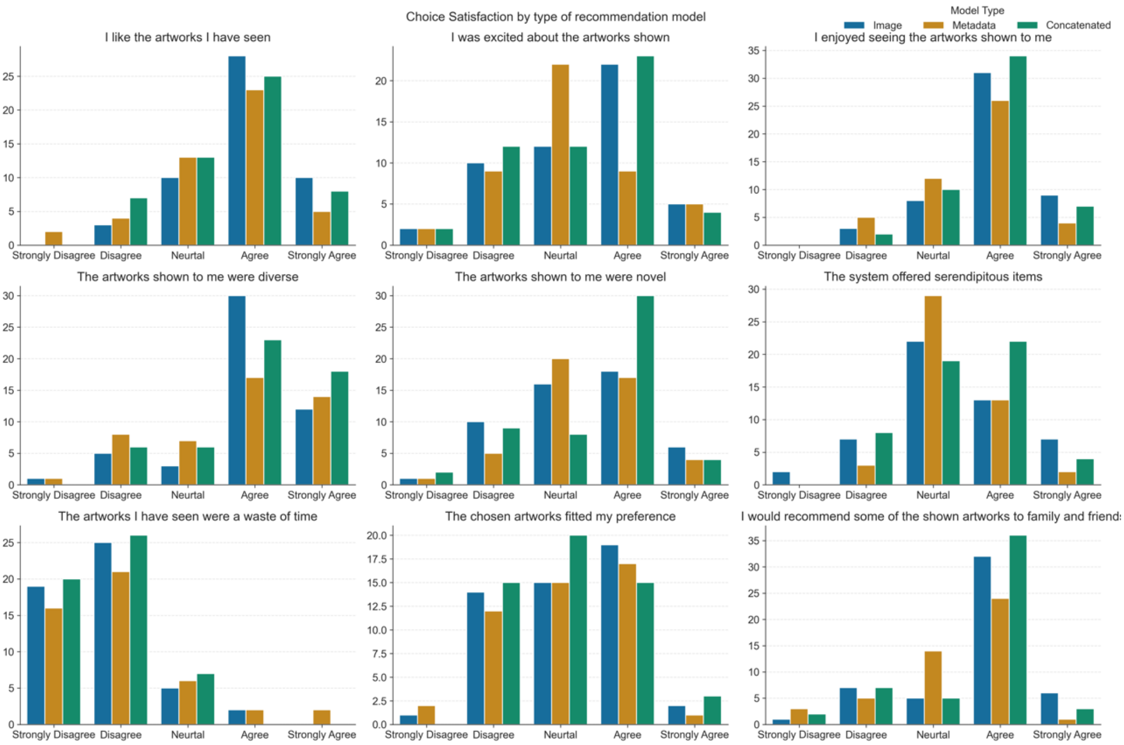


Figure 6.8. *Choice satisfaction by type of recommendation*

Test awareness by type of recommendation model

Participants indicated that they were the most aware about being shown recommendations when being served recommendations based on images, with no difference between concatenated and metadata-based recommendations (see *Figure 6.9.*). Agreement was followed by users who could not tell if the artworks were recommended or not. Users felt further that they had the impression that artworks were selected to specifically suit their choice, when having been assigned to the image condition. They were mostly not aware of the specific selection when using the concatenated model.

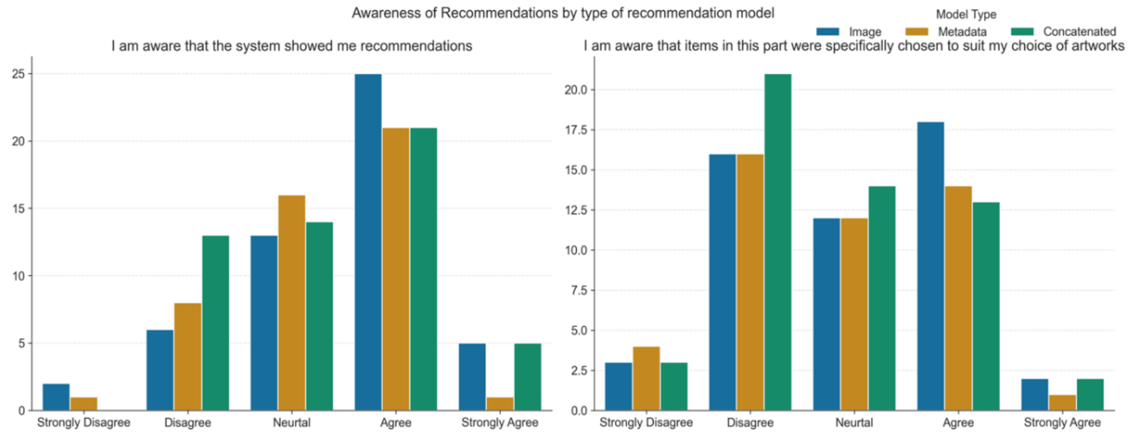


Figure 6.9. *Test awareness by recommendation type*

General questions: Intention, trust, and relevance

At the end of the post-study questionnaires, users were asked to respond to questions around their intention to provide feedback, the relevance of the system and their general opinion about museum online collections. They were further asked about their trust in technologies and data privacy.

Participants were also asked questions around trust and privacy (please refer to *Appendix C*, p. 234). The majority disagreed with the statement that technology never works and most users trusted the system, however, they also stated that technologies should always be explainable. Many users did not seem to care about the actual functioning of the technology as long as it works, and users highlighted that they generally question what happens to their personal data.

Participants were further asked a question about the relevance of the system (see *Appendix C*, p. 234). Most users did not have a preference when asked if they prefer the system or a traditional keyword search, however, there were more users who overall preferred using the system over a classic search. Users also mostly agreed that the system is suitable to display artworks. To determine their general sentiment about museum online collections, participants were asked if they find such collections generally boring and if they do not need museum online collections, both were met with strong disagreement (for plots please refer to *Appendix C,* p. 234)

When asked about if users minded having to choose artworks in the process of using the system, the vast majority did not mind (see *Appendix C,* p. 235).

6.2.2. User interaction data

User interaction data was logged for each user, throughout their time in the study. From this data, interaction metrics were calculated which summarise that time spent in the study, for example, the time spent on pages and the number of artworks visited. In total, 32,150 interaction events were collected, with the subsequent extracted metrics used for statistical analysis. The following results demonstrate relationships found within the interaction metrics. However, no significant relationships, or differences, were found between the metrics and the questionnaire responses (INT_log related to SSA, INT_q, EXP, PS, and SC).

Analysis of recommendation conditions versus random

Applying Mann-Whitney's *U* test, participants spent significantly more time interacting with the system (see *Figure 6.10.*) when using one of the recommendations compared to the randomly suggested parts (U $(N_{rec} = 161, N_{ran} = 161) = 18215.5$, $p < .001$, $CLES = 0.79$) and they also spent significantly more time looking at artworks (see *Figure 6.11.*) that were recommended to them than they had spent on non-recommended ones (U $(N_{rec} = 161, N_{ran} = 161) = 16648.5$, $p < .001$, $CLES = 0.76$). Users spent on average 13 minutes (M=13.37, STD=16.28) looking at recommended artworks, whilst the average for randomly selected artworks was 9 minutes (M=9.03, STD=46.66).

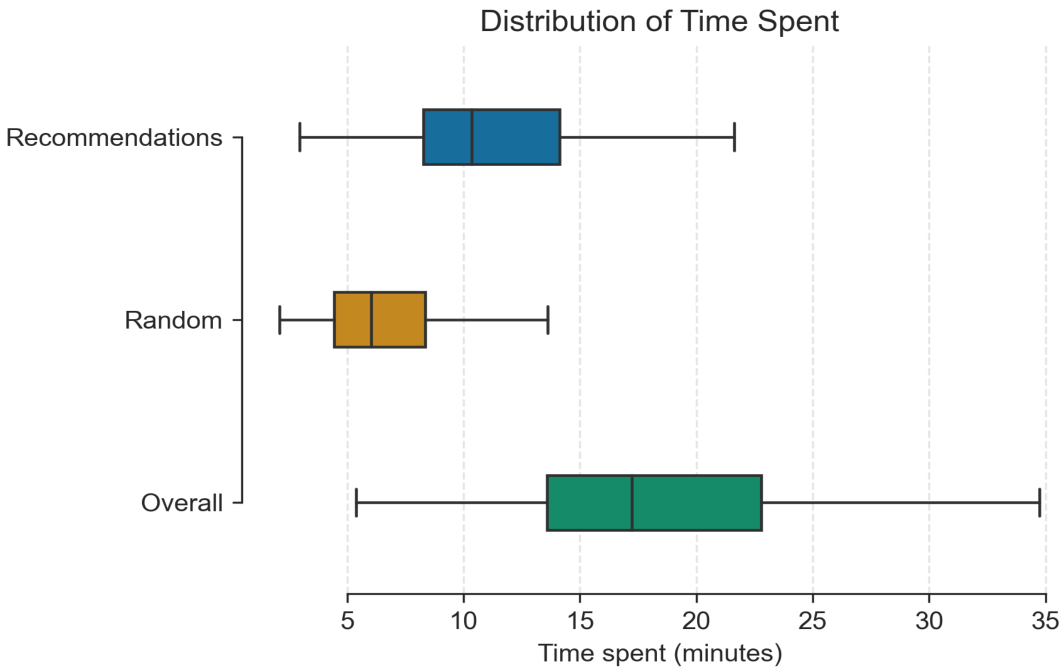


Figure 6.10. *Distribution of time spent using the system*

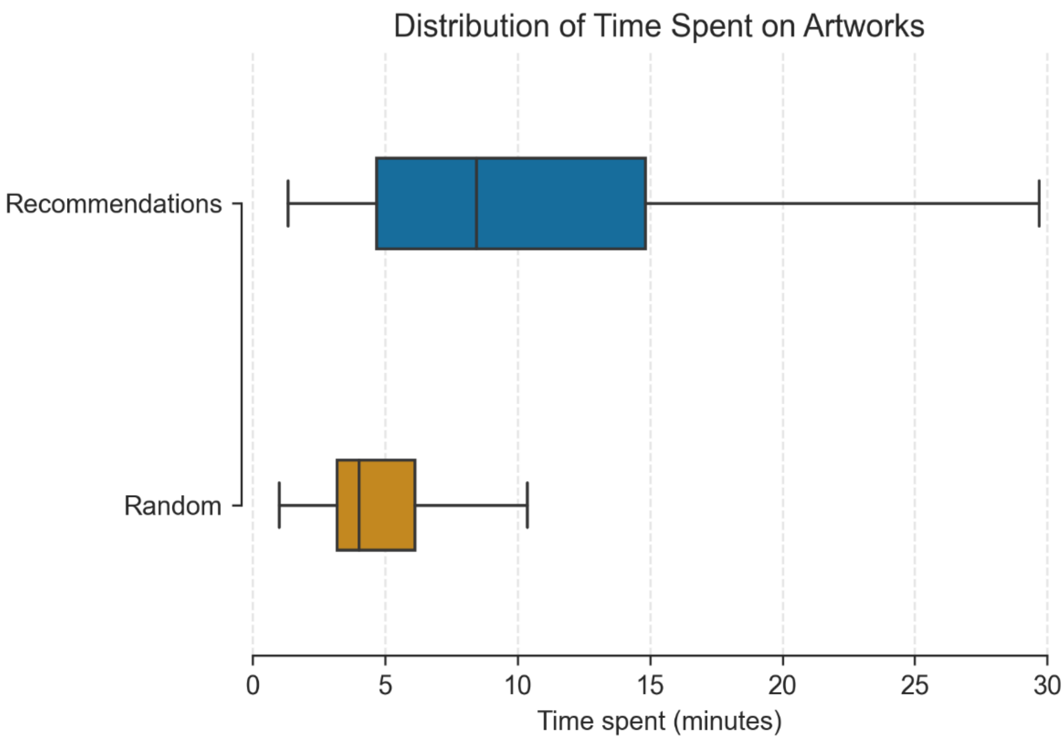


Figure 6.11. *Distribution of time spent on artworks*

As well as users spending more time in the recommendation conditions when compared to the random condition, they also visited a higher number of artworks (see *Figure 6.12.*) when they were recommended to them (U (N_{rec} = 161, N_{ran} = 161) = 17136.5, p < .001, *CLES* = 0.75). Whereas users looked at an average of 32 (M=32.47, STD=12.24) artworks in the random condition, they looked at about 43 artworks (M=42.72, STD=17.20) when recommended.

Users had the option when on an artwork single view page to click a "Show me more" button, triggering a drop-down event that displays additional information about the artworks. This option was used significantly more when looking at recommendations (U (N_{rec} = 161, N_{ran} = 161) = 17169.5, p < .001, *CLES* = 0.75) with an average usage of six times (M=6.47, STD=10.05) compared to just 2 on average (M=2.45, STD=7.68) when browsing a random selection (see *Figure 6.13.*).

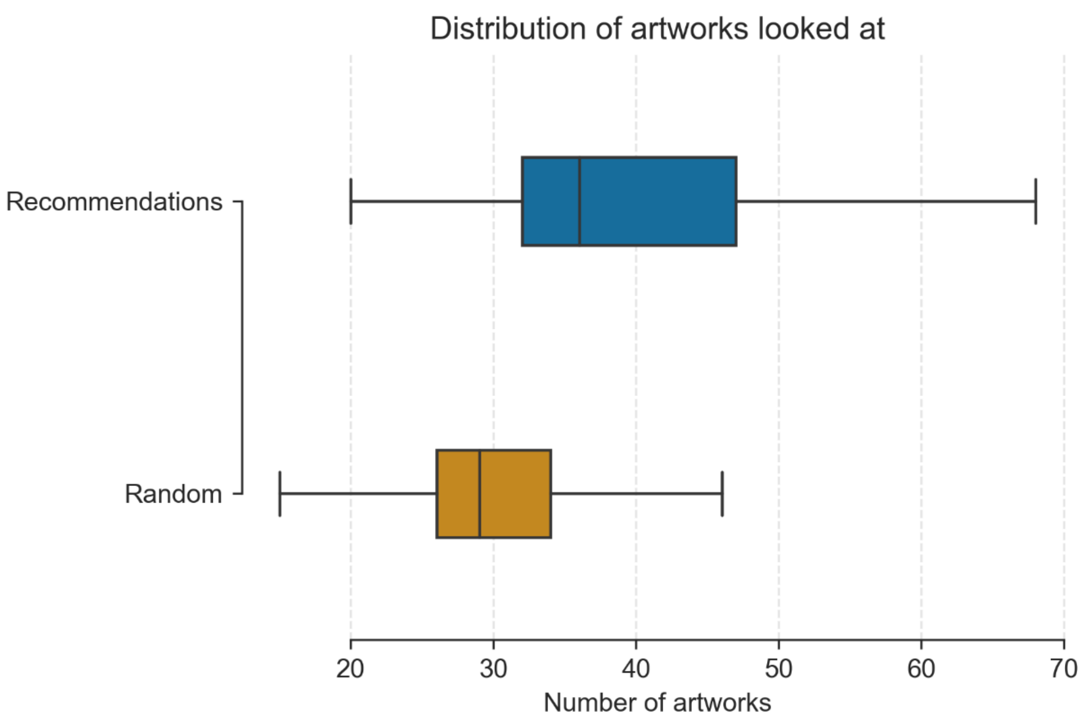


Figure 6.12. *Distribution of number of artworks looked at*

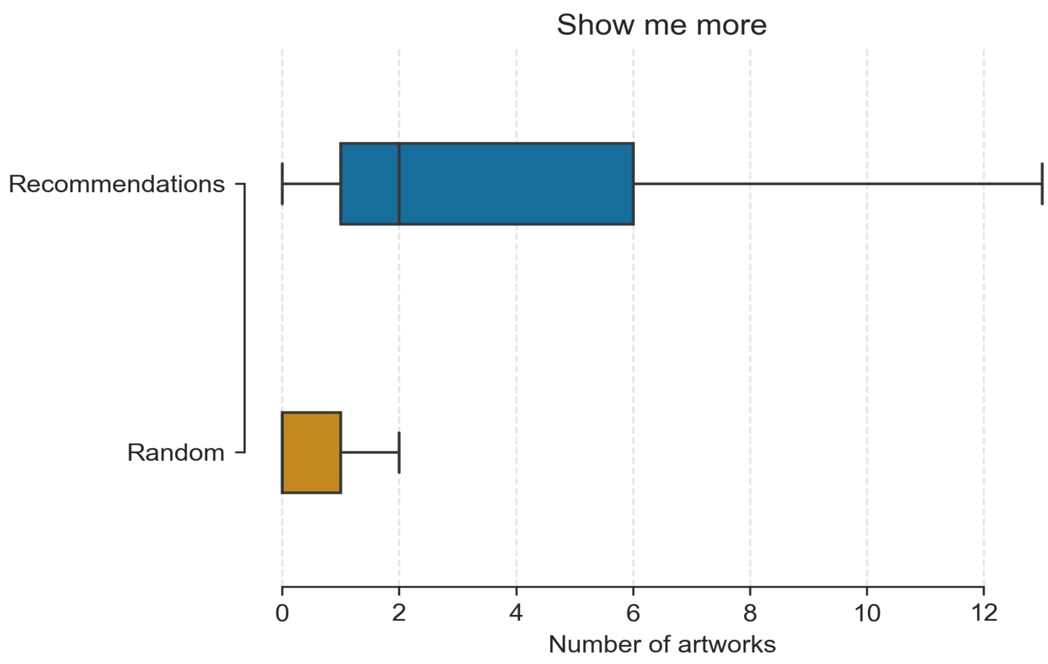


Figure 6.13. *Distribution of triggered “show me more” events*

As participants were asked to select at least five to a maximum ten of artworks “they want to see more of” (the users’ extrinsic choice of artworks to be included in the next step) before proceeding to another webpage view, data was gathered about how many artworks they have actually selected (see *Figure 6.14.*). Users included significantly more artworks they wanted to see more of when they were able to choose out of recommended ones compared to their random counterparts (U (N_{rec} = 161, N_{ran} = 161) = 19528.5, p < .001, *CLES* = 0.85). Whilst they selected 32 artworks on average (M=32.22, STD=4.83) in any of the three recommendation conditions, they selected about 26 in the random condition (M=26.30, STD=6.14).

Spearman’s rank correlation coefficient was calculated to test for possible correlations between the user interaction data and the post-study questionnaire, but no statistically significant relationship was found. Further, a linear regression model was fitted to investigate the relationship of interactions metrics and the post-study framework components, rendering no statistically significant results. Results of the tests can be found in *Appendix C*, p. pp. 231.

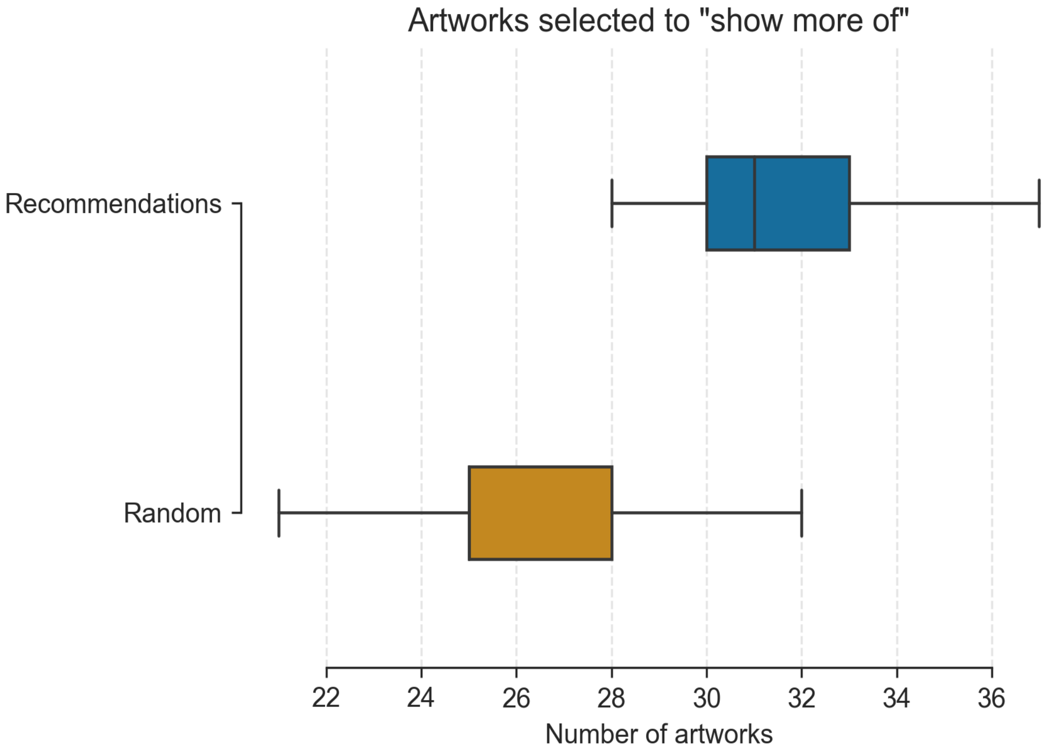


Figure 6.14. *Distribution of selected artworks "Show me more of"*

Results of recommendation conditions by model type

When comparing the between-subject conditions of the three different recommendation types (RM, RI, RMI) to the collected interaction metrics, it is evident that users who were presented with the metadata condition tended to spend slightly more time in this part than those with the two other conditions (see *Figure 6.15.*).

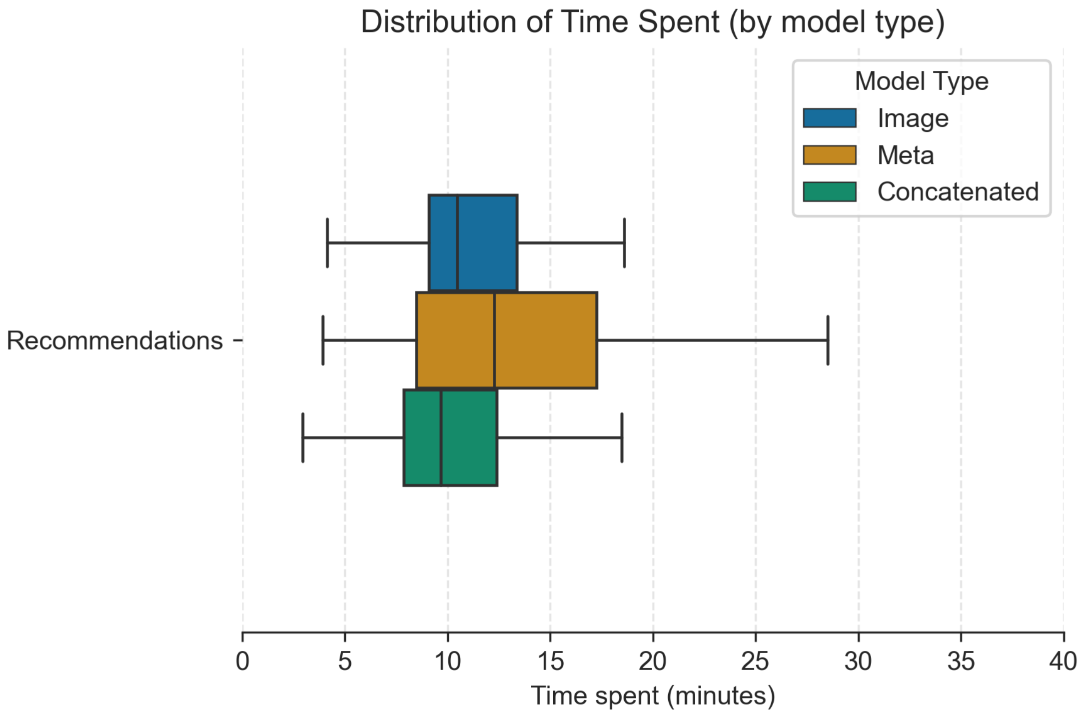


Figure 6.15. *Distribution of time spent per recommendation model*

Artwork recommendations served by the metadata model were, together with those served based on images, also looked at for longer compared to the concatenated model (see *Figure 6.16.*). The distribution of the number of artworks visited per condition shows that users visited the highest number of artworks when in the metadata condition (see *Figure 6.17.*). However, the most “Show more” events were registered in the group that was assigned to the image-based model (see *Figure 6.18.*).

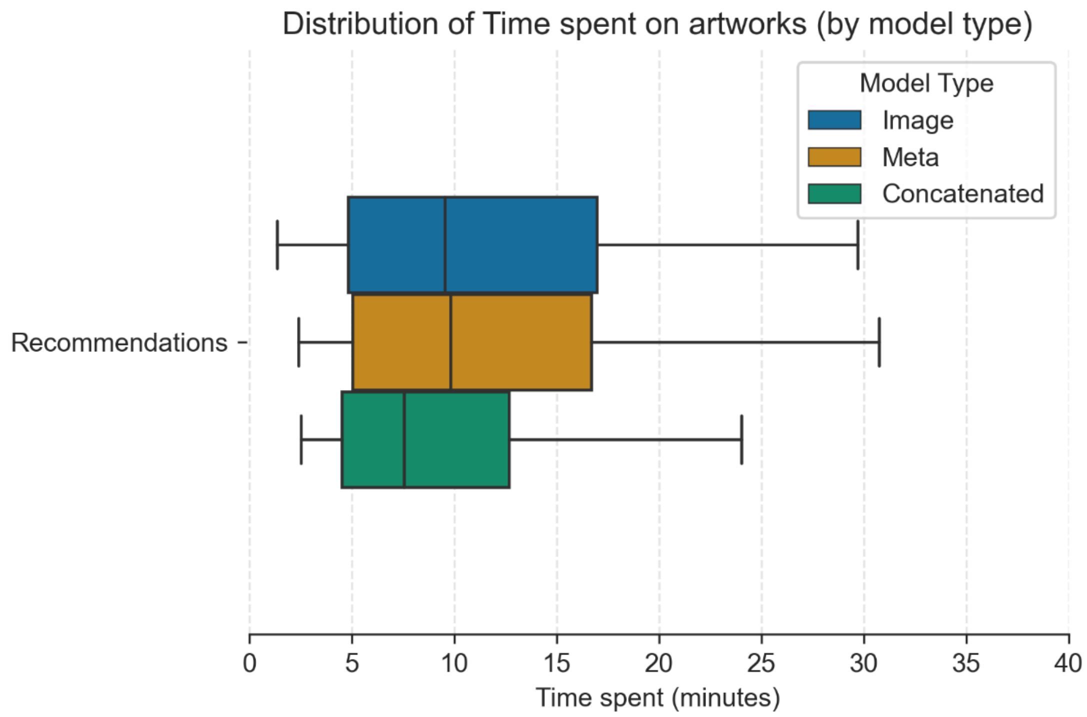


Figure 6.16. *Distribution of time spent on artworks per model type*

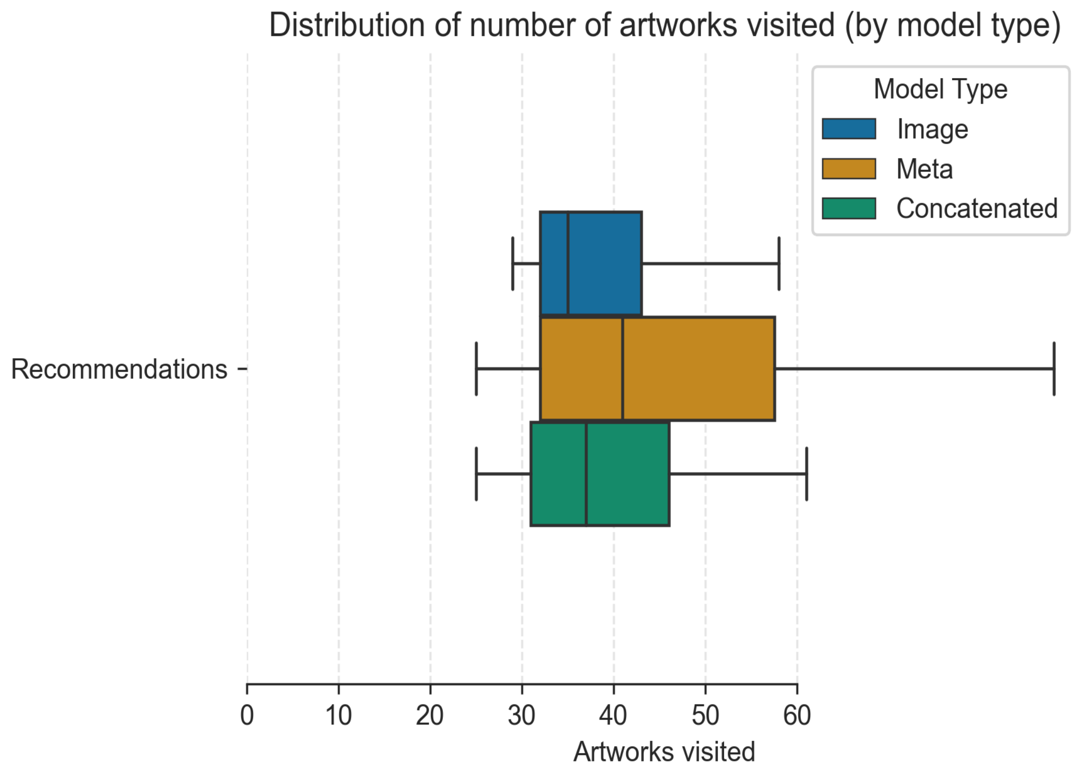


Figure 6.17. *Distribution of number of artworks visited by model type*

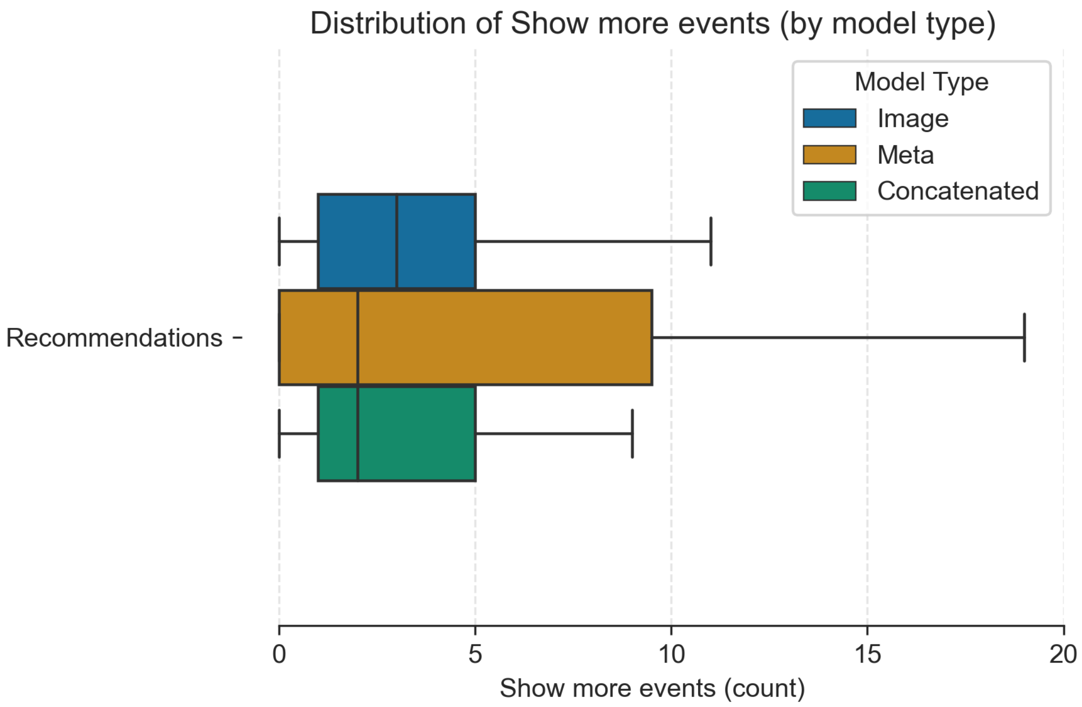


Figure 6.18. *Distribution of "Show more" events per model type*

Users generally chose an equivalent amount of artworks that they wanted to see more of before proceeding to each of the next web page views (see *Figure 6.19.*).

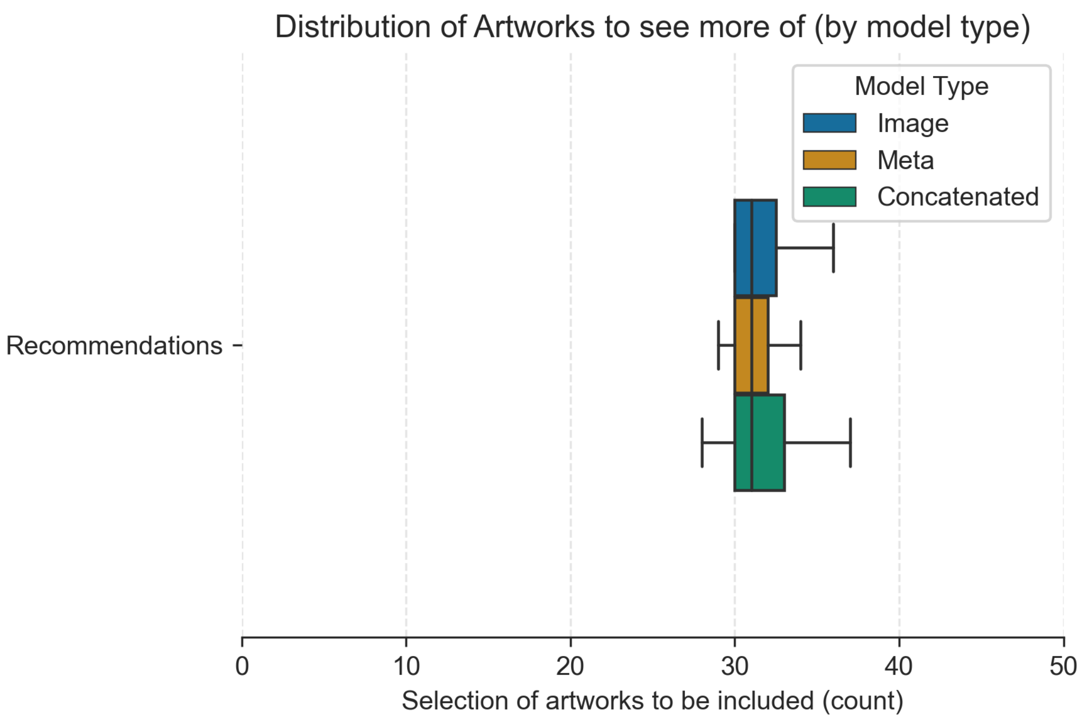


Figure 6.19. *Distribution of artworks to see more of by model type*

6.3. Discussion and summary

The study investigated the application of a museum recommender system through the collection of extrinsic and intrinsic user feedback.

In terms of extrinsic feedback from users regarding their perception and evaluation of MuseREC, the research finds that users enjoyed interacting with the system and that it positively contributed to helping them not just to discover objects that they would not have found otherwise, but, most importantly, it made them aware of their choice options. When using the system, users felt like their choices were more informed and artworks felt well chosen, fitting users' preferences and presenting content that was exciting and enjoyable, and which they would recommend to friends and family, indicating a positive experience with the system. Users liked looking at art using the RS and they showed trust in the system, but also highlighted that technology should be explainable. This makes the case for open and transparent applications in museums that give users explanations about why they see what they are seeing without necessarily giving a lot of background information about how the technology itself functions in the background.

The study found that measuring differences in subjective perceptions of the various conditions through the applied user experience framework presented challenges, which has several implications for further research and museums implementing an RS. Users' perceptions of the recommender conditions (RM, RI, RMI) did not significantly differ from their perception of the condition that served random artworks (R-). Whilst this did not negatively impact their overall user satisfaction as users liked to interact with the recommendations and the random selection, it gives reason to future considerations around the subjective perception of recommendation, i.e. seeing obvious differences between recommended features and a random selection.

The RS did not aim to test and optimise the algorithms deployed in the web app. However, it is worth noting that optimising might achieve more distinct results and greater effect sizes should museums strive for different goals, such as recommending artworks with higher relevance or aiming to achieve specified tasks that were defined apriori. Whilst such optimisations of algorithms (e.g., serving recommendations with higher relevance) are possible with the right data, museums need clear user goals and to test them in relation to the UX as optimised algorithms do not automatically result in a better experience. Besides the algorithms used in the model, the museum collection data itself can play a crucial role in how users perceive the system and the recommendations served. Heterogeneity, richness, and quality are all factors that positively or negatively contribute to recommendation models and therefore need careful investigation and consideration. However, that said, stronger effect sizes do not automatically lead to greater user satisfaction as this also depends on the goals of the session.

How user satisfaction is measured can have big impact, and the framework applied in this study is widely used in the field of recommender systems, but is not specifically aimed at museum collections as it, like many other frameworks, mostly caters to commercial applications with clear conversion or purchase goals as successful outcomes at the end of it. Museum online recommender evaluation might be further impacted due to users evaluating a system without having consumed items, meaning, that often to successfully evaluate a system, users or customers were able to consume the item (e.g., having watched the recommended film, stayed at the recommended hotel) before evaluating it (Loepp and Ziegler, 2019). This emphasises the importance of further work around museum-specific or even collection-specific user evaluation frameworks that suit the museum environment and its constituents as demands of such environments differ to commercial equivalents, particularly when aimed to offer constituents learning and educational experiences which aim to challenge the user and not confine them to constrained environments (Buder and Schwind, 2012).

Analysis of intrinsically collected user interaction data revealed interesting outcomes regarding users' experience with the system. The findings show that engagement of users with the recommended conditions were significantly higher than with the random conditions in any tested scenario.

Regarding temporal metrics, the study found that recommended artworks led to deeper engagement with the system. Users spent more time using the system in any of the recommendation conditions compared to random, indicating deeper engagement with content and heightened attention. Further, the time spent on looking at artworks also increased significantly when the objects were recommended to users.

Users were more engaged in terms of the number of artworks they were looking at (number of single artwork page views) when they were presented with a recommended set. Besides click depth, also other click events, such as triggering the “Show me more” event on the single view artwork pages, revealed that users were more inclined to retrieve further information when artworks were recommended, indicating greater interest and investment. Heightened UE with recommended artworks further led to a higher inclusion rate of artworks in the task selection “Show me more of” (i.e. the selection of 5 to 10 artworks necessary to proceed from task_1 to the task_n).

The study further showed that users interacted the most with content that was based on metadata or images, with the concatenated options having the lowest engagement. This has several important implications for future research and museums which aim to offer recommender interactives based on their collection data.

The results of the *MuseREC* user study are thought-provoking and give momentum to future considerations and accentuating the implications around using an RS in an online museum environment. Results are promising in terms of the system performance and acceptance, but also highlight areas for future research and considerations for museums that want to deploy an RS. The study demonstrates that collecting interaction data and deriving metrics within the context of museum recommendation research is useful, providing insights about how users interact with online systems and artwork collections, which will be further contextualised in the next chapter.

7. Discussion

In the preceding chapters, this thesis investigates the broader museums and AI landscape presenting an in-depth account of the perceptions and concerns with digital practices and AI technologies through a sector questionnaire survey, and focus groups which introduced museum professionals to a museum online collection RS and discussed its potential value and associated issues. The RS was also tested through an online user study *in-the-wild* to gather data around users' experiences and interaction metrics and infer their engagement. Whereas the museum environments were represented through professionals in the survey and the focus groups, they were represented by the RS itself – and user interaction with it – in the online study. Together, they build the empirical bedrock this chapter is founded on to synthesise the findings together with the postphenomenological framework. It also threads in the practice-based elements, as presented in the portfolio and the code repositories, without which neither of the empirical studies would have come to fruition. Whilst the three empirical studies established accounts of their specific research questions, and the practice-based element added the reflective and iterative experience of *doing*, this discussion chapter analyses, discusses, and situates these findings.

The next three sections are split into 7.1. *Museum environments*, 7.2. *Constituents*, and 7.3. *(AI) Technologies* to reflect the three pillars of postphenomenology's relational ontology and honour the structure used throughout the thesis.

7.1. Museum environments

Data and how it is created, stored, and processed is a major driving force of contemporary museological practice and defines if and how AI technologies can be used. What Kitchin defined as data assemblages (Kitchin, 2014), those various forms of practices, places, and forms of knowledge, amongst others are also prevalent in museum environments, rendering them as multistable constructs that are defined through the human intentionality towards them, often mediated by technologies (Jensen and Aagaard, 2018). Museum environments of the past were those of steadfast institutions that broadcast their knowledge unequivocally to a passive audience that was allowed to receive knowledge and the museum's interpretation and practices had not to be questioned. What used to be a clearly separated structure from the *outside* world through carefully curated "physical and programmatic barriers" (Heumann Gurian, 2005, p. 203) has been penetrated and became permeable, a considerable part played by the arrival of new technologies (Economou, 2008).

Museum environments are now subjected to a confluence of forces that require museums to react to the demands of the 21st century. These environments are changing, malleable, and fast-paced and the literature review highlighted that constituents enter and exit them whenever they please whilst seamlessly switching between narratives and focal points of engagement. The RS was seen as a contributor to such environments as a possible enabler of ways of engaging with collections that is personalised and caters to the networked nature of present and future constituents alike. AI technologies, through their adaptability and own way of learning from data and making sense out of it, can help to discover new perspectives in collections as they enable the exploration of different areas of museum environments. The system has also ignited a rethinking of how constituents experience and interact with collection data and therefore contribute to their perception of museum content. Through the exploration of different narratives via personalised pathways, RS could contribute towards museums' journeys becoming "spaces of self-expression and empowerment" (Parry and Sawyer, 2005, p. 39) supported by the specific ways of content interaction they enable.

However, the RS also seemed to have been a wake-up call pushing professionals to rethink current forms of data and practices related to it. Questions arose around if data is generally still correctly reflected in museum databases, often adhering to various sector standards, or if contemporary creation, processing, and access require a new understanding of data in general. Beyond internal practices, it was further evident that the RS prompted professional to reflect about what it means to make data available and whether data requires engineering in order to be meaningful, prompting participants throughout the research to consider the possible ways of meaning-making through constituents' interaction with content and the affordances unique to such systems. Apart from a positive engagement and rewarding outcomes, AI applications in museums could, on the downside, cause possibly harmful or unwanted connections. Content experienced through their output needs thorough ethical scrutiny and discussions around the agency of AI in museums. Those discussions further need to carefully evaluate who exercises the right measure of control over them without reverting to old, often authoritarian patterns.

It is evident throughout the thesis that technology is much more than simply a means to fulfil a role, it plays an active role. Whilst this thesis does not aim to trigger a general discussion around the attribution of agency to the RS itself, it assumes "that agency is distributed through human beings and technologies" (Aagaard, 2017, p. 527). This assumption establishes the role of the RS as helping to "shape the context in which it functions, altering the actions of human beings and the relation between them and its environment" (Verbeek, 2005, p. 43). It is those algorithmic operations of AI technologies and their sense of agency that enable constituents to have a "new access to reality that would be impossible without mediation [...] and constitute a new reality, a new "objectivity"" (Verbeek, 2005, p. 135). Floridi and Sanders postulate that it depends on the *Level of Abstraction (LoA)* as to if a system can be deemed an agent or not and LoAs are "determined by the way in which one chooses to describe, analyse and discuss a system and its context" (Floridi and Sanders, 2004, p. 349) which further down the line also defines if a system can be a moral agent or not. The notion around (Artificial) Agents was further theorised by Coeckelbergh, who seems to offer an approach that is suitable for the AI in museums discourse too. He proposes to abandon the discussion around "how 'moral' non-human agents *really* are by the question about the moral significance of *appearance*" (Coeckelbergh, 2009, p. 181). Thus, for future practice and a way to have rewarding discussions about AI technologies it can help to, first, rather than give in to the discussion about what really goes on inside the RS it shall be enough to deem it an agent as soon as it appears to humans as such and prioritise the interactions that are mediated, coshaped, and constituted by it. Second, an AI system is inherently anthropocentric and "there is no artificial or cultural system separate from the practices that construct, imagine, and live it" (Coeckelbergh, 2009, p. 188) as it is from the point of creation, the data it uses, and the interpretation of the outputs intrinsically intertwined with human agency. And, last, the RS can be rendered as a "social agent" (Hortensius and Cross, 2018) through its influence of constituents' behaviours, thoughts, and emotions.

The attribution of agency also requires addressing the topics of trust, authorship, and authority as witnessed throughout the research. As remarked by Parry, the liquidity of digital media has been "at odds with notions of fixity or closed authorship in the museum" (Parry, 2007, p. 107), a notion that is still prevalent 15 years later. The RS was met with curiosity and excitement, but also insipidity and mistrust. Whereas AI technologies are opening up new horizons for some and are means to drive forward museological practice, others face barriers in terms of trust and a lack of usefulness. These various considerations, be them of positive or negative in nature, fall into the technology's power of mediation.

Using the RS in a museum environment means that the technology mediates how constituents perceive content through it. Whereas the RS allows for different journeys and non-linear modes of discovery, its mediation of collection data is still one of amplification and reduction. As Verbeek highlights "technological artifacts mediate perception by excluding certain interpretations of reality and promoting others, so can they make possible certain kinds of actions and inhibit others" (Verbeek, 2005, p. 191). Those amplifications and reductions of certain voices, bias, and ethical issues around collection data are not new to museums, but the implications of them in relation to the usage with an RS, or AI systems are. Thus, they need to be considered by institutions and their use of AI technology for external facing applications as well as those destined for internal use only.

In most cases those technologies are unexplored and therefore seem to cause apprehension and, noticeably, a possible fear of losing control leading to the machine doing something that lets the reins slip out of professionals' hands. Patterns of "political, social and economic dimensions of social exclusion" (Sandell, 1998, p. 406) have always been prevalent in museums from their earliest days on and are still noticeable in contemporary practices. The RS has shown to be a possible leverage towards more inclusive ways of engagement with collection data, or as described by one of the focus groups' participants a more "humble" way, a kind of reboot of exploring museum content. Approaches to AI need to address issues of social and cultural exclusion from the start, informed by a more unpretentious and honest understanding of what data comprises and looks like. It is the non-neutrality of technologies that needs to be taken into account as their mediation of content, that is the data they are using, can strengthen or weaken specific aspects of reality (Verbeek, 2005). There is not *the* museum, there is just *our* museum as museum environments are inherently bound to their human perception and technologies mediate and shape those relationships (Verbeek, 2001).

AI technologies expand how objects and narratives are experienced and they can therefore help in "providing new perspectives and potentially generating new information" (Moens, 2018, p. 76). Enabled by such technologies, museums now have the potential to meaningfully translate their data into forms of knowledge and insight. The meaningful translation of data, that is a translation that can benefit both constituents and museums, can take on various forms of outputs that can range from a better understanding of museums audiences and a more tailored approach to content offered online, to support of research and scholarly endeavours that add value to collections and foster deeper or novel forms of engagement.

Whilst technologies have proven to at least stretch and loosen the fabric of institutions, there are still a lot of considerations to make when using them. Kidd urged not to "overplay their significance" (J. Kidd, 2016, p. 5) and it is exactly *significance* that needs scrutinising to not fall victim to the fallacy of an *institutional detachment* from technologies and AI as a means to equalise the power relationships in museums. Whereas technologies, such as the RS, can give constituents a personalised entry point to collections and a novel way to explore collection data, narratives, and prompt them to create their own meaning, it is to keep in mind that museums, in most cases, are still and will be the very operators of those systems, mainly fed by their data.

Communication, collaboration, and openness to the inputs of various constituents when deploying AI technologies seem to be the key towards safe applications of the future. To do so, museums have yet to overcome the often prevalent silence between institutions that has also not been broken by the arrival of digital media (Bearman, 2008) and one compelling reason to do so is to tackle the challenges ahead together. Whilst there are some grassroots movements driven by eager and devoted professionals who aim to use AI in the GLAM sector for the good (see, e.g., AI4LAM or the Turing Institute's AI & Arts Group[13]), museums might face hurdles – as also revealed by the enforced pivot to the digital during COVID-19 – over the next decades that cannot be addressed by a *just wait and see* mentality or other forms of institutional passivity as it will be the constituents of the future who will be demanding museums to embrace the new environments of technological ubiquity. However, there is an awakening to more collaboration and cooperation and how those relationships are taking concrete form are yet to be discovered. Funding bodies have realised that collaborations between museums need to be fostered and they are thinking of implementing levers to support collaborative projects where there are shared interests visible in museums' funding applications (M Keating, personal communication, June 30 2022). How this will be implemented is yet to be decided, but cooperation and the sharing of skills could be made mandatory to receive funding in the future and spares from unnecessary spending and strain on staff hours.

Generally, technologies are not deployed to replace current human factors that make up the very structure of museums and they are not there to harness all of the authority within the machine. However, the RS and other systems could prove useful in reshuffling and rethinking those factors and help to shift authority and enable, what Benkler hoped for the global community, museum environments through the mediation of technologies to become (at least a bit) more critical, allow for more autonomy of certain constituents, open the institution up for discourse, and a more equitable future (Benkler, 2006). Whatever roles an RS or AI technologies in general play in museum environments, they are certainly no

[13] See AI4LAM (https://sites.google.com/view/ai4lam) and Turing AI & Arts (https://www.turing.ac.uk/research/interest-groups/ai-arts)

mere intermediaries, but actively codetermine how constituents perceive the world around them, not only for the individual user, but on the level of contemporary culture.

7.2. Constituents

Contemporary constituents are fluid and are not to be cast into pre-fabricated moulds of traditional profiling anymore. They take on different roles depending on the context of interaction with technologies, swiftly switch between environments, and their engagement is often fragmented as they dive in and out of those environments, used to not "consuming 'wholes'" (Kidd, 2016, p. 5) anymore. Wholes in terms of narratives, linear forms of exhibitions from entry to the exit at the gift shop or wholes that purvey the physical presence of objects as the sole way of meaningful experience. The research around the RS shows that it has the potential to foster new relations between constituents and objects, but also between constituents and whole communities. It is those affordances that are distinctive for technologies as they not only provide more environments that humans can explore, but also provide more personal ways of experiencing those environments (Verbeek, 2005).

The RS has the potential to engage with online collections beyond standard search and browsing features towards more meaningful ways of interaction with data that resonate with constituents' own stories and emotions, supporting the evolution of museums towards places of "empowerment and affect" (Parry, 2019, p. 286).

The thesis' research shows that, when participants were speaking about the system, it was rarely about the objects that are held in the collections. Discussions, if not about the system and various forms of how to access, process, and share data, concerned the metalevel, beyond the actual holdings. It astonishingly was not about the *real thing* and an object-centric view and how AI could possibly fit into that frame, it was more about objects acquiring new functions inherently different from former object-centred practices. It was not about the object and its *concrete data* – its traditional metadata in the sense of museological documentation – anymore as the RS renders it as “a means to the end of generating an event that occurs in the psyche of the person who interacts with the object” (Hein, 2011, p. 180). The research revealed surprising topics, such as emotions, learning, and spiritual meaning, when discussing the RS and the objects and pathways of meaning that it connected users to, which went beyond the objects’ physical properties and their representation online.

Interaction with the RS suggests opening up the collections to different interpretations that all have validity as experienced through the users’ personal journeys. Hein referred to this as “people-centred” (Hein, 2011, p. 182) ways of interpretation where the museum acknowledges that various constituents can now create different meanings of which all are equal and rightful. Those specific, personal possibilities of meaning-making could further help to establish a sense of attachment to the objects that are experienced through the mediation of the RS. Whereas most professionals saw in AI technologies a way to disburden themselves from tasks and a way to delegate work with the technology working in the background, for other constituents it was the system itself that was the focus of attention. The latter can create a sense of attachment that “comes about when artifacts [the RS] invite engagement with themselves, and at the same time create scope for people to experience and interact with the world around them” (Verbeek, 2005b, p. 143).

This people-centred way of accessing collections with the RS can further foster engagement with objects or narratives in collections that are off the treaded path, increasing the probability of stumbling across content that is not part of either the high-priority audience magnets or the curators’ favourites. Guy Debord imagined exploring cities through *Dérives* - “a technique of rapid passage through varied ambiences [...] involv[ing] playful-constructive behavior and awareness of psychogeographical effects” (Knabb, 2006, p. 62) - which reminds, although computationally driven, of the paths established by the RS as it enables new ways of exploring collections and users to “let themselves be drawn by the attractions of the terrain and the encounters they find there” (Knabb, n.d.). Terrain is here to be understood as the museum environments that are presented and mediated through AI. This playful way of discovering collection data can help “to fragment the meaning of the artefact and to introduce many perspectives, many points of view" (Hooper-Greenhill, 1992, p. 204) and therefore new connections between objects and narratives, but also people.

Fragmentation is not just limited to the meaning of objects as the RS instigated thoughts of technologies enabling users to acquire different profiles and roles according to their engagement with technologies and the reasons for interacting with them, but first and foremost AI technologies, also as experienced in well-known media applications, can give users the power to self-define. Whilst the option to actively choose content is already one form of active agency in using online collections, setting up profiles, giving users the choice to acquire the persona they want to be can give constituents another personalised way to experience museum online provisions as they usually lack such options and online user profiling is based on pre-fabricated audience segments that might be outdated, not applicable, or too standardised. Simon identified that user profiles in museums can lead to "high-value outcomes" and be "the basis for a social experience" (Simon, 2010, p. 42), but it looks like a lot of museums still struggle to fully realise that option online. One of the main reasons might be that content curation according to user profiles, or in RS lingo the already in the portfolio introduced *collaborative filtering* (see *Portfolio,* p. 22), requires a lot of computational power, secure ways of storing profiles, and a system that is trained online[14] to return up to date results or it might be the struggle to give up control over the institutionalised ways of who visitors are or should be and what agency they have.

Rifts between the attribution of authoritative agency to non-professional constituents on one side, and experts - mainly curators - and a connoisseurship on the other are still prevalent, and the RS ignited discussions around who is allowed or who should be able to curate content on a system such as the RS or generally online media of museums. As observed in the focus groups, there was a notion of not just technologies mediating content between constituents and the museum environments, but also some form of professional mediation that

[14] Online trained ML models are learning in real-time and constantly adapt to new data fed to them.

supervises the algorithmic outputs. Museums feared that letting the machine agency alone present content to constituents might end in harmful and unwanted displays and connection and some form of expert moderation could sensitive the system. Whereas this is an option, it would, again, mean that power does not lie within multiple constituencies but within institutions and does not cater to contemporary approaches of new media content creation and curation which are “time-based and dynamic, interactive and participatory, generative, customizable, and variable” (Paul, 2018, p. 84). Some forms of community moderation or volunteer hosts that have oversight over content could be possible ways forward to a more democratic way of interaction that does not centre all power at the institutions.

However, not just user communities who are based online were seen as a way to equalise power relationships and as means to enrich data around collections. Constituents also produce data when attending the physical spaces of museum environments, such as learning sessions or talks. Data gathered there is mostly rich and valuable and systems such as the RS could function as a sort of terminal or gateway to feed such data to online collections and make it discoverable and interactable for a broad usership.

The following paragraphs present one of the most striking and thought-provoking findings of the user study, strategically positioned here to tie into the last section of the discussion chapter concerning the third pillar, AI technologies. The study shows users liked the system overall and they found it, amongst other positive factors, satisfying to use, enjoyable, and a way to find novel artworks that they would recommend to other people around them. Whilst this is a positive outcome regarding future systems development and offers input for further considerations, the major finding lies within the interaction with the recommendation models themselves. The study showed that subjective perception and engagement diverged significantly from users’ intrinsic behaviours. Subjectively, users perceived no significant difference between a random selection and the algorithmically generated recommendations. However, their interaction patterns diverged significantly when evaluating their actual interaction with the collection content. The algorithmic modification or presentation of collection data unnoticeably influenced their engagement, that is, an algorithmic agent worked in the background and managed to draw people in and engage them not just longer, but also more in terms of artworks and pages they looked at, and content they wanted to see more of.

This can help to carve out or investigate interests and tastes of users that are hard to define or were unknown to users, such as a lack of vocabulary to search for something or content which existence was unknown before. Further, it can be a more personal way to not just experience content and find pathways that result in longer engagement with museum online collections, but also, in combination with users' awareness of algorithmically mediated content, a more reflected experience of using AI-aided systems; in a museum and general context.

From whatever angle observed, it is without a doubt that such systems need some form of address in museums - by those directly working with AI, but also the institutions that are not working with AI to address the pervasiveness of such systems in society - and education of a broad usership, especially as a lot of AI systems operate in the background without users even knowing about their existence or their influence on outcomes, from how people experience museum collections to sensitive domains such as health care applications. Museums can be a fertile ground to experiment with and educate about AI and release it from its existence that often gets rendered as black boxes that either function or not (Verbeek, 2005) without letting constituents explore all the grey zones in between.

Whilst the thesis has shown that constituents do not necessarily want to know about how a technology works in its core, it also reveals that trust and explainability are important when interacting with AI technologies. The last section therefore synthesises the findings regarding the technology itself.

7.3. (AI) Technologies

The thesis has unearthed several implications in relation to AI technologies in museum environments. First and foremost, not just that opinions about AI technologies are hugely divergent, so are institutional capacities and set-ups. At the time of writing, only a few institutions use AI applications. The majority is still occupied with getting other digital provision up and running or generally working with systems and data that are not yet AI-ready. However, there are also those institutions that deem AI as not useful or not needed for their museums. All of these factors contribute to the widening AI gap in the sector and it risks drifting even further apart if no broader action is taken on a level that includes as many institutions as possible. With the emergence of new technologies, societies becoming more and more networked and more generations are digital natives, the question is not just how AI can be used in museums, it is what can museums contribute to the general discussion around AI. Ultimately, the pervasiveness of systems and the sheer amount of data that gets used on a daily basis impacts the life of museum constituents as well as institutional practices.

Whilst those discussions can be fruitful, they seemed to be challenging for institutions at times, especially in relation to their data. Animosities emerged in terms of content being made public as it either was not intended to be or needs some form of explanation that is currently not provided, and museums lack staff or time to do so. A conundrum as AI technologies could fill a lot of those gaps as institutions are already successfully using AI methods, such as computer vision techniques, to impute missing values (see *Portfolio,* p. 16) and tags in their collection data or NLP to enrich metadata.

The endeavour to make data publicly available and subject it to algorithmic operations caused tension as it could require institutions to address topics that might either tear open old wounds or cause new ones. Having content mediated by a technology without going through institutional oversight was hardly met with agreement, which further caused a feeling of concealment and a reluctance to make data - although basically publicly owned in the case of the partner museums - available. This might be to hide shortfalls or malpractice around collection care and documentation, the feeling of losing control of something that used to be so carefully curated and inherently institutionally mediated for nearly two centuries or admittance of difficulties and problems; historically and present. Putting data out there, however, means institutions will face a public discourse around topics that they so eagerly tried to avoid for decades, especially around those of biases of a colonial past and the marginalisation of certain groups of society. Whatever the reasons, having the means to computationally investigate the state of collections in the blink of an eye sounds daunting indeed, but having the courage to address those issues and somehow reboot collections data could lead to rewarding outcomes over time and enable new forms of engagement with AI technologies.

Perceptions range from systems being a jack of all trades, where a slider or a button can magically transform algorithms into anything someone wishes for, to very narrow applications that exceed at performing one specific task. Besides investigation of an exemplar AI system in form of the RS, the thesis also fostered wider reaching discussions around technological implementations in museums and the entanglement with them. AI technologies are mostly approached critically and explainability is key to future deployment and the development of systems that uphold principles of responsible use, such as *FAIR* (Wilkinson et al., 2016). However, more fundamentally AI technologies, their potential, and risks are often unknown in the sector and constituents, professionals and others alike, are trying to grasp what those technologies really are about.

This is aggravated by museums being at a technical crossroads, a lot of institutions were driven towards the establishment of digital strategies and strengthened their digital portfolios over the last years, whilst others seem to struggle to keep up with the pace of pervasive technologies and newly emerging environments or the penetration of the environments they are used to by new hybrid ones that seem to merge into an endless stream without clear boundaries. However, if digitally savvy, ready, getting there or reluctant, institutions should use the momentum to join the discourse around AI technologies and their applications in museums now. The thesis shows that there are plenty of topics professionals are eager to discuss and the user study highlighted that there is a necessity to address the agency of such systems as their possibly manipulative character might have no felt impact on some institutions at the moment, but will have in the future - particularly on their constituents who are using them on a daily basis.

A deeper understanding and experimental testing of technologies can establish museums as relevant partners in future technological discourse and a platform to educate and enlighten about such technologies. Whilst museums are still trusted places to seek information, they risk losing their relevance should they cease to address the challenges of the 21st century and the thesis survey has shown that professionals do attribute AI a role in future online, as well as offline, provisions of museums. If joining the discourse is not in the interest of the institution, they still should contribute to or at least observe it for the sake of their constituents.

It is exactly this trust in institutions that can be used to collaboratively explore technologies and work on issues around transparency and explainability. Even if some algorithms might be black boxes and seem obscure, there are forms of interpretation and evaluation that can contribute to more clarity and openness about the decision-making processes of such machines and invite museums to be places of critical reflection about topics that might not involve their core operations but are more of a societal importance. Museums can be both, places that can "serve the everyday needs of the user to entertain, to educate and to inform, but also to shape society, shape communication and shape political discourse and activity" (O'Neal Irwin, 2016, p. 39).
Perceiving AI technologies as socio-cultural systems that enable access to new realities and shape cultural contexts (Verbeek, 2005) also means establishing frameworks of evaluation that contribute to "steer this powerful force towards the good of society, everyone in it, and the environments we share" (Floridi et al., 2018, p. 689) including those of museums.

Using AI for the good requires robust frameworks to evaluate algorithms. However, algorithms are often focused on the specific domains they were originally applied to. Those domains are mainly commercially driven by conversion rates or purchases or other forms of consumption that might not be applicable to museums, e.g., how much of a film was watched or how long was a song listened to. Thus, whilst there are thorough frameworks for model- and user-centric evaluations in principle, it needs careful consideration if those approaches are also applicable to evaluate technologies in museums and if they have enough sensitivity to capture constituents' engagement with museum content. Development of evaluative frameworks should, just as the development of AI technologies, be a collaborative task to guarantee the inclusion of multiple perspectives from a museological and engineering standpoint. Evaluative frameworks need to step outside of existing domains driven by the dominance of commercial AI applications to bring in the wide range of constituents who have a stake in the values that AI might generate.

8. Conclusion

This thesis introduced an RS into museum environments and their constituents to explore the effects, interactions with, and reception of AI technologies and the ways they mediate engagement with museum online collections. To achieve this, the thesis addressed three main research questions; what are the roles and potential uses of RS in museum settings (RQ-1), for accessing, describing, interpreting, and enhancing existing collections, as meaningful data translators (RQ-2) and in what ways does the application of RS in museums challenge and/or enhance the public and professional perceptions of AI (RQ-3)?

Anchored in a postphenomenological framework which enabled investigation of the technology as an active mediator between constituents and the museum, the thesis presented a novel mixed-methods approach using empirical and practice research which spans the current AI and museums landscape and in-depth accounts of museum professionals to real world interaction with a museum online collection recommender system. The basis of all interventions was the development of the system during the course of the PhD which offered insight into the steps necessary to build it. Further, the iterative, changing, and unfolding nature of this practice-based element generated questions that then contributed to the scaffolding and structure of this thesis. The RS therefore became a method in its own right, whilst it also functioned as an object of scholarly investigation in and of itself.

The arrival of new technologies introduced novel ways of exploring collections, but their fast-paced nature, ubiquity, and pervasiveness also lifted the lid of Pandora’s box for museums which is accompanied by a lethargy on one hand and the struggles of the sector on the other. There are institutions that have the means to implement and use AI, there are institutions that would like to use AI, but lack basic structures and strategies to do so, and there are institutions that would rather stay away from it. For some it will be a learning of new skills and collaborating on projects that would be too big to stem alone, whilst for others it is joining discussions from the wider sector and sharing the power and potential that is often centralised in bigger, well-funded institutions. AI in museums is currently the preserve of a select few enthusiastic professionals who aim to drive the implementation of AI technologies forward and believe in its transformative force.

The future deployment of AI in museums is, in the end, a question of significance. If we want to enable the full potential of AI in museums, it needs to become significant and have relevance and value to the museums and to constituents. This can range from support in fulfilling job tasks to leisurely entertainment that excites users. Developing and deploying AI needs clear structures and goals, and an honest and collaborative approach of all stakeholders involved, to create systems that have significance and can sustain in museum environments without becoming another project or loose-ends endeavour wasting time, funds, and nerves. Another fallacy exists around the sense of an *institutional detachment* of technologies and AI as a means to equalise the power relationships in museums. Whereas technologies such as the RS can give constituents a personalised entry point to collections and a novel way to explore collection data and narratives, as well as prompt users to create their own meaning, museums, in most cases, are still, and will continue to be, the very operators of those systems, feeding technologies the data that they created in the first instance. AI technologies are not neutral; they amplify and reduce patterns in data and their application needs to be met with care and the same considerations and treatment as with any other agent in museum environments.

It is further a concern of translation. AI technologies should be able to be translated from specific projects to open-source software that can be used by other institutions. This requires common data standards and frameworks and tools to address and resolve the current problems of museum data practices. Projects such as *Towards a National Collection*, *Heritage Connector,* and *Living with Machines*[15] have realised the need for shared facilities and methods to explore future opportunities towards a common digital environment. However, those projects are in their infancy and need to gain momentum to translate the

[15] For *Towards a National Collection* and *Heritage Connector* please refer to the links on p. 96, find more info about *Living with Machines* here: https://livingwithmachines.ac.uk/

mostly scholarly work of principle research into museological practice on a larger scale. If those endeavours are successful, then it will be easier to conduct data-intensive AI operations that serve many and that can include and benefit multiple institutions to meaningfully translate their data and *re*discover their collections.

8.1. Contributions of this thesis

The main contributions of this thesis are three-fold: first, it provides a theoretical framework to critically analyse the mediative character of AI technologies in museums and their constituencies. Second, it provides in-depth empirical accounts gathered through a mixed-methods approach. Third, it presents a fully developed and functioning open-source RS with an extensive practice portfolio and code repositories. The thesis therefore contributes to a broad field of disciplines ranging from museum studies and HCI research to a more computer science orientated readership.

The thesis provides evidence that RS can play a significant role in the future of museums and that AI-supported systems have the potential to not just change how meaning and knowledge are created, but also how it is accessed and consumed by a variety of constituents. The research further shows that how such system create and translate data is yet to be fully discovered as perceptions of AI applications varied amongst research participants, making the case for further co-evolutionary research interventions and museum professionals, technological experts and other constituents working closely hand-in-hand to shape the digital future of museums together.

It particularly contributes in-depth empirical accounts of professionals giving evidence about the current status-quo and discussions around AI technologies in museum environments and their reflections of using such technologies in museums. By applying a postphenomenological framework to a museological-technological discourse, it offers a new and unique approach to analysis of museum environment-constituent-AI relations at the intersection of museums and HCI. The methodological choices further enabled investigation of the museum environments through various lenses. Whilst the survey contributes a rather zoomed out mapping of the AI landscape touching upon data practices, barriers to technologies, and use cases, the focus groups enable rich and extensive accounts of 30 museum professionals from four different institutions spread across three countries. The thesis highlighted important points for museums to consider when working with AI technologies. The findings are informative on several levels, as they give insight to the contemporary AI practices in museums, whilst can further be used to inform policy and strategies in the wider cultural domain, in both present and future.

The user study explored how constituents are interacting with the system in the wild and gathered extrinsic and intrinsic data, which is available open access including the necessary open-sourced code to analyse it. Besides, the study elicited a major point of consideration for future AI technology applications as it found that subjective perception of content was not congruent with the interaction of users, highlighting the power, but also danger of algorithmic operations and their mediative effects.

The thesis presents a novel method through the practice-based element and the development of the *5-SRLC* that can be replicated for various AI projects in museums and various other domains that work with cultural heritage data containing images and metadata. The life cycle is a thorough, step by step, data science pipeline specifically tailored to museum environments and their specificities that can find wide applicability throughout the sector and for constituents interested in applying AI, or more general data science methods, to museum collection data. To the best of the author’s knowledge, it is the first RS for museums to combine word2vec algorithms for operations on metadata with an Autoencoder to generate different sets of recommendations based on three models that applied an iterative development cycle in close partnership with museums and sector professionals to achieve a wholesome presentation and critical reflection of AI in museums in the current day.

8.2. Limitations of this thesis

Whereas this thesis aimed to be as inclusive and as widely applicable as possible, there are some limitations to this research which are illustrated in this section.

Geographically, the research for this thesis mainly focused on the application of AI in museum environments in the Anglo-American context. This is due to the partner museums and Art UK being located either in the UK or the US and therefore data being mostly drawn from institutions belonging to a similar cultural sphere.

Institutionally, the author acknowledges that the partner museums, and the Badisches which was included in the focus group research, may have different fundings structures and vary in their institution size and visitor numbers. However, they are generally well-resourced medium to super-size museums that may not reflect the structures of other museums, especially smaller or privately owned institutions.

Algorithmically, every algorithm has positives and negatives in its application. The conscious choice to use word2vec and an Autoencoder was made on the basis that these algorithms were deemed suitable to successfully modify the data to be used for this thesis' research. This does not mean that this combination was aimed at competing with other state of the art models and therefore a model-centric evaluation was neither conducted nor intended. Further, the author's decision regarding the feature engineering of the data inevitably inherits some form of subjective decisions.

Methodologically, the survey had an international outlook, but due to the research being conducted in the UK and US, and the author's established networks in these countries, the majority of participants, although self-selecting, came from these two countries. Whilst this is considered in the thesis, some of the results might not be translatable to museums outside of those regions. The focus group participants were all drawn from the partner museums, except the Badisches, which was invited upon their request. Whilst the job roles overall are varied, they do not necessarily do so on an institutional level. Participants were invited to give evidence about their personal experiences, but the missing representation of some groups at institutions might have skewed the overall institution-specific picture. This can have drawbacks in terms of wider applicability and generalisability and, rather than the thesis being witness to institutional accounts, it is to highlight that the focus lies on the personal and professional experiences of participants.

8.3. Future opportunities

This thesis was conducted at an early stage of the museums and AI relationship and, whilst these are exciting times that feature the prospect of a forthcoming metaverse and Web 3.0, the actual deployment of AI in the sector is scarce and most museums are at a stage of orientation and “getting their head around AI”.

The thesis can function as a benchmark for research yet to come and a starting point from which other projects can build. The RS itself is fully open source and can therefore be used for future research or as a web application running in the real world. The data collected during the user study can be used to model and analyse interaction, and it would be interesting to test the models further and investigate the discrepancies between subjective perception and intrinsic interaction.

This research presents a very timely snapshot of the current museums and AI landscape; it also contains valuable insights for professionals, stakeholders, and all those interested in working with AI in a cultural setting. The findings will be applicable to a variety of research disciplines and can be used to inform future work concerned with digital media, data-intensive methods, and museum collections.

The thesis’ methodology, specifically the application of a postphenomenological framework to investigate AI in museums as part of practice-based, evolutionary process which developed and evaluated a technological artefact, can also be used as a model for future research projects, beyond the museum sphere and its many constituents.

References

Aagaard, J. (2017). Introducing postphenomenological research: A brief and selective sketch of phenomenological research methods. *International Journal of Qualitative Studies in Education*, *30*(6), 519–533. https://doi.org/10.1080/09518398.2016.1263884

Ada Lovelace Institute. (2021). *The data divide. Public attitudes to tackling social and health inequalities in the COVID-19 pandemic and beyond*. https://www.adalovelaceinstitute.org/wp-content/uploads/2021/03/The-data-divide_25March_final-1.pdf

Adcock, R. (2014). *'The Increase and Diffusion of Knowledge': Constructing the Relation of the Smithsonian Institution to Politics, 1835-1866*. Annual Meeting of the American Political Science Association, Washington D.C. https://doi.org/10.13140/2.1.3178.5922

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749. https://doi.org/10.1109/TKDE.2005.99

Aggarwal, C. C. (2016). *Recommender systems: The textbook* (1st edition). New York: Springer Science+Business Media.

Agostino, D., Arnaboldi, M., & Lema, M. D. (2020). New development: COVID-19 as an accelerator of digital transformation in public service delivery. *Public Money & Management*, 1–4. https://doi.org/10.1080/09540962.2020.1764206

Albanese, M., d'Acierno, A., Moscato, V., Persia, F., & Picariello, A. (2011). A Multimedia Semantic Recommender System for Cultural Heritage Applications. *2011 IEEE Fifth International Conference on Semantic Computing*, 403–410. https://doi.org/10.1109/ICSC.2011.47

Albanese, M., d'Acierno, A., Moscato, V., Persia, F., & Picariello, A. (2013). A Multimedia Recommender System. *ACM Transactions on Internet Technology*, *13*(1), 1–32. https://doi.org/10.1145/2532640

Anderson, M. L. (1999). Museums of the Future: The Impact of Technology on Museum Practices. *Daedalus*, *128*(3), 129–162.

Anderson, S. (2020). Some Provocations on the Digital Future of Museums. In K. Winesmith & S. Anderson (Eds.), *The Digital Future of Museums: Conversations and Provocations* (1st ed., pp. 10–25). Routledge. https://doi.org/10.4324/9780429491573

Aoki, P. M., Grinter, R. E., Hurst, A., Szymanski, M. H., Thornton, J. D., & Woodruff, A. (2002). Sotto Voce: Exploring the Interplay of Conversation and Mobile Audio Spaces. *CHI '02: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 431–438.

Art Fund. (2020). *COVID-19 impact. Museum sector research findings. Summary report.* https://www.artfund.org/assets/downloads/art-fund-covid19-research-report-final.pdf

Art Fund. (2021). *Looking ahead. Museum sector research findings. Summary report.* https://www.artfund.org/assets/downloads/looking-ahead-sector-research-report-2021.pdf

Arts Council England, & Nesta. (2020). *Digital Culture 2019: Museums*. ACE and Nesta. https://www.artscouncil.org.uk/sites/default/files/download-file/DC2019-Museums-factsheet.pdf

Arvanitis, K. (2004). Digital, Virtual, Cyber or Network Museum? Searching for the term and definition. *Proceedings of the First International Conference of Museology*. Museum, Communication and New TechnologiesInternational Conference of Museology, University of the Aegean.

Arvanitis, K., Gilmore, A., Florack, F., & Zuanni, C. (2016). Data culture and organisational practice. *MW2016: Museums and the Web 2016*, 17. https://mw2016.museumsandtheweb.com/paper/data-culture-and-organisational-practice/

Baker, C., Hutton, G., Christie, L., & Wright, S. (2020, December). *COVID-19 and the digital divide*. UK Parliament POST. https://post.parliament.uk/covid-19-and-the-digital-divide/

Basile, P., Calefato, F., de Gemmis, M., Lops, P., Bux, M., Musto, C., & Narducci, F. (2008). *Augmenting a Content-based Recommender System with Tags for Cultural Heritage Personalization*. 10.

Bassett, C., Berry, D. M., Fazi, M. B., Pay, J., & Roberts, B. (2017). Critical Digital Humanities and Machine Learning: Panel Statement. *Critical Digital Humanities and Machine Learning*, 6.

Battro, A. M. (2010). From Malraux's Imaginary Museum to the Virtual Museum. In R. Parry (Ed.), *Museums in a Digital Age* (pp. 136–147). Routledge.

Bearman, D. (2008). Representing Museum Knowledge. In P. F. Marty & K. Burton Jones (Eds.), *Museum Informatics. People, Information, and Technology in Museums* (pp. 35–57). Routledge.

Beaudoin, J. (2020). Art museum collections online: Extending their reach. *Proceedings of MW20: MW20*. MW20, Online. https://mw20.museweb.net/paper/art-museum-collections-online-extending-their-reach/

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, *11*(6), 985–1002. https://doi.org/10.1177/1461444809336551

Beer, D. (2016). *Metric Power*. London: Palgrave Macmillan.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Newhaven: Yale University Press.

Bennett, T. (2005). CIVIC LABORATORIES: Museums, cultural objecthood and the governance of the social. *Cultural Studies*, *19*(5), 521–547. https://doi.org/10.1080/09502380500365416

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*(3), 351–368. https://doi.org/10.1093/pan/mpr057

Berry, D. M. (2018). The Post-Digital. The New Aesthetic and Infrastructural Aesthetics. In M. Bühler (Ed.), *No Internet, No Art* (2nd ed.). Amsterdam: Onomatopee.

Berry, D. M., & Fagerjord, A. (2017). *Digital Humanities*. Cambridge: Polity.

Blanchette, J.-F. (2011). A material history of bits. *Journal of the American Society for Information Science and Technology*, *62*(6), 1042–1057. https://doi.org/10.1002/asi.21542

Blandford, A., Cox, A. L., & Cairns, P. (2008). Controlled experiments. In P. Cairns & A. L. Cox (Eds.), *Research Methods for Human-Computer Interaction* (pp. 1–16). Cambridge University Press.

Bloor, M., Frankland, J., Thomas, M., & Robson, K. (2002). *Focus Groups in Social Research*. London: SAGE.

Boididou, C., Sheng, D., Mercer Moss, F. J., & Piscopo, A. (2021). Building Public Service Recommenders: Logbook of a Journey. *Fifteenth ACM Conference on Recommender Systems*, 538–540. https://doi.org/10.1145/3460231.3474614

Bonacchi, C., & Krzyzanska, M. (2019). Digital heritage research re-theorised: Ontologies and epistemologies in a world of big data. *International Journal of Heritage Studies*, *25*(12), 1235–1247. https://doi.org/10.1080/13527258.2019.1578989

Bowen, J. P., Bennett, J., & Johnson, J. (1998). Virtual Visits to Virtual Museums. *MW98: Proceedings*. Museums and the Web 1998, Toronto, Canada. https://www.archimuse.com/mw98/papers/bowen/bowen_paper.html

boyd, danah, & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Buder, J., & Schwind, C. (2012). Learning with personalized recommender systems: A psychological view. *Computers in Human Behavior*, *28*(1), 207–216. https://doi.org/10.1016/j.chb.2011.09.002

Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, *12*, 331–370.

Byrne, A. (2014). *Memory institutions shaping the past, present and future*. IFLA PAC Geneva, 13 -14 August 2014. Maison de la Paix.

Cameron, F. (2010). Museum Collections, Documentation, and Shifting Knowledge Paradigms. In R. Parry (Ed.), *Museum in a Digital Age* (pp. 80–95). London: Routledge.

Cameron, F., & Kenderdine, S. (Eds.). (2010). *Theorizing Digital Cultural Heritage: A Critical Discourse (Media in Transition)*. Boston, MA: MIT Press.

Candy, L. (2006). *Practice Based Research: A Guide* (CCS Report: 2006-V1.0). University of Technology.

Carayon, P. (2006). Human factors of complex sociotechnical systems. *Applied Ergonomics*, *37*(4), 525–535. https://doi.org/10.1016/j.apergo.2006.04.011

Ciecko, B. (2020). AI Sees What? The Good, the Bad, and the Ugly of Machine Vision for Museum Collections. *MW20: MW20*, 29. https://mw20.museweb.net/paper/ai-sees-what-the-good-the-bad-and-the-ugly-of-machine-vision-for-museum-collections/index.html

Coeckelbergh, M. (2009). Virtual moral agency, virtual moral responsibility: On the moral significance of the appearance, perception, and performance of artificial agents. *AI & SOCIETY*, *24*(2), 181–189. https://doi.org/10.1007/s00146-009-0208-3

Costabile, M. F., Fogli, D., Letondal, C., Mussio, P., & Piccinno, A. (2003). Domain-Expert Users and their Needs of Software Development. *Proceedings of the HCI 2003 End User Development Session*. 2nd International Conference on Universal Access in Human-Computer Interaction, Crete, Greece. http://dx.doi.org/10.13140/2.1.4737.6325

Covington, P., Adams, J., & Sargin, E. (2016). Deep Neural Networks for YouTube Recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*, 191–198. https://doi.org/10.1145/2959100.2959190

Creative Commons. (2022). *CC0 'No Rights Reserved'*. https://creativecommons.org/share-your-work/public-domain/cc0/

Creative Industries Policy and Evidence Centre. (2020). *Digital Culture: Consumption in Lockdown. Insights from the Consumer Tracking Study* (Digital Culture - Consumer Tracking Study) [Thematic Report]. Nesta. https://www.pec.ac.uk/assets/images/The-PEC-and-the-IPO-cultural-consumption-study-insights-from-the-six-week-study.pdf

Crooke, E. (2011). Museums and Community. In Macdonald, Sharon (Ed.), *A Companion to Museum Studies* (pp. 170–185). Chichester: Wiley-Blackwell.

Darzentas, D., Cameron, H., Wagner, H., Craigon, P., Bodiaj, E., Spence, J., Tennent, P., & Benford, S. (2022). Data-inspired co-design for museum and gallery visitor experiences. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *36*, e3. https://doi.org/10.1017/S0890060421000317

David, R., & Kamerling, T. (2019). Relevancy Scoring for Knowledge-based Recommender Systems: *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 233–239. https://doi.org/10.5220/0008068602330239

DCMS. (2017). *Strategic review of DCMS-sponsored museums* [Strategic review]. Department for Digital, Culture, Media & Sport. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/673938/Strategic_review_of_DCMS-sponsored_museums.pdf

Devine, C. (2015). The Museum Digital Experience: Considering the Visitor's Journey. *MWA2015: Museums and the Web Asia 2015*. MWA2015, Melbourne. https://mwa2015.museumsandtheweb.com/paper/the-museum-digital-experience-considering-the-visitors-journey/

Doherty, K., & Doherty, G. (2019). Engagement in HCI: Conception, Theory and Measurement. *ACM Computing Surveys*, *51*(5), 1–39. https://doi.org/10.1145/3234149

Dudley, S. (2012). Materiality Matters: Experiencing the Displayed Object. *University of Michigan Working Papers in Museum Studies*, *8*, 1–9.

Economou, M. (2008). A World of Interactive Exhibits. In P. F. Marty & K. Burton Jones (Eds.), *Museum Informatics. People, Information, and Technology in Museums* (pp. 137–156). Routledge.

Edmonds, E. A., Weakley, A., Candy, L., Fell, M., Knott, R., & Pauletto, S. (2005). The studio as laboratory: Combining creative practice and digital technology research. *International Journal of Human-Computer Studies*, *63*(4–5), 452–481. https://doi.org/10.1016/j.ijhcs.2005.04.012

Ekstrand, M. D., Kluver, D., Harper, F. M., & Konstan, J. A. (2015). Letting Users Choose Recommender Algorithms: An Experimental Study. *Proceedings of the 9th ACM Conference on Recommender Systems*, 11–18. https://doi.org/10.1145/2792838.2800195

Ekstrand, M. D., Ludwig, M., Konstan, J. A., & Riedl, J. T. (2011). Rethinking the recommender research ecosystem: Reproducibility, openness, and LensKit. *RecSys '11: Proceedings of the Fifth ACM Conference on Recommender Systems*, 133–140. https://doi.org/10.1145/2043932.2043958

Elahi, M., Deldjoo, Y., Bakhshandegan Moghaddam, F., Cella, L., Cereda, S., & Cremonesi, P. (2017). Exploring the Semantic Gap for Movie Recommendations. *Proceedings of the Eleventh ACM Conference on Recommender Systems*, 326–330. https://doi.org/10.1145/3109859.3109908

Erdt, M., Fernandez, A., & Rensing, C. (2015). Evaluating Recommender Systems for Technology Enhanced Learning: A Quantitative Survey. *IEEE Transactions on Learning Technologies*, *8*(4), 326–344. https://doi.org/10.1109/TLT.2015.2438867

Falk, J. (2016). Museum audiences: A visitor-centered perspective. *Loisir et Société / Society and Leisure*, *39*(3), 357–370. https://doi.org/10.1080/07053436.2016.1243830

Falk, J. H., & Dierking, L. D. (2018). *Learning from Museums* (2nd ed.). Lanham: Rowman & Littlefield.

Falk, K. (2019). *Practical Recommender Systems*. Shelter Island, NY: Manning Publications Co.

Falkowski, J. (2016). Custom collections content and generous interfaces. *MW2016: Museums and the Web 2016*. MW2016, Los Angeles. https://mw2016.museumsandtheweb.com/paper/custom-collections-content-and-generous-interfaces/

Finnis, J., & Kennedy, A. (2020). *The Digital Transformation Agenda and GLAMs. A Quick Scan Report for Europeana*. Culture24 for Europeana.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Floridi, L., & Sanders, J. W. (2004). On the Morality of Artificial Agents. *Minds and Machine*, *14*, 349–379.

Flusser, V. (2013). *Post-History* (R. Maltez Novaes, Trans.). Minneapolis: Univocal Publishing.

French, A., & Villaespesa, E. (2019). AI, Visitor Experience, and Museum Operations: A Closer Look at the Possible. In S. Anderson, I. Bruno, H. Hethmon, S. Rao, E. Rodley, & R. Ropeik (Eds.), *Humanizing the Digital: Unproceedings from the MCN 2018 Conference* (pp. 101–113).

Fresa, A., Justrell, B., & Prandoni, C. (2015). Digital curation and quality standards for memory institutions: PREFORMA research project. *Archival Science*, *15*(2), 191–216. https://doi.org/10.1007/s10502-015-9242-8

Galani, A., & Chalmers, M. (2008). Blurring Boundaries for Museum Visitors. In P. F. Marty & K. Burton Jones (Eds.), *Museum Informatics. People, Information, and Technology in Museums* (pp. 157–177). Routledge.

Galani, A., & Kidd, J. (2020). Hybrid Material Encounters – Expanding the Continuum of Museum Materialities in the Wake of a Pandemic. *Museum and Society*, *18*(3), 298–301. https://doi.org/10.29311/mas.v18i3.3565

Gaskell, I. (2003). Sacred to Profane and Back Again. In A. McClellan (Ed.), *Art and its Publics* (pp. 149–164). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470775936.ch7

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2020). Datasheets for Datasets. *ArXiv:1803.09010 [Cs]*. http://arxiv.org/abs/1803.09010

Giannini, T., & Bowen, J. P. (Eds.). (2019). *Museums and digital culture*. Springer Berlin Heidelberg.

Gilliland-Swetland, A., & White, L. (2005). Museum Information Professionals as Providers and Users of Online Resources. *Bulletin of the American Society for Information Science and Technology*, *30*(5), 23–26.

Gilmore, A., Arvanitis, K., & Albert, A. (2018). "Never Mind the Quality, Feel the Width": Big Data for Quality and Performance Evaluation in the Arts and Cultural Sector and the Case of "Culture Metrics". In G. Schiuma & D. Carlucci (Eds.), *Big Data in the Arts and Humanities. Theory and Practice* (pp. 27–40). CRC Press.

Gomez-Uribe, C. A., & Hunt, N. (2016). The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Transactions on Management Information Systems*, *6*(4), 1–19. https://doi.org/10.1145/2843948

Hallinan, B., & Striphas, T. (2016). Recommended for you: The Netflix Prize and the production of algorithmic culture. *New Media & Society*, *18*(1), 117–137. https://doi.org/10.1177/1461444814538646

Hansen, J. S. (2019). Cutting Edge and Cutting Corners: Evolving Technology, Expanding Usership, and Responsive Solutions in a Museum Database. *Advances in Archaeological Practice*, *7*(3), 234–246. https://doi.org/10.1017/aap.2019.20

Harrison, M. (1985). Art and philanthropy: T.C. Horsfall and the Manchester Art Museum. In A. Kidd & K. Roberts (Eds.), *City class and culture* (pp. 120–147). Manchester University Press.

Hassenzahl, M., & Tractinsky, N. (2006). User experience—A research agenda. *Behaviour & Information Technology*, *25*(2), 91–97. https://doi.org/10.1080/01449290500330331

Hauser, S., Oogjes, D., Wakkary, R., & Verbeek, P.-P. (2018). An Annotated Portfolio on Doing Postphenomenology Through Research Products. *Proceedings of the 2018 Designing Interactive Systems Conference*, 459–471. https://doi.org/10.1145/3196709.3196745

Hein, H. (2011). The Matter of Museums. *The Journal of Museum Education*, *36*(2), 179–187. https://doi.org/10.1080/10598650.2011.11510698

Heumann Gurian, E. (2005). Threshold Fear. In S. MacLeod (Ed.), *Reshaping Museum Space. Architecture, design, exhibtion* (pp. 203–214). London: Routledge.

HMSO. (2020). *The National Gallery annual report and accounts for the year ended 31 March 2020. Annual Report*. https://www.nationalgallery.org.uk/media/35604/the-national-gallery-annual-report-and-accounts-19-20-web-version.pdf

Holmes, H., & Burgess, G. (2020, May 6). *Opinion: Coronavirus has intensified the UK's digital divide*. University of Cambridge. https://www.cam.ac.uk/stories/digitaldivide

Hooper-Greenhill, E. (1992). *Museums and the Shaping of Knowledge*. Abingdon: Routledge.

Hortensius, R., & Cross, E. S. (2018). From automata to animate beings: The scope and limits of attributing socialness to artificial agents: Socialness attribution and artificial agents. *Annals of the New York Academy of Sciences*, *1426*(1), 93–110. https://doi.org/10.1111/nyas.13727

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425. https://doi.org/10.1007/s10683-011-9273-9

Hughes-Noehrer, L. (2022a). *MuseREC. Models and Data. Code*. https://doi.org/10.5281/zenodo.7120756

Hughes-Noehrer, L. (2022b). Museum Professionals Focus Groups. Online Resource. *University of Manchester*. https://doi.org/10.48420/21207827.v1

Hughes-Noehrer, L. (2022c). *Museum Recommender System (MuseREC) Web APP. Code*. https://doi.org/10.5281/zenodo.7120742

Hughes-Noehrer, L. (2022d). *Museums and AI Application (MAIA) Survey Analysis* [Jupyter Notebook running Python].

Hughes-Noehrer, L. (2022e). *Smithsonian Metadata Cleaner. Code*. https://doi.org/10.5281/zenodo.7120698

Hughes-Noehrer, L. (2022f). UX/UI Online Study. Online Resource. *University of Manchester*. https://doi.org/10.48420/21207851.v1

Hughes-Noehrer, L. (2022g). *UX/UI Online User Study. Analytics and Data. Code*. https://doi.org/10.5281/zenodo.7120767

Hughes-Noehrer, L., Jay, C., & Gilmore, A. (2022a). Museums and AI Applications (MAIA) Survey. *University of Manchester*. https://doi.org/10.48420/19298588.v1

Hughes-Noehrer, L., Jay, C., & Gilmore, A. (2022b). Museums and AI Applications (MAIA) Survey. Dataset. *University of Manchester*. https://doi.org/10.48420/19298588.v1

ICOM. (2020a). *Museums, museum professionals and COVID-19*. International Council of Museums. https://icom.museum/wp-content/uploads/2020/05/Report-Museums-and-COVID-19.pdf

ICOM. (2020b). *Museums, museum professionals and COVID-19: Follow-up survey*. International Council of Museums. https://icom.museum/wp-content/uploads/2020/11/FINAL-EN_Follow-up-survey.pdf

Ihde, D. (1979). *Technics and Praxis* (Vol. 24). Dordrecht: Springer.

Ihde, D. (1990). *Technology and the Lifeworld*. Bloomingdale: Indiana University Press.

Ihde, D. (2012). *Experimental Phenomenology. Multistabilities* (2nd ed.). New York: State University of New York.

International Organization for Standardization [ISO]. (2010). *ISO 9241-210:2010(en) 2.15*. https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en

Jacobson, K., Murali, V., Newett, E., Whitman, B., & Yon, R. (2016). Music Personalization at Spotify. *Proceedings of the 10th ACM Conference on Recommender Systems*, 373–373. https://doi.org/10.1145/2959100.2959120

Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York: New York University Press.

Jensen, M. M., & Aagaard, J. (2018). A postphenomenological method for HCI research. *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, 242–251. https://doi.org/10.1145/3292147.3292170

Keller, I., & Viennet, E. (2015). Recommender Systems for Museums: Evaluation on a Real Dataset. *Proceedings of the Fifth Conference on Advances in Information Mining and Management*, 65–71.

Kenderdine, S. (2016). Embodiment, Entanglement, and Immersion in Digital Cultural Heritage. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A New Companion to Digital Humanities*. Wiley.

Kidd, J. (2016). *Museums in the New Mediascape. Transmedia, Participation, Ethics.* Abingdon: Routledge.

Kidd, J. (2019). Digital media ethics and museum communication. In K. Drotner, V. Dziekan, R. Parry, & K. C. Schrøder (Eds.), *The Routledge Handbook of Museums, Media and Communication* (1st ed., pp. 193–204). Routledge. https://doi.org/10.4324/9781315560168

Kislyuk, D., Liu, Y., Liu, D., Tzeng, E., & Jing, Y. (2015). Human Curation and Convnets: Powering Item-to-Item Recommendations on Pinterest. *ArXiv:1511.04003 [Cs]*. http://arxiv.org/abs/1511.04003

Kitchin, R. (2014). *The Data Revolution. Big Data, Open Data, Data Infrastructures & Their Consequences.* SAGE.

Knabb, K. (n.d.). *Theory of the Dérive*. Situationist International Online. Retrieved 2 September 2022, from https://www.cddc.vt.edu/sionline/si/theory.html

Knabb, K. (2006). *Situationist International Anthology. Revised and Expanded Edition*. Bureau of Public Secrets.

Knijnenburg, B. P., & Willemsen, M. C. (2015). Evaluating Recommender Systems with User Experiments. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 309–352). Springer US. https://doi.org/10.1007/978-1-4899-7637-6_9

Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012). Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, *22*(4–5), 441–504. https://doi.org/10.1007/s11257-011-9118-4

Knorr Cetina, K. (2005). *The Practice Turn in Contemporary Theory.* (T. R. Schatzki, E. von Savigny, & K. Knorr-Cetina, Eds.). Routledge. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=235322

Kohavi, R., Deng, A., Frasca, B., Walker, T., Xu, Y., & Pohlmann, N. (2013). Online controlled experiments at large scale. *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1168–1176. https://doi.org/10.1145/2487575.2488217

Kohavi, R., Longbotham, R., Sommerfield, D., & Henne, R. M. (2009). Controlled experiments on the web: Survey and practical guide. *Data Mining and Knowledge Discovery*, *18*(1), 140–181. https://doi.org/10.1007/s10618-008-0114-1

Kontiza, K., Loboda, O., Deladiennee, L., Castagnos, S., & Naudet, Y. (2018). A Museum App to Trigger Users' Reflection. *CEUR Workshop on Mobile Access to Cultural Heritage*, 8.

Koukoulis, K., Koukopoulos, D., & Tzortzi, K. (2019). Connecting the museum to the city environment from the visitor's perspective. *Applied Computing and Informatics*, S2210832719301991. https://doi.org/10.1016/j.aci.2019.09.001

Kuflik, T., Stock, O., Zancanaro, M., Gorfinkel, A., Jbara, S., Kats, S., Sheidin, J., & Kashtan, N. (2011). A visitor's guide in an active museum: Presentations, communications, and reflection. *Journal on Computing and Cultural Heritage*, *3*(3), 1–25. https://doi.org/10.1145/1921614.1921618

Kuutti, K., & Bannon, L. J. (2014). The turn to practice in HCI: Towards a research agenda. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3543–3552. https://doi.org/10.1145/2556288.2557111

Kvale, S. (2007). *Doing Interviews*. London: SAGE.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction* (2nd ed.). Morgan Kaufmann Publishers.

Lehmann, J., Lalmas, M., Yom-Tov, E., & Dupret, G. (2012). Models of User Engagement. In J. Masthoff, B. Mobasher, M. C. Desmarais, & R. Nkambou (Eds.), *User Modeling, Adaptation, and Personalization* (Vol. 7379, pp. 164–175). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31454-4_14

Lindlof, T. R., & Taylor, B. C. (2017). *Qualitative Communication Research Methods* (4th ed.). SAGE.

Loboda, O., Nyhan, J., Mahony, S., Romano, D. M., & Terras, M. (2019). Content-based Recommender Systems for Heritage: Developing a Personalised Museum Tour. *Proceedings of the 1st International 'Alan Turing' Conference on Decision Support and Recommender Systems (DSRS-Turing 2019)*, 7.

Loepp, B., & Ziegler, J. (2019). Measuring the Impact of Recommender Systems – A Position Paper on Item Consumption in User Studies. *Proceedings of the 1st Workshop on Impact of Recommender Systems*, 3.

Luyten, K., Loon, H. V., Teunkens, D., Gabriëls, K., Coninx, K., & Manshoven, E. (2006). ARCHIE: Disclosing a Museum by a Socially-aware Mobile Guide. *The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST*, 6.

Maeda, J. (2020). *2020 CX Report*. CX Report. https://cx.report/2020-cxreport/

Manchester Art Gallery. (n.d.). *Our mission and vision*. Retrieved 21 August 2021, from https://manchesterartgallery.org/visit/about-us/our-vision-and-mission/

Manchester City Council. (2019a). *Manchester Art Gallery Update: Manchester City Council, Report for Information Communitiesand Equalities Scrutiny Committee. 10 October 2019*. https://democracy.manchester.gov.uk/mgConvert2PDF.aspx?ID=10650#:~:text=Manchester%20Art%20Gallery%20welcomed%20731%2C002,museum%20or%20gallery%20in%20Manchester

Manchester City Council. (2019b). *Manchester City Galleries' report and revenue budget 2019/20: Art Galleries Committee. 13 February 2019*. https://democracy.manchester.gov.uk/documents/s4573/Art%20Galleries%20Committee%202019.pdf

Manchester City Council. (2020). *Reports Spell Out Council Budget Position. 29 October 2020*. https://secure.manchester.gov.uk/news/article/8592/reports_spell_out_councils_budget_position_-_and_options_to_address_it

Manovich, L. (1999). Database as Symbolic Form. *Convergence: The International Journal of Research into New Media Technologies*, *5*(2), 80–99. https://doi.org/10.1177/135485659900500206

Manovich, L. (2002). *The Language of New Media*. Boston, MA: MIT Press.

Manovich, L. (2019). *AI Aesthetics*. Moscow: Strelka Press.

Marty, P. F. (2008a). Interactive Technologies. In P. F. Marty & K. Burton Jones (Eds.), *Museum Informatics. People, Information, and Technology in Museums* (pp. 131–135). Routledge.

Marty, P. F. (2008b). Museum websites and museum visitors: Digital museum resources and their use. *Museum Management and Curatorship*, *23*(1), 81–99. https://doi.org/10.1080/09647770701865410

McNee, S. M., Riedl, J., & Konstan, J. A. (2006). Making recommendations better: An analytic model for human-recommender interaction. *CHI '06 Extended Abstracts on Human Factors in Computing Systems - CHI EA '06*, 1103. https://doi.org/10.1145/1125451.1125660

Merritt, E. (2017). *Artificial Intelligence The Rise of the Intelligent Machine*. https://www.aam-us.org/2017/05/01/artificial-intelligence-the-rise-of-the-intelligent-machine/

Merritt, E. (2021). *TrendsWatch. Navigating a Disrupted Future*. American Alliance of Museums.

Milano, S., Taddeo, M., & Floridi, L. (2020). Recommender systems and their ethical challenges. *AI & SOCIETY*, *35*(4), 957–967. https://doi.org/10.1007/s00146-020-00950-y

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 2053951716679679. https://doi.org/10.1177/2053951716679679

Moens, B. G. (2018). Aesthetic Experience in Virtual Museums: A Postphenomenological Perspective. *Studies in Digital Heritage*, *2*(1), 68–79. https://doi.org/10.14434/sdh.v2i1.24468

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, *33*(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Morris, M. R. (2020). AI and accessibility. *Communications of the ACM*, *63*(6), 35–37. https://doi.org/10.1145/3356727

Mummolo, J., & Peterson, E. (2019). Demand Effects in Survey Experiments: An Empirical Assessment. *American Political Science Review*, *113*(2), 517–529. https://doi.org/10.1017/S0003055418000837

Munley, M. E., Roberts, R. C., Soren, B., & Hayward, J. (2007). Envisioning the Customized Museum: An Agenda to Guide Reflective Practice and Research. In J. H. Falk, L. D. Dierking, & S. Foutz (Eds.), *In Principle, In Practice. Museums as Learning Institutions* (pp. 77–90). AltaMira Press.

Murphy, O., & Villaespesa, E. (2020). *AI: A Museum Planning Toolkit | The Museums + AI Network*. Goldsmiths, University of London. https://themuseumsai.network/toolkit/

National Gallery. (n.d.). *Access policy*. Retrieved 22 April 2021, from https://www.nationalgallery.org.uk/about-us/organisation/policies/access-policy

National Gallery. (2018). *Strategic Plan 2018-2023*. https://www.nationalgallery.org.uk/media/25328/strategic-plan_2018-2023.pdf

Naudet, Y., & Deladiennée, L. (2019). Towards technology-mediated CH experiences: Some open challenges. *CI@SMAP*, 5.

Navarrete, T. (2013). Digital Cultural Heritage. In I. Rizzo & A. Mignosa (Eds.), *Handbook on the Economics of Cultural Heritage*. Edward Elgar Publishing. https://doi.org/10.4337/9780857931009

Newell, J. (2012). Old objects, new media: Historical collections, digitization and affect. *Journal of Material Culture*, *17*(3), 287–306. https://doi.org/10.1177/1359183512453534

Newman, T., Beetham, H., & Church, S. (2020). *DASH Survey Results 2020. Describing the digital attitudes, skills and organisational suppport of people working across the UK heritage sector* (p. 61). Timmus Research and Heritage Fund.

Noehrer, L., Carlton, J., & Jay, C. (2021a). Machine Learning and Museum Collections: A Data Conundrum. In: Shehade, M., & Stylianou-Lambert, T. (eds.) *Emerging Technologies and the Digital Transformation of Museums and Heritage Sites.* RISE IMET 2021. Communications in Computer and Information Science, 1432. Cham: Springer. https://doi.org/10.1007/978-3-030-83647-4_2

Noehrer, L., Gilmore, A., Jay, C., & Yehudi, Y. (2021b). The impact of COVID-19 on digital data practices in museums and art galleries in the UK and the US. *Humanities and Social Sciences Communications*, *8*(1), 236. https://doi.org/10.1057/s41599-021-00921-8

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for Reporting Qualitative Research: A Synthesis of Recommendations. *Academic Medicine*, *89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

O'Brien, H. L., & Lebow, M. (2013). Mixed-methods approach to measuring user experience in online news interactions. *Journal of the American Society for Information Science and Technology*, *64*(8), 1543–1556. https://doi.org/10.1002/asi.22871

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938–955. https://doi.org/10.1002/asi.20801

O'Neal Irwin, S. (2016). *Digital Media. Human-Technology Connection*. Lanham: Lexington Books.

O'Reilly, T. (2005). *What Is Web 2.0*. https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

Oumaima, S., Soulaimane, K., & Omar, B. (2020). Latest Trends in Recommender Systems applied in the medical domain: A Systematic Review. *Proceedings of the 3rd International Conference on Networking, Information Systems & Security*, 1–12. https://doi.org/10.1145/3386723.3387860

Padilla, T. (2017). *On a Collections as Data Imperative*. UC Santa Barbara. https://escholarship.org/uc/item/9881c8sv

Parry, R. (2005). Digital heritage and the rise of theory in museum computing. *Museum Management and Curatorship*, *20*(4), 333–348. https://doi.org/10.1080/09647770500802004

Parry, R. (2007). *Recoding the Museum: Digital Heritage and the Technologies of Change* (1st ed.). London: Routledge. https://doi.org/10.4324/9780203347485

Parry, R. (2013). The End of the Beginning. *Museum Worlds*, *1*(1), 24–39. https://doi.org/10.3167/armw.2013.010103

Parry, R. (2019). How Museums Made (and Re-made) Their Digital User. In T. Giannini & J. Bowen (Eds.), *Museums and Digital Culture* (pp. 275–293). Springer.

Parry, R., & Sawyer, A. (2005). Space and the machine. Adaptive museums, pervasive technology and the new gallery environment. In S. MacLeod (Ed.), *Reshaping Museum Space. Architecture, design, exhibitions* (pp. 39–52). Routledge.

Pasquale, F. (2016). *The Black Box Society. The Secret Algorithms That Control Money and Information*. Harvard University Press.

Paul, C. (2006). Flexible Contexts, Democratic Filtering, and Computer Aided Curating—Models for Online Curatorial Practice. In J. Krysa (Ed.), *Curating, Immateriality, Systems: On Curating Digital Media* (Vol. 3). Autonomedia.

Paul, C. (2018). Curation I Context I Archive. Presenting and Preserving New Media Art. In M. Bühler (Ed.), *No Internet, No Art* (pp. 83–94). Amsterdam: Onomatopee.

Pavlidis, G. (2019). Recommender systems, cultural heritage applications, and the way forward. *Journal of Cultural Heritage*, *35*, 183–196. https://doi.org/10.1016/j.culher.2018.06.003

Pechenizkiy, M., & Calders, T. (2007). A Framework for Guiding the Museum Tours Personalization. *Proceedings of the Workshop on Personalised Access to Cultural Heritage (Patch07)*, 11–28.

Perin, C. (1992). Chapter 7. The Communicative Circle: Museums as Communities. In I. Karp, C. M. Kreamer, S. Lavine, & Woodrow Wilson International Center for Scholars (Eds.), *Museums and communities: The politics of public culture* (pp. 182–220). Smithsonian Institution Press.

Perugini, S., Gonçalves, M. A., & Fox, E. A. (2004). Recommender Systems Research: A Connection-Centric Survey. *Journal of Intelligent Information Systems*, *23*(2), 107–143. https://doi.org/10.1023/B:JIIS.0000039532.05533.99

Pierroux, P. (2019). Learning and engagement in museum mediascapes. In K. Drotner, V. Dziekan, R. Parry, & K. C. Schrøder (Eds.), *The Routledge Handbook of Museums, Media and Communication* (1st ed., pp. 128–142). Routledge. https://doi.org/10.4324/9781315560168

Pisoni, G., Díaz-Rodríguez, N., Gijlers, H., & Tonolli, L. (2021). Human-Centered Artificial Intelligence for Designing Accessible Cultural Heritage. *Applied Sciences*, *11*(2), 870. https://doi.org/10.3390/app11020870

Portolano, M. (1999). Increase and diffusion of knowledge: Ethos of science and education in the Smithsonian's inception. *Rhetoric Review*, *18*(1), 65–81. https://doi.org/10.1080/07350199909359256

Powered by AI: Instagram's Explore recommender system. (2019). Meta AI. https://ai.facebook.com/blog/powered-by-ai-instagrams-explore-recommender-system/

Pruulmann-Vengerfeldt, P., & Aljas, A. (2014). Digital Cultural Heritage—Challenging Museums, Archives and Users. In P. Runnel & P. Pruulmann-Vengerfeldt (Eds.), *Democratising the Museum: Reflections on Participatory Technologies* (pp. 163–216). Peter Lang GmbH, Internationaler Verlag der Wissenschaften.

QSR International Pty Ltd. (2020). *NVivo* (Version 12). QSR International Pty Ltd. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Qualtrics. (2021). *Qualtrics* (August '21). Qualtrics. https://www.qualtrics.com/uk/

Quaranta, D. (2018). Domenico Quaranta on Authorship, Appropriation, Surfing Clubs and Post-Internet. In M. Bühler (Ed.), *No Internet, No Art* (pp. 51–56). Amsterdam: Onomatopee.

Rader, E., Cotter, K., & Cho, J. (2018). Explanations as Mechanisms for Supporting Algorithmic Transparency. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–13. https://doi.org/10.1145/3173574.3173677

Rastogi, R. (2018). Machine Learning @ Amazon. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 1337–1338. https://doi.org/10.1145/3209978.3210211

Reips, U.-D. (2002). Standards for Internet-Based Experimenting. *Experimental Psychology (Formerly Zeitschrift Für Experimentelle Psychologie)*, *49*(4), 243–256. https://doi.org/10.1027//1618-3169.49.4.243

Ritter, M. (2021). Postphenomenological Method and Technological Things Themselves. *Human Studies*, *44*(4), 581–593. https://doi.org/10.1007/s10746-021-09603-5

Rodney, S., & Stein, R. J. (2020). Conversation 1. Seph Rodney and Robert J. Stein. In K. Winesmith & S. Anderson (Eds.), *The Digital Future of Museums. Conversations and Provocations* (pp. 29–42). Routledge.

Rosenberger, R. (2014). Multistability and the Agency of Mundane Artifacts: From Speed Bumps to Subway Benches. *Human Studies*, *37*(3), 369–392. https://doi.org/10.1007/s10746-014-9317-1

Rosenberger, R. (2017). Notes on a Nonfoundational Phenomenology of Technology. *Foundations of Science*, *22*(3), 471–494. https://doi.org/10.1007/s10699-015-9480-5

Rosenberger, R., & Verbeek, P.-P. (2015). A Field Guide to Postphenomenolgy. In R. Rosenberger & P.-P. Verbeek (Eds.), *Postphenomenological Investigations. Essays on Human-Technology Relations* (pp. 7–42). Lexington Books.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative Interviewing: The Art of Hearing Data* (2nd ed.). London: SAGE.

Ruotsalo, T., Haav, K., Stoyanov, A., Roche, S., Fani, E., Deliai, R., Mäkelä, E., Kauppinen, T., & Hyvönen, E. (2013). SMARTMUSEUM: A mobile recommender system for the Web of Data. *Journal of Web Semantics*, *20*, 50–67. https://doi.org/10.1016/j.websem.2013.03.001

Russell, S. (2016). Should We Fear Supersmart Robots? *Scientific American*, *314*(6), 58–59. https://doi.org/10.1038/scientificamerican0616-58

Samaroudi, M., Echavarria, K. R., & Perry, L. (2020). Heritage in lockdown: Digital provision of memory institutions in the UK and US of America during the COVID-19 pandemic. *Museum Management and Curatorship*, *35*(4), 337–361. https://doi.org/10.1080/09647775.2020.1810483

Sandell, R. (1998). Museums as Agents of Social Inclusion. *Museum Management and Curatorship*, *17*(4), 401–418. https://doi.org/10.1080/09647779800401704

Sandell, R. (2005). Constructing and communicating equality. The social agency of museum space. In S. MacLeod (Ed.), *Reshaping Museum Space. Architecture, design, exhibitions* (pp. 185–200). Routledge.

Semeraro, G., Lops, P., De Gemmis, M., Musto, C., & Narducci, F. (2012). A folksonomy-based recommender system for personalized access to digital artworks. *Journal on Computing and Cultural Heritage*, *5*(3), 1–22. https://doi.org/10.1145/2362402.2362405

Simon, N. (2010). *The Participatory Museum*. San Jose, CA: Museum 20.

Skov, M., & Ingwersen, P. (2008). Exploring information seeking behaviour in a digital museum context. *Proceedings of the Second International Symposium on Information Interaction in Context - IIiX '08*, 110. https://doi.org/10.1145/1414694.1414719

Smith, B., & Linden, G. (2017). Two Decades of Recommender Systems at Amazon.com. *IEEE Internet Computing*, *21*(3), 12–18. https://doi.org/10.1109/MIC.2017.72

Smithsonian Institution. (n.d.). *Smithsonian Open Access*. Retrieved 27 March 2021, from https://www.si.edu/openaccess

Smithsonian Institution. (2014, Summer). *Delivering on the promise of the Digital Smithsonian*. Reports and Plans. https://www.si.edu/content/pdf/about/SmithsonianDigitalActionAgenda.pdf

Smithsonian Institution. (2017). *Strategic Plan: Smithsonian 2022*. Strategic Plan. https://www.si.edu/sites/default/files/about/smithsonian-2022-strategic-plan.pdf

Solas, N. (2010). Hiding Our Collections in Plain Site: Interface Strategies for 'Findability'. In J. Trant & D. Bearman (Eds.), *Museums and the Web 2010: Proceedings*. Archives & Museum Informatics. http://www.archimuse.com/mw2010/papers/solas/solas.html

Speakman, R., Hall, M. M., & Walsh, D. (2018). User Engagement with Generous Interfaces for Digital Cultural Heritage. In E. Méndez, F. Crestani, C. Ribeiro, G. David, & J. C. Lopes (Eds.), *Digital Libraries for Open Knowledge* (Vol. 11057, pp. 186–191). Springer International Publishing. https://doi.org/10.1007/978-3-030-00066-0_16

Stack, J. (2018). *Exploring museum collections online: Some background reading*. Science Museum Group Digital Lab. https://lab.sciencemuseum.org.uk/exploring-museum-collections-online-some-background-reading-da5a332fa2f8

Stash, N. (2010). *CHIP Project @ Rijksmuseum Amsterdam. Cultural Heritage Information Personalization*. https://wwwis.win.tue.nl/~nstash/presentations/siks090610.pdf

Stock, O., & Zancanaro, M. (Eds.). (2007). *PEACH: Intelligent interfaces for museum visits*. Springer.

Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Krüger, A., Kruppa, M., Kuflik, T., Not, E., & Rocchi, C. (2007). Adaptive, intelligent presentation of information for the museum visitor in PEACH. *User Modeling and User-Adapted Interaction*, *17*(3), 257–304. https://doi.org/10.1007/s11257-007-9029-6

Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, *18*(4–5), 395–412. https://doi.org/10.1177/1367549415577392

Su, X., Sperlì, G., Moscato, V., Picariello, A., Esposito, C., & Choi, C. (2019). An Edge Intelligence Empowered Recommender System Enabling Cultural Heritage Applications. *IEEE Transactions on Industrial Informatics*, *15*(7), 4266–4275. https://doi.org/10.1109/TII.2019.2908056

Suresh, H., & Guttag, J. V. (2020). A Framework for Understanding Unintended Consequences of Machine Learning. *ArXiv:1901.10002 [Cs, Stat]*. http://arxiv.org/abs/1901.10002

Taylor, J. K. (2020). *The art museum redefined: Power, opportunity, and community engagement*. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-21021-2

The Audience Agency. (2020a). *Who are museum digital visitors?* Report. https://www.theaudienceagency.org/asset/2440

the audience agency. (2020b). *The Audience Agency COVID-19 Cultural Participation Monitor Digital Findings*. https://www.theaudienceagency.org/asset/2549

The pandas development team. (2020). *pandas-dev/pandas: Pandas* (1.4.3). https://doi.org/10.5281/zenodo.3509134

The Royal Society. (2017). *Machine Learning: The Power and Promise of Computers that Learn by Example.*

The SciPy Dev Team. (2020). *SciPy Fundamental Algorithms for Scientific Computing in Python* (1.7.3.). 10.1038/s41592-019-0686-2

Travkina, E., & Sacco, P. (2020). *Culture shock: COVID-19 and the cultural and creative sectors* (Tackling Coronavirus (COVID-19): Contributing to a Global Effort, p. 55). OECD.

Vallat, R. (2018). *Pingouin: Statistics in Python* (0.5.2). https://doi.org/10.21105/joss.01026

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, *4*(2), 2053951717743530. https://doi.org/10.1177/2053951717743530

Verbeek, P.-P. (2005). *What Things Do. Philosophical Reflections on Technology, Agency, and Design.* University Park: The Pennsylvania State University Press.

Verbeek, P.-P. (2016). Toward a Theory of Technological Mediation. A Program for Postphenomonological Research. In J. K. Berg, O. Friis, & R. C. Crease (Eds.), *Technoscience and Postphenomenology: The Manhattan Papers* (pp. 189–204). Lanham: Lexington Books.

Vergo, P. (Ed.). (1989). *The New Museology*. London: Reaktion Books.

Villaespesa, E. (2019). Museum Collections and Online Users: Development of a Segmentation Model for the Metropolitan Museum of Art. *Visitor Studies*, *22*(2), 233–252. https://doi.org/10.1080/10645578.2019.1668679

Villaespesa, E., & Stack, J. (2015). Finding the motivation behind a click: Definition and implementation of a website audience segmentation. *Proceedings of MW2015: Museums and the Web 2015*. MW2015, Ontario. https://mw2015.museumsandtheweb.com/paper/finding-the-motivation-behind-a-click-definition-and-implementation-of-a-website-audience-segmentation/

Walsh, D., Hall, M. M., Clough, P., & Foster, J. (2020). Characterising online museum users: A study of the National Museums Liverpool museum website. *International Journal on Digital Libraries*, *21*(1), 75–87. https://doi.org/10.1007/s00799-018-0248-8

Wang, Y., Aroyo, L. M., Stash, N., & Rutledge, L. (2007). Interactive User Modeling for Personalized Access to Museum Collections: The Rijksmuseum Case Study. In C. Conati, K. McCoy, & G. Paliouras (Eds.), *User Modeling 2007* (Vol. 4511, pp. 385–389). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-73078-1_50

Wang, Y., Stash, N., Aroyo, L., Gorgels, P., Rutledge, L., & Schreiber, G. (2008). Recommendations based on semantically enriched museum collections. *Journal of Web Semantics*, *6*(4), 283–290. https://doi.org/10.1016/j.websem.2008.09.002

Whitelaw, M. (2015). Generous Interfaces for Digital Cultural Collections. *Digital Humanities Quarterly*, *9*(1), 15.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Wiltse, H. (2017). Mediating (Infra)structures. Technology, Media, Environment. In Y. Van den Eede, S. O'Neal Irwin, & G. Wellner (Eds.), *Postphenomenology and Media. Essays on Human-Media-World Relations* (pp. 3–27). Lexington Books.

Witcomb, A. (2007). The Materiality of Virtual Technologies: A New Approach to Thinking about the Impact of Multimedia in Museums. In F. Cameron & S. Kenderdine (Eds.), *Theorizing Digital Cultural Heritage* (pp. 35–48). The MIT Press. https://doi.org/10.7551/mitpress/9780262033534.003.0003

Wolff, J. (2013). Calico connections: Science, manufacture and culture in mid-nineteenth-century Manchester. In J. Wolff & M. Savage (Eds.), *Culture in Manchester Institutions and urban change since 1850*. Manchester University Press.

Zoom Video Communications, Inc. (2022). *Zoom Cloud Meetings* (Version 5). https://zoom.us

APPENDIX A

MAIA Survey

Q1
Welcome to the Museums and AI Applications (MAIA) Survey

Thank you for your time spent answering this survey and your valuable contribution to the community!
I have read and understood the Participant Information Sheet [available here] and agree to take part in this survey:

- Yes (1)
- No (2)

Skip To: End of Survey If Welcome to the Museums and AI Applications (MAIA) Survey Thank you for your time spent answering... != Yes

Q2 Are you a museum professional (this includes art galleries) working in either a leadership, curatorial, collection management, technical/digital, conservation/restoration, educational, or research role?

- Yes (1)
- No (2)

Skip To: End of Survey If Are you a museum professional (this includes art galleries) working in either a leadership, curat... != Yes

Q3 Are you using Artificial Intelligence* (AI) applications for your work?
**Artificial Intelligence includes symbolic (e.g. rule-based, semantic technologies, and knowledge graphs) and non-symbolic systems (e.g. machine learning, deep learning, neural networks, natural language processing, and data mining)*

- Yes, I am currently using it (1)
- Used it in the past (2)
- I intend to use it in the next 12 months (3)
- No, I am not (4)

Display This Question:

If Are you using Artificial Intelligence (AI) applications for your work?*Artificial Intelligence i... = No, I am not*

Q4 Can you tell us the main reasons why you have not used AI applications so far?

- ▢ Too expensive (1)
- ▢ Lack of staff (2)
- ▢ Lack of skills/training (3)
- ▢ Missing technology (4)
- ▢ No need (5)
- ▢ Other: (6) ____

Display This Question:

If Are you using Artificial Intelligence (AI) applications for your work?*Artificial Intelligence i... = No, I am not*

Q5 Are you personally interested in using AI for your work?

- ○ Yes (1)
- ○ Not really sure (2)
- ○ No (3)
- ○ I do not know (4)

Q6 Is someone else/an external partner using AI applications at the institution you are working for?

- ▢ Yes, internal (1)
- ▢ Yes, external (2)
- ▢ No (3)
- ▢ I do not know (4)

Display This Question:

If Is someone else/an external partner using AI applications at the institution you are working for? = No

Q7 Can you tell us why someone else in your institution/externals might not have used AI so far?

- ▢ Too expensive (1)
- ▢ Lack of staff (2)
- ▢ Lack of skills/training (3)
- ▢ Missing technology (4)
- ▢ No need (5)
- ▢ Other: (6

Q8 Please describe how you or someone else are using/will use AI in your institution and tell us about the areas of application (e.g. computer vision, classification, visitor metrics, clustering, technical art history, production of exhibits etc.)

Q9 You have indicated that you or the institution you are working for have been using AI techniques. Was that:

- On a project basis (1)
- A long-term integration (2)
- Just a trial (3)
- I do not know (4)

Q10 Was your experience working with AI positive or negative? Please explain.

Q11 Would you describe the application as successful?

- Yes (1)
- No (2)
- Not applicable (3)
- I do not know (4)

Q12 If it was successful, do you still use the application?

- Yes (1)
- No (2)
- Sometimes (3)
- Not applicable (4)
-

Display This Question:

If If it was successful, do you still use the application? = No

Q13 Why did you stop using the application?

Q14 If you have worked on an application with a partner (e.g. university, other museums, freelancer) would you describe the collaboration as positive or negative? Please give reasons.

Q15 Do you think AI applications can enhance the visitor experience on-site?

- Yes (1)
- Probably in the future (2)
- No (3)
- I do not know (4)

Q37 Do you think AI applications can enhance the user experience of museum content online?

- Yes (1)
- Probably in the future (2)
- No (3)
- I do not know (4)

Q16 Do you think the current digital capacity of your institution is sufficient to tackle the challenges of the future?

- Yes, it is (1)
- Mostly, needs minor updates (2)
- Partially, needs major updates (3)
- No (4)
- I do not know (5)

Q17 What type of software are you using to store, modify, and query information about the collections you are working with?

Q18 Do you think the software aforementioned is suitable for your work? I.e. do you find it easy to store and find relevant information about the collection.

- Yes (1)
- No (2)
- I do not know (3)
- I can do my work, but there are shortfalls/things to improve. Please describe: (4)

Q19 Are you working with ontologies (e.g. CIDOC CRM) or other methods of knowledge representation?

- Yes (1)
- No. Please explain why: (2)

Q20 Do you think current databases are suitable to describe museum information and knowledge?

- Yes, suitable (1)
- Somehow, but needs improvement. Please describe: (2)
- No, they are not suitable. Please describe: (3)
- I do not know (4)

Q21 Does your museum employ a data scientist?

- Yes, 1-5 (1)
- Yes, 6+ (2)
- Used to employ one (3)
- No (4)
- I do not know (5)

Q22 Does your museum employ a dedicated AI engineer/machine learning specialist?

- Yes, 1-5 (1)
- Yes, 6+ (2)
- Used to employ someone (3)
- No (4)
- I do not know (5)

Q23 Do you think there is a dissonance between the actual daily use of technology in museums and its perception by the senior leadership team (those are people who make strategic decisions in your institution)?

- Yes, there is a dissonance (1)
- Sometimes, depending on the application (2)
- No, SLT and daily operations are aligned (3)
- I do not know (4)

Q24 In your own words, what kind of tasks in museums can be supported by AI in the future?

Q25 Please indicate below - ranging from 'not at all' to 'very much' - if in your view AI technologies are **generally** negatively impacted by:

Not at all		Modest impact		Very high impact
0	1	2	3	4

Issues around bias ()	
Lack of explainable/interpretable technologies ()	
Effects on society ()	
Costs ()	
Environmental concerns ()	
Insufficient education about AI ()	
Other (please describe): ()	

Q26 How much impact do the below variables have on the uptake of AI applications in museums?

	No impact		Modest impact		Very high impact
	0	1	2	3	4
Funding ()					
Skills shortage ()					
Technology in museums ()					
Lack of strategy ()					
Lack of explainable technologies ()					
No need for AI applications in the sector ()					
Society is not ready yet ()					
Organisational culture ()					
Other: ()					

Q27 Which of the following would favour the uptake of AI in museums?

	Not at all				Very much
	0	1	2	3	4	5
Open-access (OA) software ()						
Data sharing partnerships between museums ()						
Plug-and-play products for museums ()						
Adoption of data standards ()						
Governing bodies and regulators especially focused on heritage institutions ()						
Other: ()						

Q28 Are museums suitable institutions to educate the general public about AI technologies and its ethical implications?

- Yes (1)
- No (2)
- I do not know (3)

Q29 What best describes your current job role?

- Leadership/strategic (1)
- Curatorial (2)
- Collection Management and Registry (3)
- Digital/technical (4)
- Conservation/Restoration (7)
- Research (5)
- Other (please describe): (6)

Q30 How long for have you been working in the museum sector?

- 1-4 years (1)
- 5-10 years (2)
- 11-15 years (3)
- > 15 years (4)

Q31 In which country is the institution you are working for based? Multi-site institutions please indicate the country your main building is based at.

▼ Afghanistan (1) ... Zimbabwe (1357)

Q32 In what area is the museum you are working for located?

- Capital city (1)
- Urban (2)
- Multi-site, but generally urban (3)
- Multi-site, but generally rural (4)
- Rural (5)

Q33 How many people are working for this institution?

- 1-5 (1)
- 6-15 (2)
- 16-30 (3)
- 31-50 (4)
- 51-100 (5)
- 101-200 (6)
- > 200 (7)
- I do not know (8)

Q34 What is the approximate annual budget of the institution you are working for? (If not included below, please convert your local currency to British Pound Sterling)

- < £25 000 (< €29 000; < $35 000) (1)
- £26 000 - £100 000 (€35 000 - €117 000; $42 000 - $140 000) (2)
- £100 000 - £1 000 000 (€117 000 - €1 171 000; $140 000 - $1 400 000) (3)
- £1 000 000 - £5 000 000 (€1 171 000 - €5 852 000; $1 400 000 - $7 000 000) (4)
- > £5 000 000 (> €5 852 000; > $7 000 000) (5)
- I do not know (6)

Q35 How many visitors does your institution attract annually? (Please indicate pre-COVID numbers in case your institution was subject to a lockdown)

- < 20 000 (1)
- < 50 000 (2)
- < 100 000 (3)
- < 300 000 (4)
- < 500 000 (5)
- < 1 000 000 (6)
- > 1 000 000 (7)
- I do not know (8)

Q36 What type best describes the museum you are working at?

- Science (1)
- Art (2)
- History (3)
- Natural History (4)
- Ethnography (5)
- Universal (holding a significant amount of various collections; multiple exhibitions with various themes) (6)
- Other (please define): (7)

APPENDIX B

Focus groups:

Artwork (a):

George Fiddes Watt
William Gordon, LLD, Town Clerk of Aberdeen (1875-1924)
1916
Oil on canvas
Aberdeen Art Gallery and Museum

Artwork (b):

Unknown artist
A Diseased Body Part
n.d.
Watercolour and pencil on paper
Wellcome Collection

Artwork (c):

William Stewart MacGeorge
Wood Cutters
n.d.
Oil on canvas
East Lothian Council

APPENDIX C

Table C.1. **Interaction metrics UX/UI online study:**

Interaction metrics
part_one_time
part_two_time
overall_time
part_one_artwork_time_m
part_one_artwork_time_std
part_two_artwork_time_m
part_two_artwork_time_std
num_artworks_part_one
num_artworks_part_two
total_events
num_revisits_part_one
num_revisits_part_two
num_visited_before_first_choice_part_one
num_visited_before_first_choice_part_two
time_before_first_choice_part_one
time_before_first_choice_part_two
show_more_part_one
show_more_part_two
artwork_selected_part_one
artwork_selected_part_two
artwork_deselected_part_one
artwork_deselected_part_two
show_more_part_one
show_more_part_two

***Table C.2.* User study artwork catalogue** [ID, title, linked terms, linked topics]

A_1

ACC_ACC_ACA_59-001
Telling the Tale
["Ashtray", "Beer", "Bench", "Cap", "Coat", "Conversation", "Drink", "Drinking", "Flat Cap", "Flat Caps", "Interior", "Lifeboat", "Man", "Men", "Model", "Model Boat", "Newspaper", "Pipe", "Pub", "Red Table", "Sitting", "Smoking", "Table", "Three Men", "Window", "Wooden Bench"]
["Daily life", "Eating and drinking (sport and leisure)", "Entertainment venues", "Everyday costume", "Groups", "Interiors", "Men"]

A_2

BCN_ELYM_1999_091-001
A View of Ely, Cambridgeshire, from St Mary's Street
["Bonnet", "Building", "Cathedral", "Chimney", "Church", "Cloud", "Firewood", "Horse", "House", "Man", "Pub", "Rider", "Road", "Sign", "Sky", "Spire", "Street", "Tower", "Window", "Woman", "Wood"]
["Places of worship", "Daily life", "Everyday costume", "Horses", "Houses", "Jewellery", "hats and accessories", "Men", "Public buildings", "Religious buildings", "Streets and squares", "Townscapes", "Women"]

A_3

TATE_TATE_T00733_10-001
Lady Macbeth Seizing the Daggers
["Royal Academician"]
["Fear and horror"]

A_4

CW_FAG_FAMAG_2004_17_27-001
Déjà vu
["Abstract", "Colour", "Curve", "Curves", "Inset", "Map", "Orange", "Pattern", "Purple", "Rectangle", "Red", "Shape", "Shapes", "Spiral", "Square", "Squiggle", "Swirl", "Wave", "Window", "Yellow"]
["Art", "Happiness and joy", "Senses", "Time"]

A_5

ASH_ASHM_WA1955_66-001
Landscape near Muiderberg
["Cloud", "Field", "Horse", "House", "Stream", "Tower", "Tree", "Landscape"]
["Countryside", "Plants and flowers", "Rivers and lakes", "Rural buildings", "Trees and shrubs"]

A_6

TATE_TATE_N05390_10-001
The Poet Reclining (Le Poète allongé)
["Barn", "Fir Tree", "Hat", "Horse", "Man", "Paddock", "Pig", "Poet", "Recline", "Shed", "Sheep", "Sky", "Stable", "Tree", "Royal Academician"]
["Animals", "farm", "Countryside", "Everyday costume", "Horses", "Literature", "Men", "Reading and writing", "Trees and shrubs"]

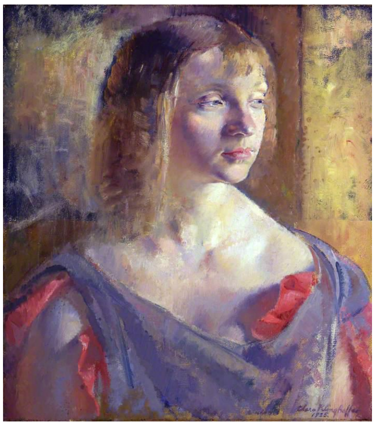
A_7

SFK_SED_MA_1992_9_395-001
Barbara Stone
["Apple", "Knife", "Woman", "Hand"]
["Eating and drinking (sport and leisure)", "Women"]

A_8

LLR_NTLMS_1965_094_1-001
Flowers in a Pottery Vase
["Carnation", "Colour", "Daisy", "Flower", "Light", "Pottery", "Shadow", "Surface", "Table", "Vase", "Wall"]
["Furniture and interiors", "Hobbies and pastimes", "Interiors", "Plants and flowers"]

A_9



ABD_AAG_ag002265-001
A Girl's Head
["Drapery", "Dress", "Face", "Girl", "Hair", "Head", "Light", "Neck", "Shoulder", "Thoughtful", "Woman"]
["Women", "Children", "Children's costume", "Drapery and classical costume", "Evening and formal costume", "Everyday costume", "Hairstyles", "cosmetics and body art", "Nudes and models"]

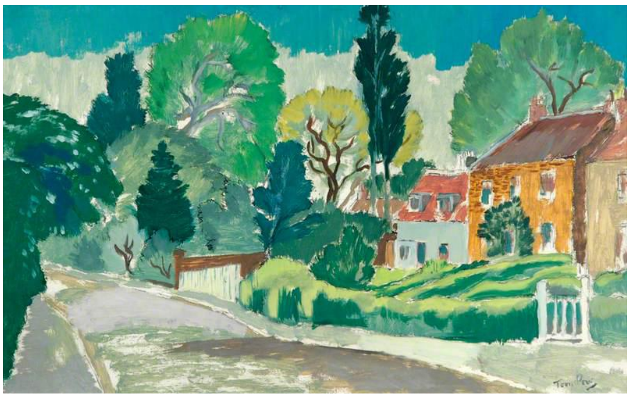
A_10

NG_NG_NG6394-001	
A Man Reading (Saint Ivo?)	
["Castle", "Church", "Document", "Ear", "Face", "Fur", "Lake", "Landscape", "Letter", "Man", "People", "Reading", "Robe", "Saint", "Shutter", "Tower", "Tree", "View", "Window"]	
["Trees and shrubs", "Clerics", "Countryside", "Drapery and classical costume", "Everyday costume", "Fortifications", "Hills and mountains", "Interiors", "Men", "Reading and writing", "Rivers and lakes", "Saints and martyrs"]	

A_11

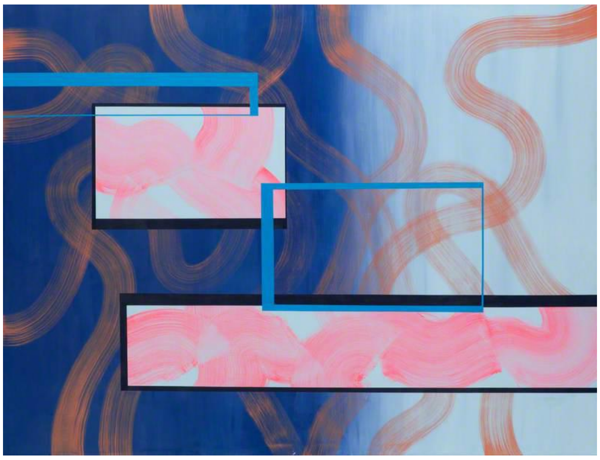
NGS_NGS_AR00230-001
Alexandre Iolas
["LGBT"]
["LGBTQ+ subjects"]

A_12



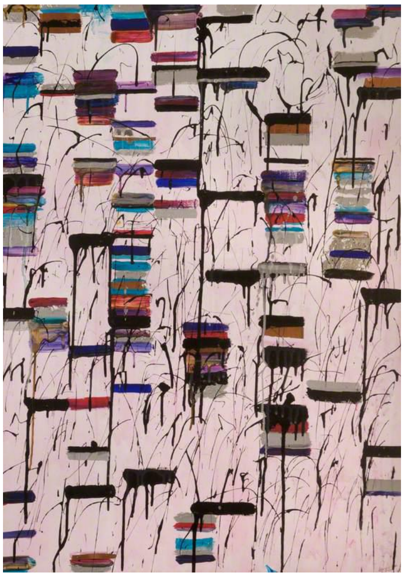
NTV_PH_66-001
Row of Terraced Cottages
["Cottage", "Garden", "Gate", "Hedge", "House", "Road", "Sky", "Tree", "Green", "Viridian"]
["Gardens and green spaces", "Houses", "Rural buildings", "Trees and shrubs"]

A_13



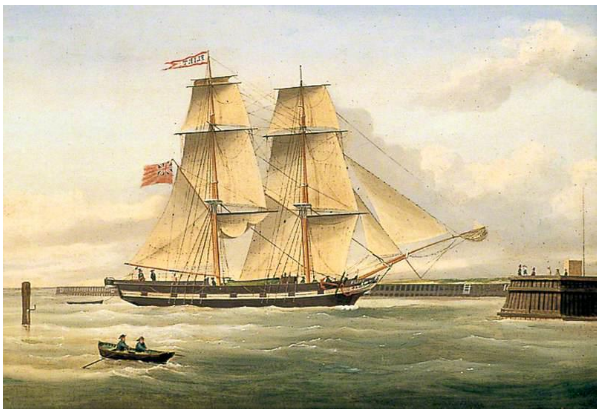
ABD_RGU_10186-001
Pink on Blue
["Abstract", "Abstraction", "Blue", "Contrast", "Curve", "Meander", "Mirror", "Oblong", "Orange", "Pink", "Rectangle", "Rectangular", "Shading", "Square", "Squiggle", "Swirl"]
["Art", "Senses"]

A_14



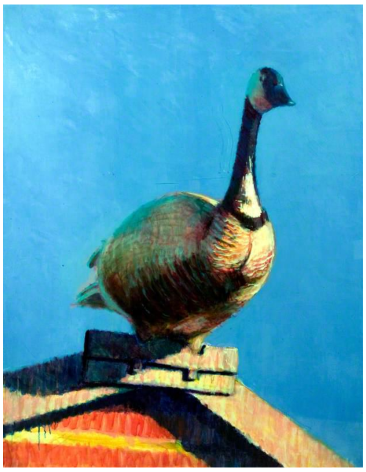
CHE_CEC_PCF119-001
DNA
["Dna", "Stripe"]
["Medicine", "Sciences"]

A_15



WS_LMT_002_022A-001
The Barque 'Flirt'
["Boat", "Cloud", "Flag", "Harbour", "Mast", "Pennant", "Quay", "Rowing Boat", "Sail", "Sailing Ship", "Sea", "Ship", "Wave", "Sailing"]
["Ports and waterways", "Seas and coasts", "Ships and boats"]

A_16



LLR_AWC_0829-001
Garden Folly (Goose)
["Canada Goose", "Folly", "Garden Folly", "Goose", "Ornament", "Roof", "Shed", "Sky"]
["Animals", "wild", "Birds", "Craft (art)", "Gardens and green spaces"]

A_17

VA_PC_2006BA0177-001
Seaford, Sussex
["Beach", "Child", "Sand", "Sea", "Coast", "Shore", "Horizon", "Royal Academician"]
["Seas and coasts"]

A_18

NWM_ALU_PCF1-001
Snowdon
["Cloud", "Crag", "Landscape", "Mountain", "Peak", "Rock", "Rugged", "Sky", "Snow", "Stormy", "Wild", "Royal Academician"]
["Wild places", "Hills and mountains", "Weather"]

A_19

GL_GM_1719-001
The Mail Coach
["Horse", "Rider", "Snow", "Tree"]
["Horses", "Seasons", "Trees and shrubs", "Weather", "Countryside"]

A_20

CAM_CCF_781-001
The Judgement of Zaleucus
["Building", "Crowd", "Judgement", "Monument"]
["Groups", "Law and order", "Virtues and vices"]

A_21

SYO_BHA_90003666-001
A Dog
["Dog"]
["Animals", "domestic", "Dogs"]

A_22

NY_MAG_HARAG_267-001
The Hop-Pickers
["Busy", "Child", "Children", "Community", "Countryside", "Dress", "Field", "Harvest", "Hill", "Hop", "Hop Bine", "Hop Garden", "Hop Picking", "Hop-picker", "Hops", "Joy", "Labour", "Men", "Oast House", "People", "Picker", "Summer", "Tree", "Vibrancy", "Woman", "Women", "Work", "Worker"]
["Men", "Plants and flowers", "Seasons", "Trees and shrubs", "Women", "Workwear", "Children", "Countryside", "Farming and fishing", "Groups"]

A_23

NTS_BROD_2009_597-001
William Warr Defeating William Wood at Navestock in Essex, 31 December, 1788
["Boxer", "Boxing", "Boxing Match", "Breeches", "Crowd", "Fight", "Man", "Referee", "Shoe", "Spectator", "Tree", "Royal Academician"]
["Countryside", "Everyday costume", "Groups", "Men", "Sporting costume", "Sports", "Trees and shrubs"]

A_24

STF_SAMS_G98_003_0002-001
Youthful Despair No. 2
["Boy", "Child", "Despair", "Girl", "Jeans", "People", "Young People", "Youth", "Youthful"]
["Men", "Sadness and grief", "Women", "Children", "Everyday costume", "Fear and horror", "Groups", "Hairstyles", "cosmetics and body art"]

A_25

ACC_ACC_ACC1_1947-001
Reconstruction
["Blue", "Chisel", "Concentration", "Cooperation", "Doctor", "Doctors", "Face Mask", "Gown", "Hammer", "Hat", "Hospital", "Instrument", "Mallet", "Man", "Mask", "Medical", "Men", "Nurse", "Operating Theatre", "Operation", "Osteotome", "People", "Reconstruction", "Retractor", "Sculptor", "Sculpture", "Sketch", "Sterile", "Surgeon", "Surgeons", "Surgery", "Tool", "White", "Woman"]
["Senses", "Sickness and health", "Women", "Workwear", "Art", "Crafts", "Groups", "Healthcare", "Life and death", "Medicine", "Men"]

A_26

HSW_DMAG_1976_94-001
The Border of the Nile Valley
["Bush", "Camel", "Cloud", "Desert", "Dune", "Hill", "Horizon", "Landscape", "Man", "People", "Plant", "Rider", "Sand", "Scrub", "Shrub", "Sky", "Valley"]
["Animals", "domestic", "Countryside", "Hills and mountains", "Men", "Plants and flowers", "Times of day", "Trees and shrubs", "Wild places"]

A_27

ACC_ACC_AC_5490-001
Milk Bar
["Abstract", "Black", "Blue", "Cow", "Green", "Line", "Milk Bar", "Shape", "Yellow"]
["Animals", "farm", "Eating and drinking (sport and leisure)", "Interiors"]

A_28

NOT_NSDC_36_70-001
Paintworks
["Barrel", "Bench", "Bottle", "Bucket", "Cloth", "Container", "Door", "Floor", "Funnel", "Hat", "Interior", "Jar", "Ladder", "Latch", "Man", "Men", "Overall", "Pail", "Paint", "Paintworks", "Pot", "Room", "Shelf", "Sink", "Table", "Tap", "Towel", "Window", "Work", "Workbench", "Worker", "Workshop"]
["Industry", "Interiors", "Manufacturing industry", "Men", "Workplaces", "Workwear", "Crafts"]

A_29

ACC_ACC_AC_1370-001
China
["Abstract", "Geometric", "Geometric Shapes", "Geometry", "Grey", "Pale Colours", "Shape", "Symmetry", "Triangle", "Royal Academician"]
["Senses"]

A_30

NG_NG_NG1079-001
Adoration of the Kings
["Adoration", "Baby", "Boot", "Casket", "Castle", "Epiphany", "Gift", "Hat", "King", "Kneel", "Magi", "Mother", "Nativity", "Robe", "Ruin", "Town", "Child", "View", "Woman", "Building"]
["Men", "Ruins", "Saints and martyrs", "Stories and people", "Townscapes", "Women", "Worshippers and congregations", "Children", "Drapery and classical costume", "Fortifications", "Jesus Christ", "Jewellery", "hats and accessories"]

Post-study questionnaire

1. <u>Post-study questionnaire</u>

   On completion of parts I and II participants will be asked to complete a post-study questionnaire, which uses adapted questions from the Knijnenburg et al. (2012) framework.

   Answers are entered via 5-point Likert scales ranging from 1 (completely disagree) to 5 (completely agree).

1.1. Please think of **PART I** (the first half of the study) when answering the following questions:

Perceived recommendation quality

- I liked the artworks shown by the system.
- The artworks fitted my preference.
- The artworks were well-chosen.
- The artworks were relevant.
- The system showed me too many bad artworks.
- I didn’t like any of the artworks shown.

Perceived system effectiveness and fun

- I have fun when I am using the system.
- I would recommend the system to others.
- Using the system is a pleasant experience.
- I can find interesting artworks with the system.
- The system showed me artworks I would usually not find.
- The system is useless.
- The system makes me more aware of my choice option.
- I make more informed choices with the system.
- I can find better items without the help of the system.
- The system showed useful items.

Choice satisfaction

- I like the artworks I have seen.
- I was excited about the artworks shown.
- I enjoyed seeing the artworks shown to me.
- The artworks shown to me were diverse.
- The artworks shown to me were novel.
- The system offered serendipitous items.
- The artworks I have seen were a waste of time.
- I would recommend some of the shown artworks to family and friends.

Test awareness

- I am aware that the system showed me recommendations.
- I am aware that items in this part were especially chosen to suit my choice of artworks.

1.2. Please think of **PART II** (the second half of the study) when answering the following questions:

Perceived recommendation quality

- I liked the artworks shown by the system.
- The artworks fitted my preference.
- The artworks were well-chosen.
- The artworks were relevant.
- The system showed me too many bad artworks.
- I didn't like any of the artworks shown.

Perceived system effectiveness and fun

- I have fun when I am using the system.
- I would recommend the system to others.
- Using the system is a pleasant experience.
- I can find interesting artworks with the system.
- The system showed me artworks I would usually not find.
- The system is useless.
- The system makes me more aware of my choice option.
- I make more informed choices with the system.
- I can find better items without the help of the system.
- The system showed useful items.

Choice satisfaction

- I like the artworks I have seen.
- I was excited about the artworks shown.
- I enjoyed seeing the artworks shown to me.
- The artworks shown to me were diverse.
- The artworks shown to me were novel.
- The system offered serendipitous items.
- The artworks I have seen were a waste of time.
- The chosen artworks fitted my preference.
- I would recommend some of the shown artworks to family and friends.

Test awareness

- I am aware that the system showed me recommendations.
- I am aware that items in this part were especially chosen to suit my choice of artworks.

1.3. **General** questions asked after the part specific ones:

Intention to provide feedback

- I didn't mind having to choose artworks.

General trust in technology

- Technology never works.
- I trust the system I have just used.
- Technology should always be explainable.
- I am not fussed about how things work in the background as long as the technology works.
- I am generally questioning what happens to my personal data

Museum online collection relevant question

- I prefer a classic keyword search compared to this system.
- The system is not suitable to display artworks.
- Museum online collections are generally boring.

I do not need museum online collections.

Table C.3. **p-values: comparison interaction data to post-study questionnaire**

- Spearman's rank for time compared to questionnaire: Those that spent more time, did they record higher scores in the survey?

	X	Y	method	alternative	n	r	CI95%	p-unc	p-corr	p-adjust	power
0	part_one_time	satisfaction	spearman	two-sided	151	0.113793	[-0.05, 0.27]	0.164160	1.0	bonf	0.286133
1	part_one_time	effectiveness	spearman	two-sided	151	0.070075	[-0.09, 0.23]	0.392556	1.0	bonf	0.137188
2	part_one_time	quality	spearman	two-sided	151	-0.025960	[-0.18, 0.13]	0.751694	1.0	bonf	0.061465
3	part_one_time	awareness	spearman	two-sided	151	-0.082004	[-0.24, 0.08]	0.316827	1.0	bonf	0.170632
4	part_two_time	satisfaction	spearman	two-sided	151	0.020124	[-0.14, 0.18]	0.806253	1.0	bonf	0.056828
5	part_two_time	effectiveness	spearman	two-sided	151	0.067232	[-0.09, 0.22]	0.412089	1.0	bonf	0.130048
6	part_two_time	quality	spearman	two-sided	151	-0.121937	[-0.28, 0.04]	0.135829	1.0	bonf	0.321211
7	part_two_time	awareness	spearman	two-sided	151	0.046771	[-0.11, 0.2]	0.568496	1.0	bonf	0.087999
8	overall_time	satisfaction	spearman	two-sided	151	0.103995	[-0.06, 0.26]	0.203823	1.0	bonf	0.246677
9	overall_time	effectiveness	spearman	two-sided	151	0.074799	[-0.09, 0.23]	0.361356	1.0	bonf	0.149758
10	overall_time	quality	spearman	two-sided	151	-0.059865	[-0.22, 0.1]	0.465285	1.0	bonf	0.113029
11	overall_time	awareness	spearman	two-sided	151	-0.017146	[-0.18, 0.14]	0.834483	1.0	bonf	0.054921

- Spearman's rank for time spent on artworks to questionnaire: Those that spent more time on artworks, did they record higher scores in the survey?

	X	Y	method	alternative	n	r	CI95%	p-unc	p-corr	p-adjust	power
0	part_one_artwork_time_m	satisfaction	spearman	two-sided	147	0.087533	[-0.08, 0.25]	0.291770	1.0	bonf	0.184217
1	part_one_artwork_time_m	effectiveness	spearman	two-sided	147	0.034205	[-0.13, 0.2]	0.680856	1.0	bonf	0.069531
2	part_one_artwork_time_m	quality	spearman	two-sided	147	-0.007118	[-0.17, 0.15]	0.931809	1.0	bonf	0.050732
3	part_one_artwork_time_m	awareness	spearman	two-sided	147	-0.080196	[-0.24, 0.08]	0.334253	1.0	bonf	0.162015
4	part_two_artwork_time_m	satisfaction	spearman	two-sided	148	0.094318	[-0.07, 0.25]	0.254186	1.0	bonf	0.207662
5	part_two_artwork_time_m	effectiveness	spearman	two-sided	148	0.072646	[-0.09, 0.23]	0.380248	1.0	bonf	0.141970
6	part_two_artwork_time_m	quality	spearman	two-sided	148	-0.001456	[-0.16, 0.16]	0.985991	1.0	bonf	0.049927
7	part_two_artwork_time_m	awareness	spearman	two-sided	148	0.055268	[-0.11, 0.21]	0.504660	1.0	bonf	0.102382

- Spearman's rank for number of artworks looked at to questionnaire: Those who looked at more artworks, did they record higher scores in the survey?

	X	Y	method	alternative	n	r	CI95%	p-unc	p-corr	p-adjust	power
0	num_artworks_part_one	satisfaction	spearman	two-sided	151	0.027937	[-0.13, 0.19]	0.733469	1.000000	bonf	0.063309
1	num_artworks_part_one	effectiveness	spearman	two-sided	151	0.004479	[-0.16, 0.16]	0.956474	1.000000	bonf	0.050236
2	num_artworks_part_one	quality	spearman	two-sided	151	-0.063103	[-0.22, 0.1]	0.441453	1.000000	bonf	0.120246
3	num_artworks_part_one	awareness	spearman	two-sided	151	-0.086793	[-0.24, 0.07]	0.289294	1.000000	bonf	0.185631
4	num_artworks_part_two	satisfaction	spearman	two-sided	151	0.018857	[-0.14, 0.18]	0.818236	1.000000	bonf	0.055979
5	num_artworks_part_two	effectiveness	spearman	two-sided	151	-0.027931	[-0.19, 0.13]	0.733524	1.000000	bonf	0.063303
6	num_artworks_part_two	quality	spearman	two-sided	151	-0.145059	[-0.3, 0.02]	0.075549	0.604393	bonf	0.429558
7	num_artworks_part_two	awareness	spearman	two-sided	151	-0.061620	[-0.22, 0.1]	0.452279	1.000000	bonf	0.116890

Table C.4. **Linear regression: Are the interaction metrics predictive of the survey results?**

- In terms of satisfaction:

	names	coef	se	T	pval	r2	adj_r2	CI[2.5%]	CI[97.5%]
0	Intercept	2.397	0.265	9.061	0.000	0.158	0.005	1.873	2.920
1	part_one_time	-87.305	103.298	-0.845	0.400	0.158	0.005	-291.713	117.104
2	part_two_time	-87.355	103.295	-0.846	0.399	0.158	0.005	-291.758	117.047
3	overall_time	87.320	103.297	0.845	0.400	0.158	0.005	-117.088	291.727
4	part_one_artwork_time_m	-0.000	0.009	-0.023	0.982	0.158	0.005	-0.018	0.018
5	part_one_artwork_time_std	0.001	0.003	0.369	0.713	0.158	0.005	-0.005	0.007
6	part_two_artwork_time_m	0.002	0.028	0.076	0.939	0.158	0.005	-0.054	0.058
7	part_two_artwork_time_std	0.003	0.006	0.482	0.631	0.158	0.005	-0.009	0.014
8	num_artworks_part_one	0.001	0.008	0.070	0.944	0.158	0.005	-0.015	0.016
9	num_artworks_part_two	-0.010	0.009	-1.123	0.264	0.158	0.005	-0.027	0.008
10	total_events	0.025	0.025	1.010	0.314	0.158	0.005	-0.024	0.074
11	num_revisits_part_one	-0.010	0.024	-0.436	0.663	0.158	0.005	-0.057	0.036
12	num_revisits_part_two	0.065	0.039	1.673	0.097	0.158	0.005	-0.012	0.141
13	num_visited_before_first_choice_part_one	-0.022	0.021	-1.026	0.307	0.158	0.005	-0.064	0.020
14	time_before_first_choice_part_one	-0.012	0.070	-0.171	0.864	0.158	0.005	-0.151	0.127
15	num_visited_before_first_choice_part_two	-0.069	0.082	-0.843	0.401	0.158	0.005	-0.231	0.093
16	time_before_first_selection_part_two	0.332	0.338	0.982	0.328	0.158	0.005	-0.337	1.001
17	show_more_part_one	-0.014	0.026	-0.550	0.583	0.158	0.005	-0.065	0.037
18	artwork_selected_part_one	-0.025	0.030	-0.843	0.401	0.158	0.005	-0.084	0.034
19	artwork_deselected_part_one	-0.015	0.051	-0.293	0.770	0.158	0.005	-0.116	0.086
20	show_more_part_two	-0.043	0.029	-1.459	0.147	0.158	0.005	-0.100	0.015
21	artwork_selected_part_two	-0.007	0.028	-0.270	0.788	0.158	0.005	-0.062	0.047
22	artwork_deselected_part_two	0.130	0.153	0.846	0.399	0.158	0.005	-0.173	0.433
23	model_type_concatenated	0.893	0.106	8.404	0.000	0.158	0.005	0.682	1.103
24	model_type_image	0.830	0.105	7.905	0.000	0.158	0.005	0.622	1.038
25	model_type_meta	0.675	0.114	5.909	0.000	0.158	0.005	0.449	0.900

- In terms of effectiveness:

	names	coef	se	T	pval	r2	adj_r2	CI[2.5%]	CI[97.5%]
0	Intercept	2.575	0.215	12.005	0.000	0.101	-0.061	2.151	3.000
1	part_one_time	-43.688	83.764	-0.522	0.603	0.101	-0.061	-209.443	122.067
2	part_two_time	-43.727	83.762	-0.522	0.603	0.101	-0.061	-209.476	122.023
3	overall_time	43.704	83.764	0.522	0.603	0.101	-0.061	-122.050	209.457
4	part_one_artwork_time_m	0.007	0.007	1.004	0.317	0.101	-0.061	-0.007	0.022
5	part_one_artwork_time_std	-0.002	0.003	-0.845	0.400	0.101	-0.061	-0.007	0.003
6	part_two_artwork_time_m	-0.006	0.023	-0.254	0.800	0.101	-0.061	-0.051	0.040
7	part_two_artwork_time_std	0.003	0.005	0.706	0.481	0.101	-0.061	-0.006	0.013
8	num_artworks_part_one	-0.002	0.006	-0.248	0.805	0.101	-0.061	-0.014	0.011
9	num_artworks_part_two	-0.006	0.007	-0.845	0.400	0.101	-0.061	-0.020	0.008
10	total_events	0.036	0.020	1.790	0.076	0.101	-0.061	-0.004	0.076
11	num_revisits_part_one	-0.001	0.019	-0.055	0.956	0.101	-0.061	-0.039	0.037
12	num_revisits_part_two	0.017	0.031	0.543	0.588	0.101	-0.061	-0.045	0.079
13	num_visited_before_first_choice_part_one	-0.016	0.017	-0.946	0.346	0.101	-0.061	-0.050	0.018
14	time_before_first_choice_part_one	0.002	0.057	0.027	0.979	0.101	-0.061	-0.111	0.115
15	num_visited_before_first_choice_part_two	0.023	0.066	0.344	0.731	0.101	-0.061	-0.108	0.154
16	time_before_first_selection_part_two	-0.067	0.274	-0.243	0.809	0.101	-0.061	-0.609	0.476
17	show_more_part_one	-0.032	0.021	-1.536	0.127	0.101	-0.061	-0.074	0.009
18	artwork_selected_part_one	-0.036	0.024	-1.503	0.135	0.101	-0.061	-0.084	0.012
19	artwork_deselected_part_one	-0.048	0.042	-1.162	0.247	0.101	-0.061	-0.131	0.034
20	show_more_part_two	-0.043	0.024	-1.819	0.071	0.101	-0.061	-0.090	0.004
21	artwork_selected_part_two	-0.030	0.022	-1.329	0.186	0.101	-0.061	-0.074	0.015
22	artwork_deselected_part_two	0.226	0.124	1.819	0.071	0.101	-0.061	-0.020	0.472
23	model_type_concatenated	0.937	0.086	10.879	0.000	0.101	-0.061	0.766	1.107
24	model_type_image	0.858	0.085	10.077	0.000	0.101	-0.061	0.689	1.026
25	model_type_meta	0.781	0.093	8.434	0.000	0.101	-0.061	0.598	0.964

- In terms of quality:

	names	coef	se	T	pval	r2	adj_r2	CI[2.5%]	CI[97.5%]
0	Intercept	2.215	0.200	11.076	0.000	0.148	-0.006	1.819	2.611
1	part_one_time	67.503	78.098	0.864	0.389	0.148	-0.006	-87.039	222.045
2	part_two_time	67.435	78.096	0.863	0.389	0.148	-0.006	-87.102	221.972
3	overall_time	-67.477	78.097	-0.864	0.389	0.148	-0.006	-222.018	87.064
4	part_one_artwork_time_m	0.002	0.007	0.342	0.733	0.148	-0.006	-0.011	0.016
5	part_one_artwork_time_std	-0.002	0.002	-0.737	0.462	0.148	-0.006	-0.007	0.003
6	part_two_artwork_time_m	-0.003	0.021	-0.137	0.892	0.148	-0.006	-0.045	0.040
7	part_two_artwork_time_std	0.004	0.004	0.925	0.357	0.148	-0.006	-0.005	0.013
8	num_artworks_part_one	-0.011	0.006	-1.877	0.063	0.148	-0.006	-0.023	0.001
9	num_artworks_part_two	-0.004	0.007	-0.661	0.510	0.148	-0.006	-0.018	0.009
10	total_events	0.017	0.019	0.903	0.368	0.148	-0.006	-0.020	0.054
11	num_revisits_part_one	0.020	0.018	1.123	0.263	0.148	-0.006	-0.015	0.055
12	num_revisits_part_two	0.032	0.029	1.090	0.278	0.148	-0.006	-0.026	0.090
13	num_visited_before_first_choice_part_one	0.016	0.016	1.011	0.314	0.148	-0.006	-0.016	0.048
14	time_before_first_choice_part_one	-0.065	0.053	-1.224	0.223	0.148	-0.006	-0.171	0.040
15	num_visited_before_first_choice_part_two	-0.012	0.062	-0.187	0.852	0.148	-0.006	-0.134	0.111
16	time_before_first_selection_part_two	0.282	0.256	1.104	0.272	0.148	-0.006	-0.224	0.788
17	show_more_part_one	-0.019	0.020	-0.953	0.342	0.148	-0.006	-0.057	0.020
18	artwork_selected_part_one	-0.015	0.023	-0.646	0.520	0.148	-0.006	-0.059	0.030
19	artwork_deselected_part_one	-0.016	0.039	-0.412	0.681	0.148	-0.006	-0.093	0.061
20	show_more_part_two	-0.022	0.022	-0.978	0.330	0.148	-0.006	-0.065	0.022
21	artwork_selected_part_two	0.003	0.021	0.155	0.877	0.148	-0.006	-0.038	0.045
22	artwork_deselected_part_two	0.084	0.116	0.730	0.467	0.148	-0.006	-0.145	0.314
23	model_type_concatenated	0.741	0.080	9.232	0.000	0.148	-0.006	0.582	0.900
24	model_type_image	0.709	0.079	8.928	0.000	0.148	-0.006	0.552	0.866
25	model_type_meta	0.765	0.086	8.867	0.000	0.148	-0.006	0.595	0.936

- In terms of awareness:

	names	coef	se	T	pval	r2	adj_r2	CI[2.5%]	CI[97.5%]
0	Intercept	2.476	0.456	5.425	0.000	0.093	-0.071	1.573	3.379
1	part_one_time	81.534	178.212	0.458	0.648	0.093	-0.071	-271.114	434.183
2	part_two_time	81.600	178.206	0.458	0.648	0.093	-0.071	-271.038	434.237
3	overall_time	-81.551	178.210	-0.458	0.648	0.093	-0.071	-434.197	271.095
4	part_one_artwork_time_m	0.000	0.016	0.015	0.988	0.093	-0.071	-0.031	0.032
5	part_one_artwork_time_std	-0.001	0.005	-0.176	0.860	0.093	-0.071	-0.012	0.010
6	part_two_artwork_time_m	-0.042	0.049	-0.860	0.392	0.093	-0.071	-0.139	0.055
7	part_two_artwork_time_std	0.005	0.010	0.526	0.600	0.093	-0.071	-0.015	0.025
8	num_artworks_part_one	-0.003	0.014	-0.244	0.808	0.093	-0.071	-0.030	0.024
9	num_artworks_part_two	-0.022	0.015	-1.412	0.160	0.093	-0.071	-0.052	0.009
10	total_events	0.016	0.043	0.383	0.702	0.093	-0.071	-0.068	0.101
11	num_revisits_part_one	0.042	0.041	1.041	0.300	0.093	-0.071	-0.038	0.123
12	num_revisits_part_two	0.091	0.067	1.367	0.174	0.093	-0.071	-0.041	0.223
13	num_visited_before_first_choice_part_one	-0.017	0.037	-0.472	0.638	0.093	-0.071	-0.090	0.055
14	time_before_first_choice_part_one	0.012	0.121	0.099	0.922	0.093	-0.071	-0.228	0.252
15	num_visited_before_first_choice_part_two	0.000	0.141	0.000	1.000	0.093	-0.071	-0.279	0.279
16	time_before_first_selection_part_two	0.310	0.583	0.531	0.596	0.093	-0.071	-0.844	1.464
17	show_more_part_one	-0.016	0.045	-0.352	0.725	0.093	-0.071	-0.104	0.073
18	artwork_selected_part_one	0.014	0.052	0.274	0.784	0.093	-0.071	-0.088	0.116
19	artwork_deselected_part_one	-0.073	0.088	-0.824	0.412	0.093	-0.071	-0.248	0.102
20	show_more_part_two	-0.004	0.050	-0.080	0.937	0.093	-0.071	-0.104	0.096
21	artwork_selected_part_two	-0.032	0.048	-0.668	0.505	0.093	-0.071	-0.126	0.063
22	artwork_deselected_part_two	0.127	0.264	0.480	0.632	0.093	-0.071	-0.396	0.649
23	model_type_concatenated	0.778	0.183	4.247	0.000	0.093	-0.071	0.416	1.141
24	model_type_image	0.960	0.181	5.298	0.000	0.093	-0.071	0.601	1.318
25	model_type_meta	0.738	0.197	3.748	0.000	0.093	-0.071	0.348	1.128

Trust and privacy

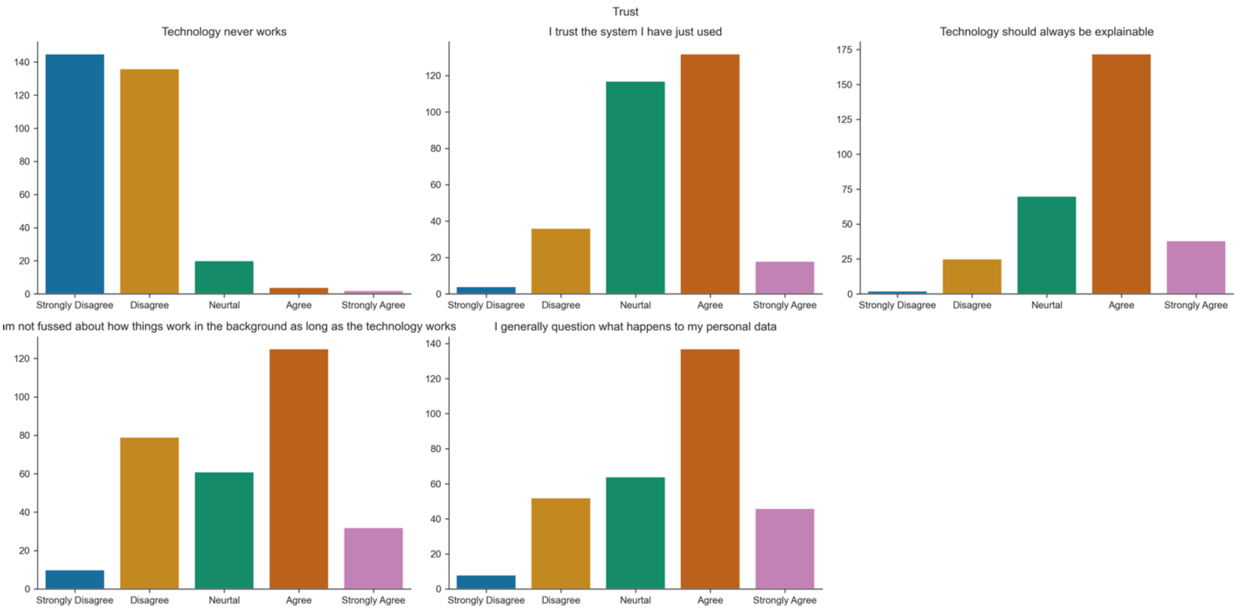


Figure C.1. *Trust and privacy*

Relevance

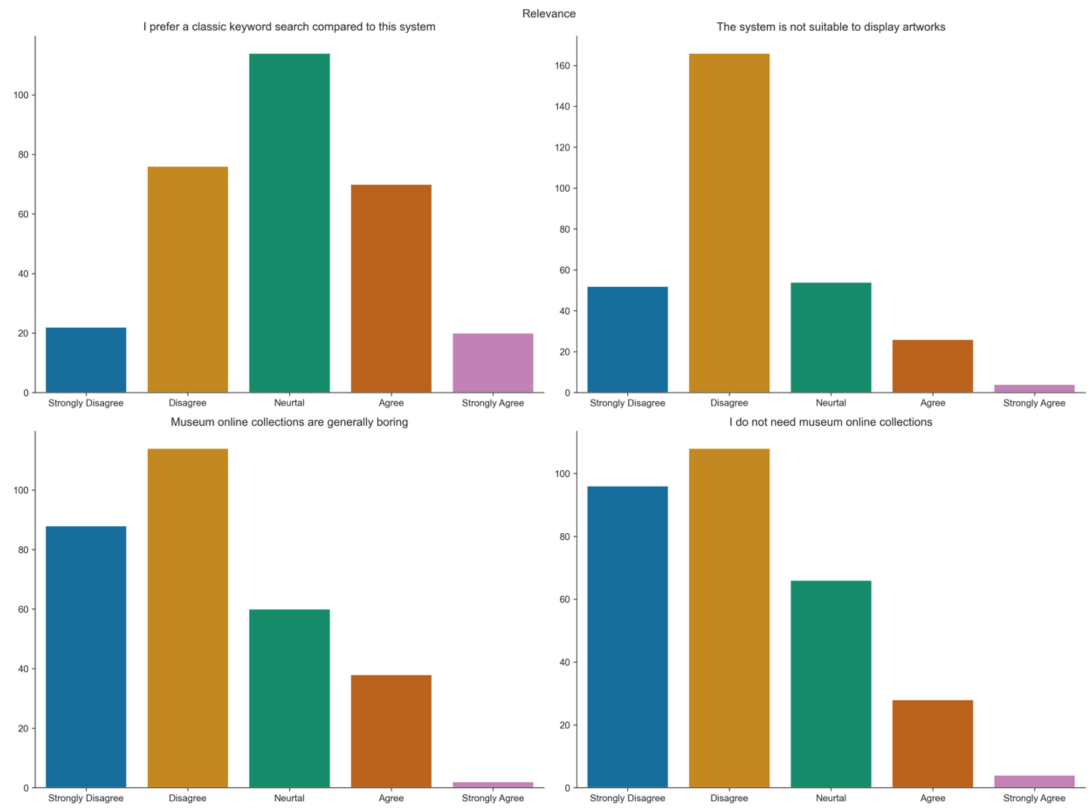


Figure C.2. *Relevance*

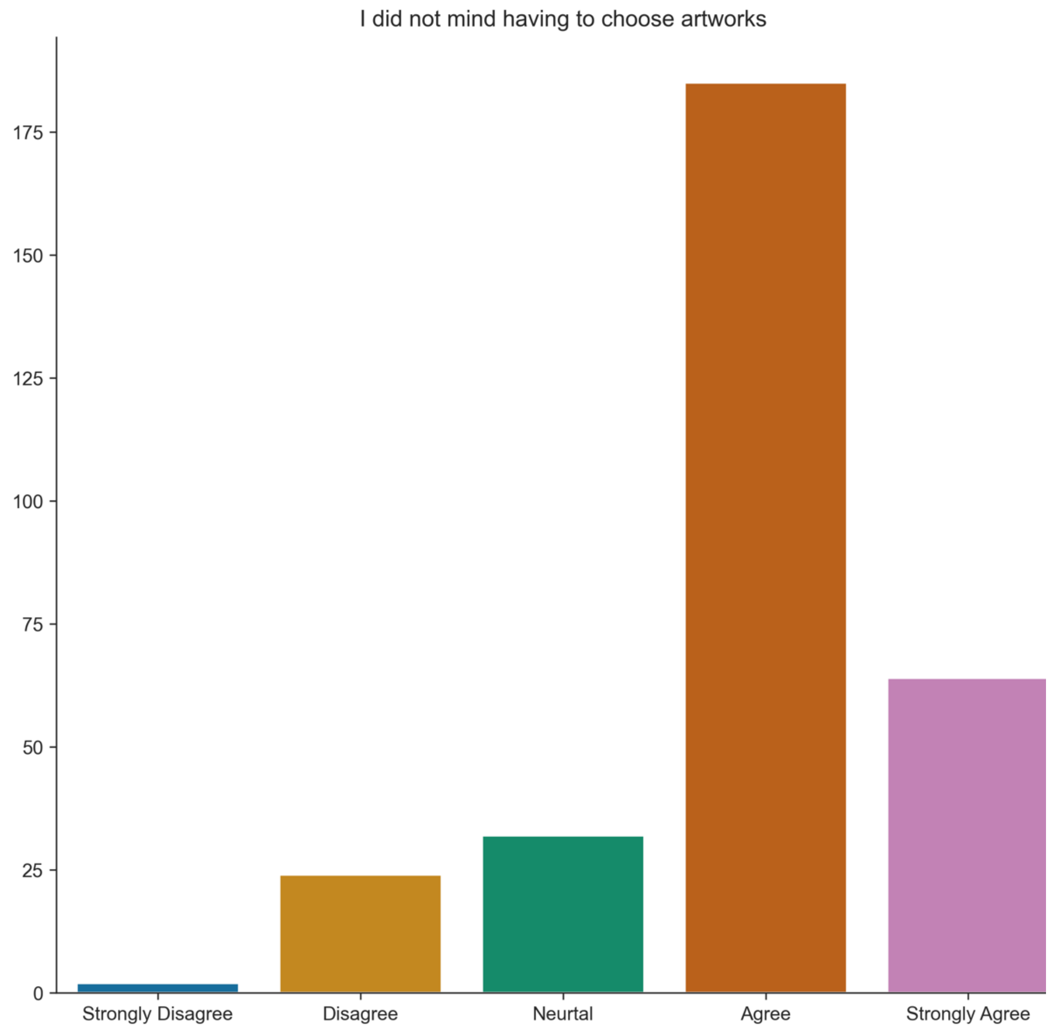


Figure C.3. *User intention to provide feedback via the system*