Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

ZANNA FRIBERG[1]

Preface

For this issue of *Tidskrift för ABM* we invited explorations of the transformative role of digital technologies in relation to cultural heritage, knowledge access and information futures. The issue contains two research articles, a review, an editorial note and two dissertation summaries. The contributions cover a diverse range of topics relating to information futures: personal information privacy, preservation of deplatformed content, challenges and opportunities of AI, digital participation in cultural heritage, and content moderation.

In the editorial note *Digitalization - a hot research topic for students*, Anna-Karin Rabe and Biyanto Rebin highlight common themes in ALM student works published in 2023. The Department of ALM at Uppsala University offers two master programmes, one in ALM (Archives, Libraries and Museums) and one in Digital Humanities. Across these programs Rabe and Rebin have identified a broad interest in digital technologies, through master theses covering themes of digital access, social media, virtualization, strategies and implementation.

This issue's two peer-reviewed research articles by Rikard Friberg von Sydow and Simone Magaard, Daniel Markus and Ann-Sofie Klareld both make inquiries into contemporary issues and their implications when looking

1 Zanna Friberg is a PhD candidate in Information Studies at the Department of ALM at Uppsala University and the editor of the 2023 issue of *Tidskrift för ABM*. Friberg's research interests are many, including archaeological documentation, collection management, knowledge organisation and the history of cultural heritage management.
E-mail: zanna.friberg@abm.uu.se

towards the information future, albeit on two fairly different topics.

In “A Conspiracy of Lost Content” The Case of ”Q” and QAnon, Friberg von Sydow engage with the question of preservation in the digital age. The author engages with the question of how we are to preserve networks of linked information on the internet without ending up with dysfunctional copies full of broken links and empty add spaces. The article particularly focuses on the case of QAnon, which is a movement central to recent political events, most notably the storming of the United States Capitol in Washington DC in 2022. In its notoriety, it is also a movement that has been banned and deplatformed from central social network platforms, making it vulnerable from an archival and preservation perspective. In the article, Friberg von Sydow explores the question of what we need to preserve about an Internet based movement such as QAnon in order to understand it in retrospect. Furthermore, how do we go about preserving knowledge about gaps in these networks that arise when the links that make up the network’s nodes are broken? As deplatforming becomes an increasingly more common societal response to groups spreading misinformation and disinformation, we risk losing our ability to study these groups and movements. Friberg von Sydow looks at the future implications of a contemporary example of information loss while Magaard, Markus and Klareld shines a light on the implications of new types of information being recorded.

Today, there is a number of providers of DNA testing on the market, allowing anyone to purchase an insight into their identity or heritage through a mapping of their genetic lineage and connections. In *Genomic privacy? Direct-to-consumer genetic testing companies’ and the processing of sensitive personal data* Magaard, Markus and Klareld investigates privacy aspects of direct-to-consumer genetic testing (DTCGT), focusing on the providers’ adherence to the General Data Protection Regulation (GDPR). The authors study GDPR adherence, transparency and consent in the privacy policies of the companies *MyHeritage*, *FamilyTreeDNA* and *MyDNA*. Additionally, the authors explore the conceptual relationship between the right to privacy and data protection. Furthermore, they engage with the peculiarities of privacy when it comes to genetic information, which is shared between several individuals. DNA is personal information not only for the person buying

a direct-to-consumer genetic test but also their relatives, past present and future. Implications of DTCGT services were a topic featured in last year's issue of *Tidskrift för ABM* as well, where Widholm and Andersson (2022) explored the implications of DNA testing for the relationship between genealogists and archival institutions. The article by Magaard, Markus and Klareld further illustrates the many aspects of genetic information to pay attention to, now and looking ahead.

Another type of information that is a major component in contemporary ideas of information futures is AI generated content. General use of Artificial Intelligence (AI) has become less and less of a distant future and speculation around the directions of AI technologies going forward are livelier than ever. A recent example the launch of the virtual assistant ChatGPT by the company OpenAI and the debate about use of AI generated writing in higher education and academia. In *Challenges or Opportunities: When Artificial Intelligence is Applied to Digital Humanities*, Haoyue Zhao explores these challenges through a review of ChatGPT4. The author provides the AI tool with a prompt to write a 2000-word text on Digital Humanities. In this essayistic review of AI, the author delves into both the challenges and opportunities of AI and the nature of the Digital Humanities as a field. While an initial look at the text provided by ChatGPT4 may look sound, Zhao notes inconsistencies and issues in both form and content. Zhao concludes that while some challenges in the AI's performance may be more easily solved, like the formatting of references, other deeper issues connected to creativity and ethics remain dependent on human interference.

Tidskrift för ABM is a journal published by the Department of ALM at Uppsala University and these past two years, the department celebrated three colleagues defending their dissertations, Amalia Jeneström, Inge Zwart and Ina-Maria Jansson. In this issue, Jeneström and Jansson publish summaries of their PhD theses.

Jeneström's PhD thesis, *Content Moderation and Fact-Checking: A Study of Journalists' Information Practices in the Contemporary News Media Landscape*, is a compilation thesis of four articles on content moderation and fact-checking amongst journalists and media professionals.

Jansson's PhD thesis, *A Past for Everyone? Digital Participation as Method for Democratic Cultural Heritage Production*, is a compilation thesis of four articles. Across the

articles Jansson explores digital participation around cultural heritage collections in relation to democratic ideals and inclusivity.

Both Juneström's and Jansson's thesis work, just as the other contributions to the 2023 issue, identify central aspects of the many possible information futures ahead. We will likely face increasingly complex information landscapes with new and old kinds of information, each with seemingly endless possibilities along with great challenges. The themes of the ephemeral character of digital information, how to deal with difficult content in the digital space and the usefulness and imagined dangers of AI will probably continue to occupy our thoughts in the near future. So will the push and pull of access versus privacy, inclusivity versus moderation, and other themes that are not new to the digital space but takes on new meanings in new mediums. This issue of *Tidskrift för ABM* aims to serve as inspiration for further exploration of this vast and engaging topic.

Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

ANNA-KARIN RABE [2]
BIYANTO REBIN [3]

Digitalization - a hot research topic for students

Digital technologies offer many opportunities in the LAM (Libraries, Archives and Museums) and Digital Humanities sector, and the current technological improvement rate is spurring exploration. One example is how physical cultural heritage institutions are transformed into the digital sphere, enabling people to consume and enjoy culture regardless of physical location. Digital technologies are also fundamentally changing how we preserve, sort, present, and access information, and consequently how we interpret the world. Search engines have for example moved from being largely directory-based, and thus controlled by human editors, toward automated indexing where information is classified, sorted and presented based on statistical relevance (best guesses). This creates fast and flexible knowledge management systems, but also increases the risk of bias and filter bubbles. Users also have very limited insight into the logic behind the decisions and how information is sorted, and sorted out, and it has been argued that information

2 Anna-Karin Rabe is a student member of the editorial board for *Tidskrift för ABM* and a master student in Library and Information Studies in at the department of ALM at Uppsala University
Email: annakrabe@gmail.com
3 Biyanto Rebin is a student member of the editorial board for *Tidskrift för ABM* and a master student in Digital Humanities at the Department of ALM at Uppsala University
Email: biyanto.rebin@gmail.com

literacy education should extend to AI literacy for both users and information professionals.
The Department of ALM at Uppsala University offers two master programs, one in Digital Humanities and one in ALM (Archives, Libraries and Museums). An overview of master theses published at the Department of ALM in 2023 shows that new information carriers, technologies and behaviors are key topics of interest for students in both the Digital Humanities and ALM programs. Below you find a selection of theses investigating these issues, grouped in themes identified by the authors of this editorial. There are many more interesting findings to explore, so please visit the Diva search portal for a complete overview of all theses.

Digital access

Natanya B. Newton, Piecing Together the Past - A Study about the significance of digitally accessing family records between Australia and England, and the key players responsible

Natanya B. Newton investigates the importance of accessing archival information between Australia and England and how the international gaps between countries can make ancestral research difficult. The results show that Australians studying family history rely on access to England's records because their pasts are otherwise incomplete. The study also shows that ancestral research needs both archivists, genealogists and librarians who can find and connect the information.

Jacob Leimar, Digitizing a minority and its history : A study in accessibility and digitization in Jewish cultural heritage collections and Holocaust memory

Jacob Leimar explores the challenges and considerations of digitizing Jewish cultural heritage and Holocaust memory collections in Sweden. The thesis argues that digitization offers new opportunities to make these materials more accessible, but that there are also challenges, such as funds limitations and regulations. Lemar also argues that the newly established Swedish Holocaust Museum is placed in a unique spot in challenging the current status quo in making cultural heritage deemed vulnerable accessible.

Feiye Yin, Investigating Experience of Mobile Application in Museum : The Case Study of the Nationalmuseum

Feiye Yin explores how the context of museum visit frames the experience of usage of smartphone applications. This case study finds that the visitor's experience is complicated and can be affected by multiple factors. The outcomes of this study bring some insights into the visitor experience of using museum-related mobile applications. The four contexts for the museum visit are proved to be applicable for exploring this topic from the perspective of the visitors, and thus can be regarded as a lens for understanding the visitor experience of usage of museum applications.

Social media

Monika Israelsson, “I do like the algorithm, for better or for worse”: A phenomenographic analysis of the information landscapes of adolescents on social media

Monika Israelsson studies adolescents’ conceptions of information literacy on social media, with the aim of shedding light on their everyday information life and shared information landscape. The study shows that teenagers engage in a range of information literacy practices, which in part differ from the practices taught in information literacy education. The study also indicates a gap between the adolescents’ information literacy practice on social media and the way they perceive information literacy and source evaluation.

Matilda Eriksson, ”To TikTok or not” TikTok:s usefulness for Swedish museums in appealing Generation Z

Matilda Eriksson investigates to what extent Swedish museums perceive TikTok as a useful application to reach Generation Z, and how young people are influenced by social media in relation to museums. The study shows that very few museums in the studies selection group use TikTok, and that the majority of Gen Z informants would be

motivated to visit museums if they were inspired enough by its content on TikTok or other social media. The results also indicate that there is a high possibility that the types of videos museums publish on TikTok will influence how Gen Z perceives them.

Tong Liu, Exploring the online public spheres of Zhihu, a Chinese social media platform

Tong Liu explores the discourse and sentiment characteristics of the thematic public sphere by analyzing posts around the topic of COVID-19 vaccine booster. The study showed that users were willing to use the platform for information confirmation, that the answers covered various presentations, such as videos and pictures, and that there were fewer professional people involved in the discussions. It also showed that the public sentiment was generally positive, but many users expressed dissatisfaction with the epidemic prevention measures and doubts about the effectiveness of booster vaccines.

Virtualization

Ludvig Stone, The Most Boring Game in the World : A study of World of Warcraft as a means for social interactivity within an enclosed group

Ludvig Stone explores the social facet of gaming through the lens of a World of Warcraft guild. By using Political Discourse Theory (PDT), the study found that respect for other guild members' time is the most important trait among players. This respect is primarily expressed through understanding that other members have lives outside of the game and therefore cannot devote the majority of their time to the game. The study provides insights into how social rules and codes influence gaming practices and the desirable behaviors among players.

Nikolaos Gkizis Chatziantoniou, From Pixels to Culture : Gamification and Extended Reality in the Modern Museum

Nikolaos Gkizis Chatziantoniou explores the use of gamification and extended reality (XR) in the cultural sector, specifically in museums. The thesis examines the impact of gamification and virtualization on information dissemination, visitor experience, and challenges arising from the use of these technologies. It also looks at the role of digitalization in memory institutions and how gamification and virtualization can be used to preserve cultural heritage. The findings showcase that gamification and virtualization are not just generic additions to museum exhibitions, they signify the crossing towards a new type of learning, while experiencing culture more critically and shifting the focus not only towards the visitor but also to a new type of visitor.

Strategy and implementation

Julia Unterstrasser, Linked Data and Libraries : How the Switch to Linked Data Has Affected Work Practices at the National Library of Sweden

Julia Unterstrasser employs an interview study at the National Library of Sweden, the first national library worldwide that has adopted linked data as its core data-model, to provide deeper insights into how linked data is affecting the current work practices of library professionals from their own perspectives. The findings suggest that linked data is fundamentally changing knowledge and information organization in the digital age. Furthermore, the results of the study suggest that linked data is only part of a paradigm-shifting change currently happening in the knowledge and information organization community.

Amanda Holm & Pernilla Hjelm Andersson, Contracting for digital archives: A study of Swedish municipalities requirement specifications

Amanda Holm and Pernilla Hjelm Andersson investigate the contracting of digital archives for Swedish municipalities and the needs and requirements that govern this. The study concludes that there still are considerable steps to take before

reaching an e-archive solution that can be used on a broad scale to accommodate many different needs and restrictions. At the same time, the development of long-term preservation systems for digital records would benefit from more standardized guidelines on a national scale.

Aijia Zhang, Attributing Digitalization Decisions in Museums: A Multiple Case Study of Swedish Public Museums

Aijia Zhang's study employs the Digital Curation Lifecycle Model to conceptualize the process mapping for digitalization projects and the attribution theory to identify the internal and external factors affecting digitalization decisions. The findings indicate that the attribution of museums' digitalization decisions is complicated and can be affected by multiple factors and can be multiple attributed. The study underscores the importance of contextuality and underpins the emerging trend in museums' digitalization practices.

Wilma Enström, Digital Differences within Cultural Cornerstones: A Case Study of the Digital Tools and Implementations at the Museum of Ethnography in Stockholm

Wilma Enström explores the role of digital tools and implementations in the museum exhibitions and examines the potential discrepancies of both museum employees and visitors. Through interview, on-site observations, and one participatory observation, the study finds that digital tools and implementations are highly regarded as invaluable assets. The thesis contributes to the field of Digital Humanities by providing new insights into the effective utilization of digital tools in cultural heritage museums.

Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

RIKARD FRIBERG VON SYDOW[4]

”A Conspiracy of Lost Content” - The Case of ”Q” and QAnon

Abstract: Starting in November 2017, the anonymous poster ”Q” filled its followers with messages regarding American politics. The QAnons – the movement formed around the messages from Q – were fierce online discussants occupying almost every possible social media platform. Although, this online presence would not last as platform after platform banned the QAnons. The author uses a theory and method in which the Q-posts are described as part of a digital ecosystem and tries to answer questions regarding if it is possible to preserve the content of a deplatformed movement thus exploring the limits of digital preservation of interlinked content created by a community.

Keywords: deplatforming, digital humanities, digital preservation, interlinked content, kickbanning, social media, QAnon,

4 Dr Rikard Friberg von Sydow is Senior Lecturer in Archival Science and Head of the department of Archival Science and Library and Information Science at Södertörn University, Stockholm. His research focuses, among other things, on the history of information management and on studies of internet culture, especially in its more extreme forms.
E-mail: rikard.friberg.von.sydow@sh.se

Introduction

> ”The uh human species...Homo sap… (laughter) is perhaps two million years old… prehistorians keep pushing our birthdate further back… perhaps an abortion would be the uh simplest solution… (laughter) but the incidence of clearly recognizable artifacts dates back only fifty to a hundred thousand years. In that modest span gentlemen, we have come from stone axes and spears to intercontinental missiles with nuclear warheads… the same principle as the spear but rather more efficacious… (laughter). Is it not feasible that other cultures may have traveled the same road and disappeared without a trace? Nor can we rule out the possibility that artifacts were deliberately destroyed. The river people of New Guinea fashions masks for their festival which are burned once the festival is consummated. And what would a historian make of pseudo artifacts of modern art? Who is that artist who does a barrelful of nuts and bolts? He went on to burnt kitchen chairs … oh yes … Armand … How could our future scholar know that this artifact commemorates the sale of a name. It’s an Armand and worth so much just as the coppers of Kwakiutl potlatches were valued according to the transfers they had accreted.”
>
> William S. Burroughs, *The Place of Dead Roads*

From November 2017 to December 2020 the mystical anonymous poster ”Q” filled their followers with messages regarding American politics. Q acted like the person behind the handle was someone who had insight into both the intelligence community and important political circles. Q posted on different image boards and as the number of followers grew different web pages were created where canonical Q-posts (usually called ”drops” or Q-drops”) were preserved, numbered, tagged, and interlinked. The QAnons – the movement formed around the messages from Q – were fierce discussants on the Internet occupying almost every possible social media platform. QAnon is a baseless movement combined with a syncretic conspiracy theory presenting an all-you-can-eat-buffet of recycled conspiracies; lizard men, a blood-drinking paedophile elite, underground tunnels where secret people live etcetera. Baseless because it does not have a hierarchical organization, consisting mostly of isolated individuals in front of internet-connected screens (Bloom & Moskalenko 2021). Syncretic

because it recycles a collection of different earlier conspiracy theories and incorporates them into a compilation of QAnon-approved narratives. These different conspiracy theories dwell within the movement and is why it sometimes is described as a "Big Tent Conspiracy Theory" (Roose 2021). It is a tent under which lots of different world views can mix and find common ground – a form of "Meta Conspiracy Theory" (Bloom & Moskalenko 2021, p. 33). There is also an idea in media and on internet forums that QAnon is not an ordinary conspiracy theory, that it is an internet-based role-playing game gone wrong. A way of sorts to hack reality and create a substratum where a virtual reality can grow, where it is possible to perform an advanced role play. If this is the case, the movement itself may not be aware of this, the QAnons being a sort of non-player personalities (NPCs) tricked into performing in a certain way. Behaving like an internet-based flash mob with instructions, yet with no insight into the background of their acting (Financial Times 2020). Nevertheless, this has not been proven other than on a mere speculative level and will not have any impact on the work performed in this article.

It is often claimed that the QAnon movement reached its peak when the QAnons spearheaded the attack against the Capitol on the 6th of January 2021. During the autumn of 2020 almost all big social media platforms started to purge QAnons, kickbanning them and thus removing their accounts and their created content (BBC 2022). After the purge the movement now lives on in a crippled way – in the open on smaller platforms like GAB and Bitchute – and in a more covert way among those followers whose digital presence survived the purge and still holds accounts on mainstream social media platforms. Instant messaging applications like Telegram have also become a haven for QAnons, especially using the application's possibility to create chat groups.

Previous research

QAnon has of course been investigated by others, although its brief history makes the list of publications short in comparison to other fringe movements. Mia Bloom and Sophia Moskalenko, a professor of communications and a researcher in psychology, published what to my knowledge was the first academic book dedicated to QAnon, *Pastels and*

Pedophiles – Inside the Mind of QAnon in 2021. Besides describing the movement and its history, they focus on the psychological processes behind the followers and the consequences for families when a member becomes entangled in conspiracist belief (Bloom & Moskalenko 2021). If we search through research articles regarding QAnon some disciplines stand out; psychology, and other disciplines within the behavioural sciences, informatics, and computer science. Within the behavioural sciences, there is a will to explain why the movement has appeared and who joined it. This can be done by using Jungian analytical psychology as in Jim Kline's article *Eat My Flesh, Drink My Blood Archetypal Elements in the QAnon Conspiracy* or using theories regarding political aggression as in Michael A. Jensen and Sheehan Kane's *QAnon-inspired violence in the United States: an empirical assessment of a misunderstood threat* (Kline 2021, Jensen 2021). Another perspective from the social sciences that has been investigated is the social classes within the QAnon-movement, in Isaac Kamola's article *QAnon and the Digital Lumpenproletariat* (Kamola 2020).

In Informatics and Computer Science examples worth noting are Hanly et al *No Calm in The Storm: Investigating QAnon Website Relationships* in which relationships between different QAnon websites are investigated through a web crawling-technique in which a program follows links in QAnon-content from different message boards to social media sites and onward. This research gives us the possibility to understand the ecosystem in which QAnon resides (Hanly et al 2021). There are also several interesting works in the interdisciplinary field that sometimes is named Digital Humanities. Peter L. Forberg's article *From the Fringe to the Fore: An Algorithmic Ethnography of the Far-Right Conspiracy Theory Group QAnon* is one of those that investigates QAnon from an interdisciplinary perspective in which both ethnography and analysis of algorithmic techniques are used. In this case with a focus on how the QAnons use search algorithms to boost the spread of the movement on the web (Forberg 2021). Another example from this field is Paul Bleakley's article *Panic, pizza and mainstreaming the alt-right: A social media analysis of Pizzagate and the rise of the QAnon conspiracy* in which he investigates how the QAnon-movement used the Pizzagate-hashtag to boost their relevance on the net, and how the use of debunked conspiracy theories doesn't seem to hinder a movement from growing (Bleakely 2021). QAnon, being a

contemporary political movement connected to different political incidents, has of course gained a lot of attention from journalists. At least one example needs to be mentioned here. Journalists connected to the independent international research collective Bellingcat have written about the plasticity of QAnon and how the ”big tent” conspiracists seem to have an advantage through the possibility of pacing back and forth between different views and opinions, having an ideological kernel that doesn’t need to apply to the unwritten rules of political movements (Tian 2021). So far there are, to my knowledge, no research directed towards the preservation of the content left behind by the movement and this is where this article fits into the puzzle of QAnon-related research.

If we want to preserve a digital source, it will require regular tending. If it is to be preserved over a longer period we have to control our possibilities to read and interpret the material even though the technological environment has changed (Smith Rumsey 2016, p. 148f). If not tended, several problems can be discovered by a later retriever of the source; the delivery technology (the tape, the server) could have malfunctioned or the file format is not readable anymore (Jenkins 2006, p. 13). Or it could be gone, forgotten or misplaced. When we are dealing with sources being curated by a controversial movement we need to ensure that the sources, a website, a forum, presence on social media etcetera, do not disappear suddenly. This could happen if the internet service provider decides to remove the webpage due to its content, if the owner forgets to pay the bill for website hosting, or if a social media platform decides to ban certain expressions in text or in memes – and for many other reasons. We can very fast be faced with a situation where certain sources suddenly are removed. *MySpace*, the once popular social media platform, lost 12 years of user-uploaded music in 2019 when files ended up corrupted after server migration (Hayes 2019). There are even examples of internet services that focused on preserving digital material that has gone bankrupt leading to users losing their material. Like the Swedish life-logging service *Narrative-Memento* (Frigo 2017, p. 18). If we want to preserve a source, especially if we don’t control it ourselves as curators or creators, we need to perform some preventive work regarding that source to keep it from being lost in the future. I have performed some work regarding QAnon related to preservation, but when doing this preservative work I have

also seen the limits of the possibilities to preserve internet content. What I claim that those limits are will be clear to the reader further on in this article.

Preservation of cultural heritage

In January 2021 I started to preserve all Q-drops by downloading them in raw text format and saving them as individual files. The content was also saved in picture format, through a screenshot of each post from one of the sites that preserve the canonized drops. It is a very interesting material, full of dog-whistling, echo-chamber-like links to supporters on different social media platforms, and memes whose messages are very internal to the movement. Downloading the posts as text created the possibility to use different text processing tools to investigate the corpus. It also created an insight into what happens with preserved internet content when links suddenly are broken and the movement surrounding the content is banned. This insight transformed into the main questions that will lead the discussions in this text: 1) How do we preserve (and can we preserve?) the content of a deplatformed movement? and 2) What are the limits of digital preservation of interlinked content created by a community?

The title of the article suggests that the Q-posts are part of a cultural heritage. Cultural heritage is "the sites, things, and practices a society regards as old, important, and worthy of conservation" (Brumann 2015). In general, cultural heritage organizations consider digital objects as worthy of conservation, although there is an insight that the problems of conserving digital objects have not been fully understood yet (Cameron & Kenderdine 2010). Interlinked content is one of the problems of conserving digital objects. The reader might have encountered the "Wayback Machine" of the Internet Archives (Internet Archives 2023). This is a site that provides the service of archiving any website on a specific date and gives the user a possibility to surf a website as it was on a specific date – if someone has made a copy of it. If you enter the "Wayback machine" you will notice one of the problems with preservation of interlinked content. Most of the links on the websites are dead and linked content as advertising are empty boxes. If we consider the World Wide Web as an actual web of content linking to other content this does not resemble what the user experienced when visiting

the websites before they were preserved in the archive. What we have left is a dysfunctional copy of the website. Without the majority of its web-like characteristics.

It might be time to wrap up the introduction and the description of previous QAnon-research and continue with a description of the outline of what is going to take place below. After this introduction, the theoretical concepts used in the study will be explained. This is followed by a description of the methods used. After theory and method, we are fit to investigate some problems with preserving the deplatformed QAnon movement and its maimed remnants. Finally, some conclusive remarks and possible further research will be presented.

Theoretical concepts; Ecosystem, Authenticity and Provenance

There are at least three theoretical concepts that are suitable when investigating and analyzing the remnants of Q and QAnon from an archival science perspective. First what we could call an "Ecosystem", related to the interlinked qualities of the content. The other two are the concepts of "Provenance" and "Authenticity" that are connected to, among others, the discipline of Archival Science, and are used to describe qualities of archival content. The two concepts are interdependent on each other, we cannot assess authenticity without a perception of the provenance and we cannot build a perception of the provenance without assessing authenticity. The idea of a digital ecosystem can be argued, if it is described as I will describe it below, to be connected to both provenance and authenticity. I will start by describing the three concepts separately, but there are, as the reader will notice, several overlaps between them.

Ecosystem

In this text parts of the Internet will be described as environments with certain traits; an ecosystem. An Ecosystem is a system containing a community of animals, humans, or other living organisms (usually called an ecology) and a physical environment in which this community lives, a rotting log, a social media forum (Townsend et al 2008). What makes the Internet an

ecosystem is the connections between different nodes within the system and the possibility, under the right circumstances, for one node to affect another. This is because of a network structure in which nodes are "connected through a disordered pattern of many different interactions." (Caldarelli & Catanzaro 2012). The network structure is evident on both the technical level of the Internet and on the social level. On the technical level, the TCP/IP protocol connects servers and clients in a network. Within the web servers the social interaction, the social networks, has its foundation. It provides interfaces for the users to interact and databases in which these interactions are stored. The Internet consists of, to use Christian Fuchs's words: "a technological system and a social subsystem that both have a network structure" (Fuchs 2017, p. 41).

It could be easy to claim that the technical level is the starting point of the World Wide Web as a network, but even before the technical foundations were created, and before our present social global connections, there were ideas of local interlinking of information and of possibilities to collaborate in the development of information resources. Vannevar Bush wrote his paradigmatic essay "As We May Think" in 1945 using his experience of the need for fast information research during World War 2 to imagine a machine with infinite information storage. More important though – Bush discussed the possibility of "trailblazing", the act of creating codes that would link two entities of information to each other. And this is before zeroes and ones in any practical way entered the information storage scene (Bush 1945). Ted Nelson was the researcher who during the 1960s designed a large part of the idea of linking information between different text (Barnett 2015, p. 61). However, links as many of us are used to today, with some sort of highlighting et cetera, were introduced in 1984 by one of Nelson's disciples (Schneiderman 2015, p. 7).

Technically, when we are linking something on a web-based platform, a simple line of code orders the computer to make one part of the screen, often a piece of text, become clickable and when clicked, move the viewer to another document. In HTML this is done by a href-tag, but on most social media platforms there is no need for the user to know what happens behind the scene. Socially there are more components. A user decides that a piece of information should refer to another piece of information thus motivating the creation of a link between the two documents that

contain the two different entities of information. Often a user can do this without any need for knowledge about the technique behind, copying and pasting a link. The social internet today has the same role as our vocal cords – we can use them without really knowing how they work or even without knowledge regarding if they even exist or not. The interfaces we use hide the technical aspects. Link upon links can then be created by myriads of users with different purposes, from referring to proof in an argument to recommending a text, video, etcetera that another user should see. One of the more prominent internal jokes of the Internet culture is the so-called "Rick-rolling", tricking someone into clicking on a link that leads to a video of Rick Astley's 1987 hit "Never Gonna Give You Up" (Brown 2022). Together the links and the users create an ecosystem that feeds on both technical and social premises. A link can break both for technical reasons – a moved webpage, a crashed server, and for social reasons, a removed post, or a banned webpage. In QAnon's case, there is a history of links being broken because the content linked is often seen as controversial outside the movement. QAnon is a movement that has been banned from different platforms and this kind of broken linkage can be connected to the social part of the description of a network.

Authenticity and Provenance

Provenance is, as mentioned earlier, a concept found in, among others, the discipline of Archival Science. It describes the context in which a unit of information is created. Which organization that has created it and, during which time and so forth. It is an important organizing principle in Archival Science and generally the key organizing principle in most archives, globally, since the 19th century (Douglas 2017, p. 26). Provenance can also be used as a foundation for the next principle I will describe; Authenticity. A concept that provides us both with a context and a creator of a specific unit of information.

In the discipline of information security, authenticity is a criteria connected to proper attribution; the ability to prove that a certain unit of information is created by a certain organizational entity (Andress 2015, p. 8). In the discipline of Archival Science this definition of the concept is commonly criticized as being too simple. Authenticity cannot be argued to be an intrinsic character of an

information entity, authenticity is created through different circumstances surrounding the entity. This is an experience that Archival Science has inherited from the medieval discipline of Diplomatics, a discipline focusing on evaluating legal documents, often legal document related to ownership. At the end of the nineteen-eighties, Archival Science scholar Luciana Duranti used concepts from Diplomatics to analyze digital content (Duranti 1989, p. 12). This led to an ongoing debate in Archival Science. Often there are two divisions of authenticity used in Diplomatics: legal and historical authenticity. Legal being if the information (often, in the pre-digital era some kind of paper document) has been ruled legitimate or false by a court of law, historical, if there is evidence of fidelity towards the entity of information: if people generally have believed it as authentic (Mak 2012, p. 4ff). I will later on present my definition of authenticity in relation to the Q-drops, leaning more towards the concept of historical authenticity.

Methodology

Two methods need to be explained here. First, the method which is connected to the work in this article. The desired result is to describe the ecosystem of QAnon and the processes within that ecosystem in a way that makes the reader understand where the information exist in this ecosystem and thus what we need to preserve to have a possibility to understand the movement in retrospect. I will do this in two parts, first describing the ecosystem and the processes within it, then describing the breakdown and what it did to the material. The next method is connected to the qualities of the preserved material, the Q-posts. What in this material we can preserve and how the material can be assessed using the two concepts of provenance and authenticity. The first method is thematic and descriptive. Describing the QAnon movement with the premise that it is an ecosystem and that our understanding of the movement's processes will give us a clearer view of how it works. Describing the breakdown of the movement will supply us with the premise that the deplatforming broke certain processes within this ecosystem. The second method uses the concepts of provenance and authenticity to perform an analysis of the QAnon movement with a focus on the Q-drops. Using these concepts, we will have a clearer view of

the qualities of the preserved drops and how they can, and cannot, be used in research in the future.

Posts interlinked within posts interlinked; an ecology

If we study the canonized Q-drops it is possible to create a timeline of Qs career. Q first published 4953 drops starting on the 28th of October 2017 and later became quiet on the 8th of December 2020. The silence lasted until 2022, when Q published a number of drops slightly different than during the earlier period. These later drops are less connected to contemporary politics than those from the earlier period and more esoteric in comparison with earlier posts (OperationQ 2023). During the first period, ending in 2020, Q and its followers eventually started to develop a pattern. I call this pattern the Q-process. It starts when Q posts a drop on one of the image boards used as a "dead drop", mainly 4Chan and 8Chan (later called 8Kun). "Dead drop" is a term used in espionage describing a secret place where a spy can drop information that later can be picked up by another spy (without a need for the two persons to meet each other). The followers identify the post as a Q-drop mainly by the use of a specific user ID, but also through style, content etcetera. They pick it up and start to analyze it. The part of the process that started after Q publishes a drop has been described with the expression "Breadcrumbs and Bakers". The idea behind this description is that Q leaves "breadcrumbs", clues regarding what happens behind the scenes in politics. These breadcrumbs are picked up by "bakers" who put them into the context of the conspiracy theories included in QAnon's big tent (Bloom & Moskalenko 2021, p. 9). The bakers are people that are invested in QAnon and have a platform (a podcast, a Youtube-channel et cetera) where they have an audience. The bakers create bread that the QAnons, the affiliates of this conspiracy theory, consume. Bread could be described as processed and contextualized information. Information that descends from the Q-drops. Even if this is a rather uninformed way of describing the work process of a real bakery, it is the way this group describes their process of information gathering and distribution (Schwartz 2018). Each part of the process, the drop, the canonization, and the

analysis are nodes in a network containing links, assumptions, references and so forth.

Take these broken links; the breakdown

QAnon eventually got purged from all major social media platforms, beginning with *Reddit* as early as 2018 and continuing with *Facebook*, *Twitter*, and *Youtube* during the fall of 2020 (Bloom & Moskalenko 2021, p. 45). Purge in this case means practically that their subreddits were deleted, their groups were deleted, and certain users that spread QAnon-related material had their accounts deleted. This meant not just that the ecosystem that Q and the QAnons had created was destroyed, however it also meant that vast amounts of content were gone. One of the distinctive traits of the collapse of a network is that it is very hard to get an overview of the results of the collapse. Especially to conquer this overview before the actual collapse happens (Caldarelli & Catanzaro 2012, p. 42). So, when the QAnon movement was virtually destroyed on most prominent social media platforms, it was not only the movement that was destroyed. A lot of information regarding the movement fell with it – who the affiliates were, what they had posted on the platforms, and how they were connected.

Practically what happened is that nodes disappeared. Suddenly links were broken and accounts were non-existent. The World Wide Web is a lazy discourse with a tradition of linking information rather than retyping it. When a post is removed, the content is lost. Deplatforming is a common practice in contemporary society regarding groups that spread false or hateful messages. In this context, when the existence of a movement and the movement's communication are almost identical the deplatforming will also affect the sources needed to investigate the movement.

As mentioned at the beginning of this article, in early 2021 I saved all Q-drops in txt-format creating the possibility to use text-analysis tools such as GREP and AWK to analyze the content. These tools make it possible to find specific words, link-types etcetera searching all individual text-files simultaneously. This work has resulted in three poster presentations, one regarding the existence of direct

references to established conspiracy theories in the Q-drops, a scope which does not have any larger impact on the analysis performed in this article (Friberg von Sydow 2021). Poster presentation 2 describes an empirical investigation of linked content in the Q-drops using Twitter links as an example. In the 4953 Q-drops, there are 1509 links to tweets. Often used by Q as the beginning of an argument or as examples and references. Of these tweets, 43% belong to suspended accounts and cannot be retrieved today. Added to this 3.8% of the tweets are deleted (by Twitter or the tweeter) and 3% are retweets from suspended accounts – making around half of the tweets linked in the Q-drops unavailable today (Friberg von Sydow 2022). The third poster presents a similar work as in poster 2, but uses Youtube.com as the empirical example. The result is similar – a great portion of the linked content is missing today, although only a few years has passed (Friberg von Sydow 2023). In general most of the dead links are due to suspended accounts making the broken network, and the loss of content and context, a result of the purge against the QAnon movement on Twitter and other mainstream social media providers.

Is it possible to preserve this? And what are we preserving?

Even without the purge the content of these kind of movements are very hard to preserve. There is also an important question that need to be discussed: what should actually be included when preserving the digital remnants of a network-based movement? To answer this we need to analyze the Q-drops using the two concepts that are vital to preserve the qualities of a collection of information; authenticity and provenance. Regarding provenance, there are not any credible knowledge regarding the actual source of the Q-drops. This is one of the main ingredients in the concept of Q – that the poster is anonymous. We have to build provenance on other arguments, and lean more towards the audience and their view of how to define a Q-drop. There are some uncanonized early Q-drops that are not included on collection-sites as “OperationQ” (Bellingcat 2021). There are also investigations analyzing the language style used in the Q-drops that point towards that they are the

product of more than one writer (Aliapoulios et al 2021). Despite that, the Q-drops are still a collection of information entities that a group of followers put their trust in. If we want to understand the Q-drops and their use it seems realistic to ignore these problems of provenance and authenticity, and build arguments that are rooted in the belief of the movement; the QAnons. The content of the Q-drops is, outside the QAnon-movement, generally argued to be false – the first Q-drop – as an example – claims that Hillary Clinton is going to be arrested soon (posted in October 2017). This is known to be false, so it would be problematic to try to build arguments regarding authenticity using the content of the drops themselves. I have constructed a model describing different categories of authenticity that is possible to use to analyze Q-drops. "Authenticity through content" is generally not usable on the Q-drops. They are believed in by the QAnons, and this is the important aspect. It is much more plausible to argue towards an "Authenticity through the faith of followers", the follower's faith in the drops being what makes them stand out from other posts on 4Chan and other image boards, that the QAnons will ignore. It is also possible to claim relevance through another possible categories of authenticity. A "Process-based authenticity" – claiming that authenticity is created through a process of Q posting a drop on a platform, the QAnons claiming it as a true Q-drop thus including it in the Canon of Q-drops. This category would also focus on canonization itself, ignoring the process that appears before the actual inclusion among the canonized drops. Using these categories it is possible to argue that the Q-drops are authentic and it is also possible to argue that there are traces of provenance that knit them together as a collection. Being a part of QAnon is to have access to a (belief-dependent) information classification system, that can be used to analyze content related to such diverse phenomenon as World politics or gossip regarding Hollywood stars. The Q-drops are a part of this classification system. The classification system of the QAnons is not in essence, different to other classification systems. A classification system, to use the words of the library- and information scientists Geoffrey C. Bowker and Susan Leigh Star is "a spatial, temporal or spatio-temporal segmentation of the world" (Bowkers & Leigh Star 2000, p. 10). The spatio-temporal segmentation, in the case of QAnon, is the internet-content, the nodes that make up the network – the ecology.

Can we preserve an ecology? No - a net breaks if we start to remove the central nodes. It rips if we start to remove nodes on the edges. We have to sew it together again if it is impossible to recreate the node that we have lost. If we want to preserve the totality of the drops – including links – we have to perform an appraisal of the material and choose what to preserve and not to preserve. Linked content can change over time, which makes nothing but a slice of the total ecology possible to preserve. We have to choose what we preserve and when we preserve it. Make a collection of relevant material. Using a preservation model, as OAIS, the Open Archival Information System, we have tools to describe such a collection, describing what has been preserved and when it was preserved (DLM 2017). More important – We would need to describe what is missing – where the net is broken and ripped. In the corpus-analysis described above, where broken links were identified those rips and breaks are shown. Such a collection can then be used by future researchers, who through the descriptions of its premises will be aware of the qualities of its content.

When creating such a collection some kind of active preservation during the period of time when the movement is active, would be preferred in order to not lose large amounts of information due to purges et cetera. This is not an easy method though, it requires timing, starting the preservation in an early phase of the movement's existence. It requires great knowledge of in which platforms the movement is active and how to monitor these platforms and preserve the content posted. In many ways an active preservation under these circumstances requires multiple skills both the skills of a researcher in the field of online political movements in order to identify which content that is useful for future research, and skills of digital preservation in order to preserve the content properly. There are several risks involved in using such a method if the goal is to create preserved digital content that can be used by future researchers. Two risks worth mentioning are 1) identifying the right material, which always will be a task involving some speculations and 2) preserving the material in such way that links and other connections within the material is intelligible for future researchers.

Conclusive remarks

The quote from William S. Burroughs late novel *The Place of Dead Roads* found at the beginning of this article describes the confusion that the future observer of ancient artefacts can be taken with. A broken link in digital research material should affect the future observer in a similar way. A piece of the puzzle is missing and the distant observer, probably with less knowledge of the context than a contemporary observer, will have great problems understanding the material. Deplatforming is a very efficient tactic in the wars of the information society. It disarms the combatant, making them voiceless in a war that mostly is performed through discussions and disruptions. However, it also threatens to destroys the scarce sources of future research in political movements – especially if these movements are seen as controversial in their contemporary setting. The loss of information that is a consequence of deplatforming can also affect the possibility of proving what a movement really stood for. It might remove hate speech – yet it also removes the proof that the hate speech occurred in the first place. Deplatforming as a method to stop online movements who spread hatred or false information, can seem uncontroversial, but those recommending it should reflect on its consequences: Before tearing down the enemy fortress of content, reflect on the content's possible future use. As evidence of both possible crimes and as a cultural artefact. Don't let earlier cultures travel down the road and disappear without a trace. It's a Tweet and valued through how many times it was retweeted - just as the coppers of Kwakiutl potlatches were valued according to the transfers they had accreted. We should preserve the knowledge of its value for the future.

References

Aliapoulios, Max et al (2021) *The Gospel According to Q: Understanding the QAnon Conspiracy*, https://arxiv.org/abs/2101.08750 ,Visited 2023-10-30.

Andress, Jason (2015) *The Basics of Information Security – Understanding the Fundamentals of Infosec in Theory and Practice*, 2nd edition, Oxford: Syngress.

Barnett, Belinda (2015) ”The Importance of Ted’s Vision” in Dechoff, Douglas R & Struppa, Daniele C I*ntertwingled – The work and Influence of Ted Nelson*, Cham: Springer Open Access.

BBC (2020) *Twitter suspends 70,000 accounts linked to Qanon*, https://www.bbc.com/news/technology-55638558 , Visited 2023-10-30.

Bellingcat (2021) *How Q’s 'Lost Drops' Undermine the QAnon Myth*, https://www.bellingcat.com/news/rest-of-world/2021/04/22/how-qs-lost-drops-undermine-the-qanon-myth/ Visited 2023-10-30.

Bleakely, Paul (2021) *Panic, pizza and mainstreaming the alt-right: A social media analysis of Pizzagate and the rise of the QAnon conspiracy*, Current Sociology, 1-17.

Bloom, Mia, Moskalenko, Sophia (2021) *Pastels and Pedophiles – inside the mind of Qanon*. 1th ed, Stanford: Redwood Press, 2021.

Brown, Helena (2022) *Never Gonna Give You Up — how Rick Astley’s 1987 hit became a global meme*, Financial Times. https://ig.ft.com/life-of-a-song/never-gonna-give-you-up.html, Visited 2023-10-30.

Bowkers, Geoffrey C & Leigh Star, Susan (2000) *Sorting Things Out – Classification and its consequences*, Cambridge: MIT Press.

Brumann, Christoph (2015) ”Cultural Heritage” in *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier.

Burroughs, William S. (2015) , *Place of Dead Roads,* London: Penguin Books.

Bush, Vannevar (1945) *As We May Think*, The Atlantic, July 1945. https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/ , visited 2023-10-30.

Caldarelli, Guido & Catanzaro, Michele (2012) *Networks – A very short introduction*, Oxford: Oxford University Press. 2012.

Cameron, Fiona, Kenderdine, Sarah (2010) *Theorizing Digital Cultural Heritage : A Critical Discourse,* Cambridge MA: The MIT Press.

DLM (2017) *Archival Standards Board General SIP Specification.*

Douglas, Jennifer (2017) ”Origins and Beyond: The Ongoing Evolution of Archival Ideas about Provenance” in MacNeil, Heather & Eastwood, Terry *Currents of Archival Thinking*, Santa Barbara: Libraries Unlimited.

Duranti, Luciana (1989) *Diplomatics: New uses for an old Science*, Archivaria 28.

Financial Times (2020) *Is QAnon a game gone wrong?*, https://www.ft.com/video/372cac40-0f6f-498b-8c19-7b635142296e , visited 2023-10-30.

Forberg, Peter L (2021) *From the Fringe to the Fore: An Algorithmic Ethnography of the Far-Right Conspiracy Theory Group Qanon*, Journal of Contemporary Ethnography, 1-27.

Friberg von Sydow, Rikard (2021) The Q-drops – A content analysis, European Conference of Information Literacy.

Friberg von Sydow, Rikard (2022) *Take these broken links – Twitter, the Q-drops and the collapse of a digital Ecosystem*, Digital Humanities in the Nordic and Baltic States.

Friberg von Sydow, Rikard (2023) *Preserving the cultural heritage of conspiracists: The Case of ”Q” and Qanon*, ICA CONGRESS 2023, Abu Dhabi.

Frigo, Alberto (2017) *Life-stowing from a Digital Media Perspective: Past, Present and Future*, Stockholm: Södertörn University.

Fuchs, Christian (2017) *Social Media – A Critical Introduction*, Los Angeles: Sage Publishing.

Hanly, Hans W.A, Kumar, Deepak, Durumeric, Zakir (2021) *No Calm in The Storm: Investigating QAnon Website Relationships*, ICWSM.

Hayes, Dade (2019) *MySpace Loses 12 Years Of Uploaded Music Due To A Server Issue*, Deadline,

https://deadline.com/2019/03/myspace-loses-12-years-of-uploaded-music-due-to-a-server-issue-1202577303/, visited 2023-10-30.

Internet Archives *The Wayback Machine*, https://archive.org/web/ ,visited 2023-10-30.

Jenkins, Henry (2006) *Convergence Culture – Where old and new media collide*, New York: New York University Press.

Jensen, Michael A & Kane, Sheehan (2021) *QAnon-inspired violence in the United States: an empirical assessment of a misunderstood threat*, Behavioral Sciences of Terrorism and Political Aggression.

Kamola, Isaac (2020) *QAnon and the Digital Lumpenproletariat*, New Political Science, vol 43, no 2.

Kline, Jim (2021) *Eat My Flesh, Drink My Blood Archetypal Elements in the QAnon Conspiracy*, Jung Journal, 15:3, 36-52.

Mak, Bonnie (2012) *On the use of Authenticity*, Archivaria 73.

OperationQ, https://operationq.pub/ ,visited 2023-10-30.

Roose, Kevin (2021) *What is Qanon, the viral Pro-Trump Conspiracy Theory?* New York Times, 2021-09-03 https://www.nytimes.com/article/what-is-qanon.html ,Visited 2023-10-30.

Schneiderman, Ben (2015) ”Odes to Ted Nelson” in Dechoff, Douglas R & Struppa, Daniele C *Intertwingled – The work and Influence of Ted Nelson*, Cham: Springer Open Access.

Schwartz, Mathathias (2018) *A Trail of ‘Bread Crumbs,’ Leading Conspiracy Theorists Into the Wilderness* , New York Times Magazine, https://www.nytimes.com/2018/09/11/magazine/a-trail-of-bread-crumbs-leading-conspiracy-theorists-into-the-wilderness.html visited 2023-10-30.

Smith Rumsey, Abby (2016) *When we are no more – How digital memory is shaping our future*, New York: Bloomsbury Press.

Tian, Edward (2021) *The QAnon Timeline: Four Years, 5,000 Drops and Countless Failed Prophecies*, Bellingcat. https://www.bellingcat.com/news/americas/2021/01/29/the-qanon-timeline/ visited 2023-10-30.

Townsend, Colin R, Begon, Michael, Harper, John L (2008) *Essentials of Ecology*,Oxford: Blackwell Publishing.

Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

SIMONE MAGAARD[5]
DANIEL MARKUS[6]
ANN-SOFIE KLARELD[7]

Genomic privacy? Direct-to-consumer genetic testing companies and the processing of sensitive personal data

5 Simone Magaard received her master's degree in archival science from Lund University in 2021 and has been working as an archivist since. Currently she works for the Region Skåne archives.
Email: simone_magaard@hotmail.com
6 Daniel Markus received his master's degree in archival science from Lund University in 2021. Today he is working as an archivist for the Swedish municipal administrative authority VA SYD.
Email: daniel_markus@hotmail.com
7 Ann-Sofie Klareld is a senior lecturer at the Department of Arts and Cultural Sciences, Lund University. She has a PhD in Archives and Information Science from Mid Sweden University. Klareld teaches at the Master's Programme in Archival Studies, Library and Information Studies, Museum Studies (ALM). Her current research interests include archives in society, the professional role of archivists, and contemporary family history.
Email: ann-sofie.klareld@kultur.lu.se

Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

Abstract: This paper explores the processing of sensitive personal data for genealogical purposes by studying how the privacy policies of three direct-to-consumer genetic testing companies comply with the General Data Protection Regulation (GDPR). The methodological approach is content analysis. The privacy policies are analysed in relation to the three core principles of 1) adherence to GDPR; 2) transparent processing of personal data; and 3) consent. The paper also discusses privacy in relation to third parties whose sensitive personal data is processed indirectly due to the nature of DNA as shared between family members. The main result is that the privacy policies rarely mention GDPR; when mentioned it is done in a brief way; the privacy policies studied convey insufficient transparency as regards data processing; and the current legal framework does not suffice to protect the privacy of genetic relatives of users whose personal data is being processed indirectly.

Keywords: personal data, privacy, genetic data, GDPR, genealogy

Introduction

Direct-to-consumer genetic testing (DTCGT) are DNA-testing products that consumers can purchase over the counter or online. DTCGT is a rapidly growing industry developed alongside advances in genetic and genomic science, arguably transforming family history by enabling genetic genealogy (Stallard & De Groot, 2020). Genetic data warrants a high standard of privacy protection since 1) it may be used to identify predispositions, disease risk, and predict future medical conditions; 2) it may reveal information about the individual's family members, including future children; 3) it may contain unexpected information or information of which the full impact may not be understood at the time of collection; and 4) it may have cultural significance for groups or individuals (Future of Privacy Forum, 2018). The General Data Protection Regulation (GDPR) prohibits processing genetic and biometric data with several exceptions, including when "the data subject has given explicit consent to the processing of those personal data for one or more specified purposes", and when "processing relates to personal data which are manifestly made public by the data subject" (GDPR, 2016, Article 9).

DTCGT companies do try to avoid, to a reasonable extent, health informative markers and privacy policies exist to ensure that consumers can make informed decisions. However, researchers have raised concern regarding transparency related to confidentiality, privacy, and secondary use of data (Howard, Knoppers & Borry, 2020), compliance with international guidelines (Laestadius, Rich & Auer, 2017), and information provided to third parties (Wallace et al, 2015). A recent Swedish report indicated problems concerning one DTCGT company's' privacy policy and their terms and conditions by concluding "a number of the terms that Swedish consumers agree to should be considered unreasonable" (Swedish Consumers' Association, 2020, our translation).

On the basis of international research pointing to problems with privacy policies in general and the Swedish Consumers' Association pointing to specific problems regarding one DTCGT company operating in Sweden, this paper explores the privacy policies of three DTCGT companies offering their products and services to Swedish customers. The privacy policies are read in relation to GDPR, and their implications are discussed in relation to personal integrity. Note that the paper does not claim to assess the industry as a whole, nor point out any company as 'worse' than any other. It does, however, take a critical standpoint as the guiding principle, discussing privacy protection in general and sensitive personal data in particular.[8]

Aim and purpose

The paper contributes to the field of research that concerns the tensions between sensitive personal data, law, and new technical possibilities. Our purpose is to explore and problematize what three selected DTCGT companies say in their privacy policies about their processing of genetic personal data, and discuss the limitations of the legal framework in relation to the nature of DNA as shared between different individuals. The research questions are:

[8] The concepts used in the paper are defined according to GDPR, Chapter 1, article 4.

1. How do the three companies' privacy policies comply with the General Data Protection Regulation?
2. What are the limitations of the current legal framework when considering the nature of genetic data as shared between individuals?

Background

GDPR came into force in 2018 with the aim to protect citizen's integrity. Provisions existed previously in extant European Data Protection Law under an EU Directive (95/46/EC). GDPR asserts that "the protection of natural persons in relation to the processing of personal data is a fundamental right" (GDPR, 2016). It includes definitions of what personal data is, and how this data may be processed. The main principle is that natural persons are to have control over their own personal data, and that sensitive data merit specific protection. GDPR applies both when the entity processing the personal data is established within the EU and when an entity outside the EU offers goods and services to people within the Union, even if the data are being stored or used outside of the EU. Consequently, GDPR has had a significant influence on privacy and data protection law worldwide (Phillips, 2019, Zaeem & Barber, 2020). GDPR applies in principle to every kind of operation and activity and regardless of who carries out the processing. Thus, while none of the companies in focus of the present paper are EU entities, they are still obliged to follow GDPR.

Technological development, having "transformed both the economy and social life", is stated as the main reason why GDPR was created (GDPR, 2016). DNA analysis combined with digital data processing is an example of such transformative technology. As King puts it, "DTCGT represents a new frontier in the collection of personal information" (King, 2019, p. 2).

Commercial DNA-testing has been used since the 1980s to assist with criminal investigations (Aronson, 2007) and determine familial relations (Cole, 2001). Today, tests are advertised for a variety of purposes, including genealogy. Like traditional records, the results need to be interpreted and evaluated regarding accuracy and significance (Donovan, Pasquetto & Pierre, 2018). Taking a DNA-test does not automatically reveal all one wishes to know about familial relations, and it can take a lot of time and effort to,

for example, find the genetic grandfather one is looking for (Classon, 2019). Nevertheless, the advent of DNA testing has undeniably opened up a new information resource for genealogists to explore. According to Stevens, "DNA sequencing machines are computerised instruments that produce DNA as bits and bytes. In this sense, a human genome sequence is already a new media object – it is born online, to be stored in databases and transmitted across the Internet" (Stevens, 2015, p. 394).

The DTCGT companies' value lie in providing a platform through which users can exchange information and utilise the logic of crowdsourcing when they organise and make accessible large volumes of data. Opting in to make yourself searchable and available for matching with others is a prerequisite for using the test result to its fullest extent. The privacy policies that users need to approve before using the services that DTCGT-companies provide should convey clear appreciation and understanding of all relevant facts, implications, and consequences. However, a recent case indicates that these requirements are not always met. In 2020, the Swedish Consumers Association (SCA), reported MyHeritage to the Swedish Authority for Privacy Protection. A review of MyHeritage's contract terms, concluded that the terms and conditions were unreasonable, violated the 3 § of the Swedish Contract Terms Act (SFS, 1994:1512), and did not follow GDPR (Swedish Consumers' Association, 2020). SCA said it was almost impossible for consumers to understand the terms and conditions, including the privacy policy; that the texts were too long and contained difficult words and concepts normally only used by lawyers and researchers within the field of medicine; that the terms and conditions as well as the privacy policy contained misleading wording and disclaimers; that the English and the Swedish versions did not correspond; and that the company was careless with personal data. MyHeritage also demanded that users waive the right to a trial in Sweden, insisting that consumers would not have the right to report the company anywhere but Israel, where the head office was located. According to MyHeritage's contract terms, it would therefore be impossible to bring a lawsuit against the company in Swedish court. In an email to SVT Nyheter [Swedish Television News] MyHeritage responded briefly to some of the criticisms levelled at them, announcing that they had never shared personal data with

third parties and that they were working to produce clearer agreements (Sveriges Television, 2020).

The example above points to a reason why DTCGT companies do not follow GDPR, namely that their business stem from and operate within other legislations that are not compatible with GDPR. Nevertheless, any person or entity operating within the EU should, according to GDPR, adapt their processing of personal data to the principles and rules established in the regulation.

Related research

Zaeem & Barber (2020) attempted to quantify how GDPR had improved privacy policies around the globe. Their findings indicated that though some progress had been made, more was necessary. Noncompliance often occurred "in the form of failing to explicitly indicate compliance, which in turn speaks to an organization's lack of transparency and disclosure regarding their processing and protection of personal information" (Zaeem & Barber, 2020, p. 1). Similar problems have been shown in studies that focus on other legal frameworks than GDPR. In a study focusing on the differing legal frameworks of England, Wales, Germany, and South Korea, Lee found that though the laws were different the problems of how genetic data could be used was very similar, suggesting that clearer guidelines are needed to balance the use of DNA technology with the laws (Lee, 2016). In a study focusing on federal statutes and regulations in the USA, Wright Clayton et al. criticised the existing juridical framework, arguing that "it has been difficult to develop broadly applicable legal principles for genetic privacy" and concluded that "few, if any, legal doctrines or enactments provide accurate protection or meaningful control to individuals over disclosures that may affect them" (Wright Clayton et al., 2019, p. 1). Phillips reviewed contracts of 71 DTCGT companies that provide tests for health purposes, concluding that it was questionable if the contracts were adequate to comply with UK law (Phillips, 2017 & 2019). Laestadius, Rich, & Auer (2017) assessed compliance with international transparency guidelines concluding that these were not consistently met in relation to transparency, privacy, and secondary use of data. Stevens examined the interdependence between biotechnology and new media in the development of

DTCGT companies, maintaining that the Web "allows genetic testing services to operate within an ambiguous regulatory framework" (Stevens, 2015, p. 395).

Noncompliance and legal ambiguity are however not the only problems - research on online contracts and wrap agreements show that users might not read them due to their length and to their rate of recurrence in the online environment (Obar & Oeldorf-Hirsch, 2020), and that even when reading privacy policies, there is "a significant level of misunderstanding on the part of consumers" on their meaning and effect (Phillips, 2015a, p. 72). In relation to DTCGT, Phillips argued that privacy policies need to be "(…) more comprehensive and address the issues of data sharing, sale, storage, and security in much greater depth and explicitly draw consumers' attention to companies' privacy practice" (Phillips, 2015b, p. 63), and suggested that contracts could be improved by making them more interactive, drawing attention to key clauses, and offering more options to opt out of services (Phillips, 2015a, p. 73). Surveying privacy policies of 90 DTCGT companies operating in the USA, Hazel & Slobogin observed "tremendous variability across the DTCGT industry in the quantity and quality of information provided to consumers concerning the collection, use, and sharing of their genetic data" (Hazel & Slobogin, 2018, p. 66), and though the majority made vague references to potentially applicable laws in their policy documents, they provided little or no information regarding their scope or potential applicability to DTCGT. Benoliel & Becher conducted linguistic readability tests to the five hundred most popular websites in the United States that use sign-in-wrap agreements, arguing that "Whereas consumers are expected and presumed to read their contracts, suppliers do not generally have a duty to draft readable contracts" (Benoliel & Becher, 2019, p. 2256). Yet, despite the shortcomings of privacy policies, an interview study conducted by King (2019) indicated that users are prone to lack risk awareness, trust DTCGT companies with their sensitive data, and assume that anonymity will protect them from potential privacy harm.

The use of DNA databases also creates several new ethical, legal, and integrity related challenges due to the possibility to reveal sensitive personal information about genetic relatives. Law enforcement agencies have been able to identify criminals by accessing databases managed by DTCGT companies. Often it is not a case of direct matches,

but of using genetic data from relatives to hone in on a suspect. Thus, though an agreement between two parties – the DTCGT company and the user – might seem unproblematic it potentially affects other persons as well, regardless of them knowing about it. This is because, as Ram puts it, "genetic information is not like other forms of private or personal information because it is shared— immutably and involuntarily— in ways that are identifying of both the source and that person's close genetic relatives" (Ram, 2015, p. 873). Extending this perspective to a societal level, King has argued that DTCGT "could be the final nail in the coffin of anonymity, as even those people who don't have a traceable digital presence find themselves drawn into a genetic human map through their relatives" (King, 2019, p. 27). It has been argued that DNA technology needs to be complemented and limited with a better and clearer legal framework and that cooperation is required between companies, law enforcement, bioethicists and others involved to create a responsible use of people's DNA (Kennett, 2019).

Theory

This section provides a brief overview of privacy as a theoretical concept and its relation to genetic data which is the type of data we focus on specifically in this paper.

Privacy is a normative concept, inscribed in article 12 of the Universal declaration of human rights (United Nations, 1948). The right to privacy refers to "the ethical and legal principles that recognize the importance of limited access to an individual or information about an individual" (Wright Clayton et al., 2019, p. 5). According to GDPR, limiting access is a challenge due to "rapid technological developments and globalisation" which have brought new obstacles for the protection of personal data, not least due to the increased scale of data collection and sharing (GDPR, 2016, p. 2). Disclosure or misuse of personal data can negatively affect individuals' relationships, reputation, employability, insurability, financial status, and more (Nissim & Wood, 2018).

Privacy was originally associated with secrecy or concealment, a "right to be let alone" (Warren & Brandelis, 1890, p. 193). The current understanding of privacy focuses more on privacy as a matter of control in relation to data privacy, "a personal right to control the use of one's data"

(Schwartz, 2000, p. 816). This "privacy control-paradigm", Schwartz argues, is a "liberal autonomy principle" that places the individual at the centre of decision-making through "individual stewardship of personal data", and "encourages a property approach to personal information that transforms data into a commodity" (ibid,. p. 820). Personal data can even be described as the "currency of the Internet" (Zaeem & Barber, 2020, p. 17). GDPR fits into this paradigm as it is an attempt to balance the interest of businesses with the interest of the individual. Though GDPR revolves around information privacy, the concept of privacy is not used in the regulation. The individual is at the centre of the regulation, and is referred to as "data subject", a term that occurs 406 times, or "natural person", a term that occurs 131 times. Exchanges and flows of personal data are in GDPR connected to the proper functioning of the internal market - creating trust is described as essential to "allow the digital economy to develop" (GDPR, 2016).

There are different types of personal data and some are considered sensitive, such as genetic data, "personal data relating to the inherited or acquired genetic characteristics of a natural person which result from the analysis of a biological sample from the natural person in question" (GDPR, 2016). Theoretically, a difference can be made between genetic material and the information rendered from analysing it, but these entities can also be seen as indistinguishable: De Witte and ten Have discussed the ownership of genetic material and information, maintaining that "On the level of scientific research there is hardly a distinction between genetic material and genetic information (genetic material is genetic information)" (De Witte & the Have, 1997, p. 58), and therefore: "With regard to genetic material and genetic information a reason to acknowledge ownership is not so much the possibility of selling genetic material and information, but the right to prevent others from having access to one's genetic material and information" (Witte & ten Have, 1997, p 57f). Wright Clayton et al. also highlighted that the genome is "a uniquely individual assemblage of widely shared common elements" that "imbues it with a dual private and public significance that confounds any discussion of policy addressing genetic privacy" (Wright Clayton et al., 2019, p. 2). This complicates GDPRs' notion of personal data as tied to and owned by one individual.

In the context of DTCGT the genome can be seen as an information resource or even a "book of life" (Kay, 2000) that people can explore to learn more about themselves and their familial relations. In their study about how DNA was used among attendees at DNA collection sessions in northern England, Scully, King & Brown (2013), conceptualised genetic code as "a form of archival memory" connected to narratives of identity, which in turn relate to macro narratives such as nationality, collective memory, and domesticity. According to Stevens, "DNA histories connect individual history (genealogy) to grand stories of human migration and exploration" and provide "powerful ways of forming group identities at a time when older narratives of identity (centred on nation, ethnicity, culture, and language) are being eroded" (Stevens, 2015, p. 391). DTCGT companies offer to give test takers insight into their genetic heritage on an aggregate level, and also on a more personal level tied to familial relations with other test-takers. However, once two test takers are connected, for example by revealing that they are half-siblings, the relationships to other family members, such as genetic parents, are also revealed - regardless of their consent or knowledge.

Material and method

Our material consists of privacy policies from three selected DTCGT companies. To use the services that the companies provide, users must agree to two types of policy documents where one is the privacy policy, and the other is the terms and conditions. The choice to focus primarily on the first type is that this is where the companies describe what data they collect, how they process personal data, and how they adhere to GDPR. The companies were selected on the basis of a list compiled by the Swedish Society for Genetic Genealogy (SSGG) of the top companies active in Sweden. This association's purpose was to support genealogists and accelerate the use of DNA for genealogical purposes in Sweden and we thus regarded SSGG as a solid basis for knowledge about DTCGT companies operating in the country. Further, the companies chosen are not only known in genealogical circles but are household names that continuously expand their business not only in Sweden but also in other countries in the European Union and other parts of the globe. MyHeritage, an Israeli company founded

in 2003, was in 2021 according to SSGG the company with the largest number of tested persons from Europe and among the more popular in Scandinavia. FamilyTreeDNA, founded in Texas in 2000 and in 2021 bought by the Australian company MyDNA, was according to SSGG was the company with the largest number of tested Swedes in 2021. Ancestry, founded in USA 1996, was in 2021 the largest company among the three. Their database had a preponderance of American history and genetic information which had proven especially useful for Swedes searching for emigrant ancestors (SSGG, 2020).

The versions of the privacy policies studied were from July 6, 2021 (MyHeritage), July 5, 2019 (FamilyTreeDNA), and August 3, 2021 (Ancestry). The amount of text ranged from about 5000 to 7000 words. The result was obtained through qualitative content analysis (Krippendorf, 2013), facilitating systematic reading and thematizing. The material was read repeatedly to acquire a good grasp of the whole. The units of analysis (Neuendorf 2017) selected were occasions where the privacy policies mentioned GDPR or explained how sensitive personal data was processed. Processing entails to collect, register, organise, preserve, change, read, transfer, spread, delete, or destroy data (IMY, 2021b). The occasions found were extracted, then sorted deductively according to the three principles of (i) adherence to GDPR, (ii) transparent processing of personal data, and (iii) consent. The material was also used to understand the problems of ensuring the privacy of genetic relatives of individuals using the services of DTCGT companies. To help interpret the privacy policies in relation to GDPR we used writings from the Swedish Authority for Privacy Protection ([Integritetsskyddsmyndigheten] IMY), designated by the Swedish government as the supervisory authority under GDPR (IMY, 2021a).

Result and analysis

The GDPR principles informing the examination of the privacy policies are presented under the respective sub section below. The core of GDPR is that personal data should only be processed when there are good reasons for it – it must be justified why an organisation holds a collection of personal data, and controllers should weigh their own interests against those of the data subjects.

ADHERENCE

It is not enough to state *that* GDPR is followed, it must be clear *how*. Controllers need to "demonstrate the compliance of processing activities with this Regulation" (GDPR, 2016, p. 14). This section is therefore guided by the question of how the companies show that they adhere to the principles of GDPR. The result shows that only one of the companies mention GDPR in their privacy policy.

Ancestry did not mention GDPR but said, "Under EU and UK law, we are required to specify the purposes for which we process your Personal Information and the legal bases which we rely on to do this". The legal bases were accounted for in a separate document, where "Legitimate interests" was stated as the legal base to personalise, understand, maintain, develop, and improve Ancestrys' services and to conduct scientific, statistical, and historical research on aggregated genetic information. What was meant by legitimate interests was not explained further than saying "Where we rely on legitimate interests to process your Personal Information, you have the right to object to such processing (meaning that you can ask us to stop). You can use your privacy settings to control certain ways in which we process your data". Other legal bases found in the document seemed more straightforward, such as explicit consent (to undertake scientific research on biological samples), consent-device based permissions (to collect geolocation information from users' devices), and compliance with legal obligations (to detect and defend against fraudulent, abusive, or unlawful activity).

FamilyTreeDNA did not mention GDPR either, instead they referred to Privacy Shield Framework, an agreement that they said allowed American companies to process personal data from Europe. However, in its judgement of 16 July 2020 (Case C-311/18), the Court of Justice of the European Union had ruled that the EU-US Privacy Shield was no longer a valid mechanism to transfer personal data from the European Union to the United States. The annulment of Privacy Shield meant that personal data controllers in the EU were no longer allowed to transfer personal data to recipients in the US based on Privacy Shield. The European Commission and the US Government had started negotiations on a successor arrangement to the EU-US Privacy Shield to comply with the judgement of the Court. A possible explanation as to why this agreement was still in the privacy policy was that the latest update when the

document was accessed was made in 2019. However, GDPR was already adopted by then. Thus, FamilyTreeDNA referred to an obsolete regulation.

MyHeritage said they had "taken steps to ensure compliance with all applicable privacy laws, including the general data protection regulation (GDPR)". However, the Terms and conditions (2020) still asserted that: "This Agreement and any dispute regarding the Service shall be exclusively governed by the laws of the State of Israel, without regard to conflict of law provisions, and you agree that any legal proceeding about the execution, performance and/or enforcement of this Agreement shall be brought exclusively to the courts located in Tel Aviv, Israel." In the privacy policy, under the heading "Legal grounds for the processing of personal information" MyHeritage aimed to explain their justification for processing personal data, which included when "the processing is in our legitimate commercial interests subject to your interests and fundamental rights". What was meant by legitimate commercial interests or what basic rights consumers was not explicitly stated.

Transparency

This section focuses on linguistic clarity. GDPR says "any information and communication relating to the processing of those personal data be easily accessible and easy to understand, and that clear and plain language be used" (GDPR, 2016, p. 9).

The number of words in the privacy policies ranged from 5000 to 7000. Counting other referenced documents (terms and conditions, cookie policy, extended details on privacy preference) added 8500 - 15 500 words. Users thus needed to read a significant amount of text to grasp how their data was being processed. All three companies stated that they might change their privacy policy at any time, requiring users to keep updated. A general problem was that the privacy policies did not offer a clear distinction between personal data and sensitive personal data. Another problem was that sweeping phrases – "for example", "among other things", "etcetera" – were used when explaining ways in which data was processed. All companies included disclaimers regarding data security, saying that while using reasonable efforts, they could not guarantee that loss,

misuse, or alteration of data would not occur. All companies also said they disclosed personal information to third parties, and used personal data to develop new products and services.

Ancestry defined Personal Information as “information that can identify you directly or indirectly”; DNA data as “machine-readable biometric data”; and genetic information as “DNA data and any information derived from it such as ethnicity estimates, communities, traits, and genetic relative matches”. A 7-point bullet list was provided of the primary purposes for which Ancestry used genetic information, which included “studying aggregated genetic information to better understand population and ethnicity-related health, wellness, ageing, or physical conditions”, and to “conduct scientific, statistical, and historical research”. Ancestry emphasised that they only shared genetic information with selected third parties, including Ancestry Companies, other users, and service providers which would “have access to some of your information, including some of your Personal Information, in their systems”. It was highlighted that biological samples were not considered personal information and that users could request to have their biological samples destroyed.

FamilyTreeDNA defined personal information as “information that can help identify you”, and genetic information as “readable DNA Data”, which is converted to from extracted DNA. “Sensitive information” was a term used but not defined. FamilyTreeDNA said they used genetic information for purposes which “may include, in addition to other things not specifically mentioned” to “perform statistical, scientific, and historical research”. Third party service providers, not specified, would “have some of your information in their systems”. Aggregate information was defined as “Information that has been combined with that of other users and analysed or assessed as a whole, such that no specific individual may be reasonably identified”. It was unclear what reasonably meant. The possibility of being identified was mentioned in relation to pseudonymous information, of which it was stated: “Information that has been stripped of your Account Information and other identifying data, such that you cannot easily be identified as an individual to the public and is instead only identifiable by a kit number or other alphanumeric sequences.” What was meant by easily identified remained unclear. GDPR does not apply to anonymous information. However, it states that

"Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly" (GDPR, 2016, p. 5).

MyHeritage did not define personal information other than exemplifying with "types of personal information", including "DNA information" of which it was said "DNA-related information is generated and stored when you use our DNA Services". A disclaimer stated "You acknowledge that you provide your personal information at your own risk". MyHeritage explained how they used personal data in a section which contained the most comprehensive information, but it was stated that the privacy policy should be read in conjunction with the Terms and Conditions. The privacy policy said, "We only collect information we believe is necessary for our legitimate business interests, to provide you with the Service". It was unclear what was meant by legitimate business interests. MyHeritage stated that they processed health related information not only of the data subject but also their family: "Before providing you with the DNA Health Reports, we collect certain self-reported family health history information from you, about you and your family members". This information was said to be required in the USA but collected from all users "for data consistency". The privacy policy did not require consent from family members before reporting their health history despite this being personal data that could be highly sensitive.

Consent

According to GDPR "Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her" (GDPR, 2016, p. 6). This can include ticking a box when visiting an internet website or choosing technical settings for information society services. If the controller intends to process the personal data for a purpose other than that for which they were collected, the controller "should provide the data subject prior to that further processing with

information on that other purpose and other necessary information" (GDPR, 2016, p. 12).

All three privacy policies declared that by creating an account, users were confirming that they understood that the company would collect, process, and share the users' personal data. Additional documents covered informed consent for research, of which it was stated that consent could be revoked at any time, unless the research had already occurred or was underway.

Ancestrys' Informed consent for research covered "research that may be performed by AncestryDNA, a third-party researcher or in collaboration between AncestryDNA and a third-party researcher" (Ancestry, 2018) under the umbrella of Ancestry Human Diversity Project, which covered a broad spectrum of research related to the study of human genetics, genealogy, anthropology, and health. The research agreed to could be performed by employees of AncestryDNA, as well as "researchers from other organisations and companies who share our Purpose. This might include academic institutions as well as non-profit and for-profit businesses or government agencies". This statement is problematic since consent should be given only for specific purposes. Potential research participants are to be given information about the key elements of a research study and what their participation will involve. Ancestry's informed consent seemed to be much broader. The consent covered the use of "all data that you provide to us when you use our Services, including Biological Samples and any data derived from those samples". Self-reported health and trait data may include information about family members that had not given their consent since it consisted of "lifestyle or other traits of you and your family members".

FamilyTreeDNAs' privacy policy said DNA samples were stored "so that they can be attainable for future testing", if the data subject had agreed to it. FamilyTreeDNA said one of the primary purposes for which users' genetic information was processed was to "perform statistical, scientific, and historical research", but that the company would not share users' personal information with third parties without additional consent "before we collect and process Sensitive Information as part of your interaction with the Services". However, it was unclear what was meant by Sensitive Information and how this related to what was defined as genetic information or DNA data. FamilyTreeDNA described their consent process for

research as "Users wishing to participate in genetic research projects may indicate so and will be included in a candidate pool. If you become a potential candidate for a research project, you will be contacted to grant specific consent for every individual research project opportunity".

MyHeritages privacy policy said, "special category data or sensitive personal data is processed on your explicit consent", and that "If you voluntarily agreed to the DNA Informed consent we may use your information (such as DNA Results and other DNA Information) for the purposes of research as specified therein". Their document DNA Informed Consent — MyHeritage DNA Research Project (2021) explained that The MyHeritage DNA Research Project was "a program designed by MyHeritage to better understand human genetics, genealogy, and anthropology". Consent gave MyHeritage the right to use personal data in a broad scope: "research studies designed to further our understanding of genealogy, anthropology, cultures, human evolution and migration, human genetics, population genetics, epidemiology, population health issues, and regional health issues".

GDPR acknowledges that it is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection, why "data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research" (GDPR, 2016, p. 6). Only FamilyTreeDNA said they would ask for additional consent in this way.

Privacy of genetic relatives

"Data subject" is a concept used frequently in GDPR, mentioned 88 times, but not defined. We understand it as referring to the person who the personal data concerns. GDPR defines personal data as "any information relating to an identified or identifiable natural person" which in the context of DTCGT is the person who has taken a DNA test, but which may also, due to the nature of DNA as partially shared between individuals, extend to genetic relatives. In relation to DTCGT, the individual taking a DNA test can be seen both as a data subject, a consumer and as a part of the company's assets since the business is based on people sharing their data with others. Sharing genetic data is however not limited to being a privacy risk for the person

who has taken the test but also for genetic relatives. GDPR does not address this issue but is founded on the idea that my personal is mine alone to do with as I please. One of the privacy policies stated that the company "collect certain self-reported family health history information from you, about you and your family members" which from what we can understand would mean that the person reporting the health history of others thereby would become a data processor "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller". This would require consent from concerned family members whose health data was being reported. Wallace et al. has suggested a "generational consent" model and recommended that "companies should make it clearer that clients should inform third parties about their plans to participate, that third parties' data will be provided to companies, and that that data will be linked to other databases, thus raising privacy and issues on use of data" (Wallace et al, 2015, p. 1). However, which third parties that are concerned is not always clear from the start since there are differences between social family and genetic family, thus consent from a social relative might still breach the privacy of a genetic relative.

An immediate consequence for genetic relatives is that family relations are revealed in an unexpected and unwanted way. Theunissen found that "notions of family were frequently challenged with unexpected DNA test results causing shifts in personal and social identities, especially in their family and biological identities" (Theunissen, 2022, p. 1).

Discussion and concluding remarks

The first research question was: How do the three companies' privacy policies comply with the General Data Protection Regulation? The result indicated three types of shortcomings in the companies' approach to the regulation:

- The privacy policies seldom mentioned GDPR. When mentioned, it was done briefly and in one case we found references to invalid international agreements. This conflicts with GDPR's requirement that "the controller shall implement appropriate technical and organisational measures to ensure and to be able to

demonstrate that processing is performed in accordance with this Regulation" (GDPR, 2016, Chapter IV Article 24).

- The privacy policies were often vague and even potentially misleading in their wording. The consequence was that users could not fully understand what personal data the companies collected or what would happen to the data. This conflicts with GDPR's requirement that information about data processing should be provided "in a concise, transparent, intelligible and easily accessible form, using clear and plain language" (GDPR, 2016, Chapter III Article 12).
- All companies said they may change or update the terms without notice. This might undermine what the data subject had first agreed to and thus conflicts with GDPR's requirement that personal data should only be collected "for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes" (GDPR, 2016, Chapter II Article 5).

The second research question was: What are the limitations of the current legal framework when considering the nature of genetic data as shared between individuals? Given the shortcomings we were able to identify, one important point becomes clear:

- The privacy policies studied all make a distinction between genetic material and genetic data, as does GDPR: "Genetic data should be defined as personal data relating to the inherited or acquired genetic characteristics of a natural person which result from the analysis of a biological sample from the natural person in question" (GDPR, 2016, Article 4). This becomes problematic due to the nature of genetic data as partially shared, which has the consequence that sensitive data such as maternity or paternity can be revealed without consent of all concerned parties. As De Witte & ten Have pointed out it is even problematic to talk of genes as parts of an individual person's body since "is not specific for an individual but refers to a pedigree; similar material is shared with relatives and genes as physical entities are in fact common to all people" (De Witte & the Have, 1997, p. 58).

In sum, we found that there was not always a clear difference between what GDPR says about how personal data should be

processed and how the companies said they processed personal data. The problems appeared between the lines – in unclear formulations and inaccessible language, in translated versions of the privacy policies that were not legally binding, in seemingly trivial settings that could have far-reaching consequences, and in conditions that could be changed at any time. Thus, it was a challenge to specify the risks that the privacy policies had in relation to privacy and applicable law. However, we could estimate a few problematic features.

The first relates to the relationship between law and technical development. The services DTCGT companies offer are historically unprecedented. Genealogy and family history are not new phenomena, but the possibility to turn to genetics to the extent done today is based on methods developed in a relatively short time where it is unclear if juridical frameworks have kept up. This is not to say that genealogy based on traditional records or oral narrations is unproblematic. However, contemporary genealogy done by analysing and sharing genetic information increases the privacy risks, both for the individual tested and for their genetic relatives. In one case the company asked users to provide family health history information about not only the user but also about the user's family members, which adds to the privacy risks of people who are not customers.

We also found a lack of a coherent account of what sensitive personal data that the companies processed. Often there were hints as to what type of personal data the privacy policies referred to, but these hints were scattered across the text under different headings. An adjacent problem was that in the two cases where the privacy policies were translated to Swedish, the companies stated that the translation was just for convenience, and that only the English version is legally binding. Not all users have the linguistic skills to read English. The companies place much responsibility on the user. This is also relevant in relation to consent. To give consent, the user must understand the privacy policy. Small and seemingly insignificant setting options at the web pages of companies can lead to large consequences for the user. For example, how personal data can be received or shared via the user's profile to others. These settings also hold the potential to have consequences for genetic relatives that have not given their consent to having their data processed. The offer that services can be customised might be problematic since it is difficult for the individual to comprehend the consequences of all settings made. It is hard to predict solely based on the privacy policies how the data will be processed in the future and which consequences that may follow. Our results confirm King's conclusion that "the collective identification presented by DTCGT takes the challenge of managing collective privacy to a new level while also raising critical questions of consent" (King 2019, p. 26). More knowledge is needed about how to handle privacy of genetic relatives in cases where individuals provide genetic data to DTCGT companies. Future research could

investigate what specific privacy consequences DNA testing has had for genetic relatives of customers using the services of DTCGT companies, and how the legal framework could be adapted to respond to the development of commercial services based on new advances in genomic and digital technology.

References

Ancestry. *Your privacy*. Effective Date: 3 Aug 2021.Accessed 2021-09-20 from https://www.ancestry.com/cs/legal/privacystatement

Ancestry. *AncestryDNA Informed Consent*. Effective Date: July 24, 2018. Accessed 2022-01-10 from https://www.ancestry.com/dna/lp/informedconsent-v4-en

Aronson, Jay D. (2007) *Genetic Witness: Science, Law, and Controversy in the Making of DNA Profiling*, Rutgers University Press.

Benoliel, Uri, and Shmuel I. Becher (2019). "The Duty to Read the Unreadable." *Boston College Law Review* 60, no. 8 (November 1, 2019): 2255–96.

Clayton, Ellen Wright, Barbara J. Evans, James W. Hazel & Mark A. Rothstein (2019). "The Law of Genetic Privacy: Applications, Implications, and Limitations." *Journal of Law and the Biosciences* 6 (1): 1–36.

Classon, Roland (2019). "Hon topsade sig - och fick en ny farfar" ["She swabbed her cheek - and got a new grandfather"]. *Sydsvenskan*. 19 december 2019. Accessed 2023-11-07 from https://www.sydsvenskan.se/2019-12-19/hon-topsade-sig--och-fick-en-ny-farfar

Cole, Simon A. (2001) *Suspect identities : A history of fingerprinting and criminal identification*. Harvard university press.

Court of Justice of the European Union (2020). *Judgement in Case C-311/18 - Data Protection Commissioner v Facebook Ireland Ltd and Maximillian Schrems*.

De Witte, Joke I. and Henk Ten Have (1997). "Ownership of Genetic Material and Information." *Social Science and Medicine* 45, no. 1: 51–60. doi:10.1016/S0277-9536(96)00309-7.

Donovan, Joan, Irene Pasquetto & Jennifer Pierre (2018). "Cracking Open The Black Box of Genetic Ancestry Testing Technologies". *Proceedings of the 36th Annual Hawaii International Conference on System Science (HICSS)*.

FamilyTreeDNA. *Privacy statement*. Last updated 05/07/2019. Accessed 2021-09-20 from https://www.familytreedna.com/legal/privacy-statement

FamilyTreeDNA. *Group Project Participation Informed Consent*. Effective as of June 5, (2018). Accessed 2022-01-10 from https://www.familytreedna.com/legal/terms/group-project-participation

Future of Privacy Forum. "Privacy best practices for consumer genetic testing services". (2018). [accessed 2022-03-29 from https://fpf.org/blog/privacy-best-practices-for-consumer-genetic-testing-services/]

Hazel, James W., and Christopher Slobogin (2018). "Who Knows What, and When: A Survey of the Privacy Policies Proffered by U.S. Direct-to-Consumer Genetic Testing Companies." *Cornell Journal of Law and Public Policy* 28, no. 1 (January 1, 2018): 35–66.

Howard, Heidi C, Bartha Maria Knoppers, and Pascal Borry (2010). "Blurring Lines. The Research Activities of Direct-to-Consumer Genetic Testing Companies Raise Questions about Consumers as Research Subjects." *EMBO Reports* 11, no. 8 (August 2010): 579–82. doi:10.1038/embor.2010.105.

Integritetsskyddsmyndigheten [Swedish Authority for Privacy Protection] *About us*. (2021a). Accessed 2021-02-20 from https://www.imy.se/en/about-us/swedish-authority-for-privacy-protections-assignment/

Integritetsskyddsmyndigheten [Swedish Authority for Privacy Protection]. *Detta är känsliga personuppgifter. [This is sensitive personal data]* (2021b). Accessed 2021-02-20 from https://www.imy.se/lagar--regler/dataskyddsforordningen/kansliga-personuppgifter/detta-ar-kansliga-personuppgifter/

Integritetsskyddsmyndigheten [Swedish Authority for Privacy Protection]. *Sanktionsavgifter och varningar. [Sanction fees and warnings]* (2021c). Accessed 2021-03-17 from https://www.imy.se/om-oss/vart-uppdrag/sa-arbetar-vi-med-tillsyn/vad-kan-tillsynen-leda-till/

Kay, Lily E. (2000) *Who Wrote the Book of Life? : A History of the Genetic Code*. Writing Science. Stanford University Press.

Kennett, Debbie (2019). "Using Genetic Genealogy Databases in Missing Persons Cases and to Develop Suspect Leads in Violent Crimes." *Forensic Science International* 301 (August 1, 2019): 107–17. doi:10.1016/j.forsciint.2019.05.016.

King, Jennifer (2019). "Becoming Part of Something Bigger : Direct to Consumer Genetic Testing, Privacy, and Personal Disclosure." *Proceedings of the ACM on Human-Computer Interaction*. 2019;3(CSCW):1-33. doi:10.1145/3359260

Krippendorff, Klaus (2013). *Content analysis : An introduction to its methodology*. 3rd ed. SAGE.

Laestadius, Linnea I., Jennifer R. Rich, and Paul L. Auer (2016). "All Your Data (Effectively) Belong to Us: Data Practices among Direct-to-Consumer Genetic Testing Firms." *Genetics in Medicine* 19, no. 5: 513–20. doi:10.1038/gim.2016.136.

Lee, Jungnyum (2016). "The Presence and Future of the Use of DNA-Information and the Protection of Genetic Informational Privacy: A Comparative Perspective." *International Journal of Law, Crime and Justice* 44 (March 1, 2016): 212–29. doi:10.1016/j.ijlcj.2015.10.001.

MyHeritage. *Terms and conditions*. Last updated on April 18, 2020. Accessed 2022-01-10 from https://www.myheritage.se/terms-and-conditions

MyHeritage. *MyHeritage Privacy Policy*. Last updated July 6 2021. Accessed 2021-09-20 from https://www.myheritage.se/FP/Company/popup-privacy-policy.php?dispLang=EN

MyHeritage. *DNA Informed Consent — MyHeritage DNA Research Project*. Updated on October 28, 2021. Accessed 2022-01-10 from https://www.myheritage.mk/dna-informed-consent-agreement

Neuendorf, Kimberly A. (2017). *The Content Analysis Guidebook*. 2. ed. SAGE.

Nissim, Kobbi & Wood, Alexandra (2018). "Is privacy privacy?" *Philosophical Transactions of the Royal Society A*. vol 376, issue 2128. https://doi.org/10.1098/rsta.2017.0358

Obar, Jonathan A., and Anne Oeldorf-Hirsch (2020). "The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services." *Information, Communication & Society* 23, no. 1 (January 1, 2020): 128–47. doi:10.1080/1369118X.2018.1486870.

Official Journal of the European Union (1995). *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995*

on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995

Official Journal of the European Union (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016*

Phillips, Andelka M. (2015). "Think before You Click: Ordering a Genetic Test Online." *SciTech Lawyer* 11, no. 2 (January 1, 2015): 8–12.

Phillips, Andelka M. (2015). "Genomic Privacy and Direct-to-Consumer Genetics: Big Consumer Genetic Data - What's in That Contract?" *Proceedings - 2015 IEEE Security and Privacy Workshops, SPW* 2015, 60–64. doi:10.1109/SPW.2015.19.

Phillips, Andelka M. (2019). *Buying Your Self on the Internet : Wrap Contracts and Personal Genomics*. Future Law. Edinburgh University Press, 2019.

Ram, Natalie (2015). "DNA by the Entirety." *Columbia Law Review* 115, no. 4 (May 1, 2015): 873–940.

Schwartz, Paul M. (2000). "Internet Privacy and the State." *Connecticut Law Review* 32, no. 3 (March 15, 2000): 815–60.

Stallard, Matthew and Jerome De Groot (2020). "'Things Are Coming Out That Are Questionable, We Never Knew About': DNA and the New Family History." *Journal of Family History* 45, no. 3: 274–94. doi:10.1177/0363199020906853.

Scully, Marc, Turi King and Steven D. Brown (2013). "Remediating Viking Origins: Genetic Code as Archival Memory of the Remote Past." *Sociology* 47, no. 5 (October 1, 2013): 921–38.

Stevens, Hallam (2015). "Genetimes and Lifetimes: DNA, New Media, and History." *Memory Studies* 8, no. 4: 390–406. doi:10.1177/1750698015581074.

Svenska Sällskapet för Genetisk Genealogi [Swedish Society for Genetic Genealogy], (2020). *Vilket test ska jag välja? [Which test should I chose?]* Accessed 2021-09-08 from http://ssgg.se/vilket-test-ska-jag-valja/.

Sveriges Konsumenter [Swedish Consumers Association] (2020a). *Anmälan om oskäliga avtalsvillkor gällande DNA-tester från MyHeritage Ltd [Report on unreasonable terms and conditions relating to DNA tests from MyHeritage]* [2021-09-08].

Sveriges Konsumenter [Swedish Consumers Association] (2020b). *Vi anmäler MyHeritage för brott mot GDPR [We report MyHeritage for breach against GDPR]*
[press release], 13 March 2020.

Sveriges Television [Swedish television] (2013). ”Dna-testning anmäls för GDPR-brott” [DNA-testing reported for GDPR violation]. News article, *SVT Nyheter [Swedish Television News]*, April 13, 2020.

Swedish Code of Statutes. *Lag (1994:1512) om avtalsvillkor i konsumentförhållanden [Act (1994: 1512) on contract terms in consumer relations]*

Theunissen, Catherine Agnes (2022). “The Effects of DNA Test Results on Biological and Family Identities”. *Genealogy*. 2022;6(17):17. doi:10.3390/genealogy6010017

United Nations (1948). *Universal declaration of human rights*.

Wallace, Susan E., Elli G. Gourna, Viktoriya Nikolova, and Nuala A. Sheehan (2015). “Family Tree and Ancestry Inference: Is There a Need for a ‘Generational’ Consent?” *BMC Medical Ethics* 16 (December 9, 2015): 1–9. doi:10.1186/s12910-015-0080-2.

Warren, Samuel D., and Louis D. Brandelis (1890). “The Right to Privacy.” *Harvard Law Review* 4, no. 5 (December 15, 1890): 193–220. doi:10.2307/1321160.

Zaeem, Razieh Nokhbeh & Baber, Suzanne K. (2020). “The effect of the GDPR on privacy policies: Recent progress and future promise”. *ACM Transactions on Management Information Systems*, Vol. 12, No. 1. https://doi.org/10.1145/3389685

Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

HAOYUE ZHAO[9]

Challenges or Opportunities: When Artificial Intelligence is Applied to Digital Humanities

Abstract: The aim of this review is to provide a dialectical reflection on the phenomenon of using AI tools to digital humanities, especially in academic writing, and a series of scenarios for the way of applying AI to digital humanities in future. To write this review, I give direct instructions to ChatGPT to generate a 2000-word essay, which defines the digital humanities and describes the field. Then, in the review, I analyze the argument, argumentation methodology and data of the essay. Through these analyses, I conclude that AI is not a replacement for humans in future research and development of digital humanities, but the emergence of AI does make a difference in the way traditional academic writing is done. By a series of comparisons, I argue that AI tools provide new opportunities for digital humanities and suggest some possible ways for AI engagement in digital humanities scholarship.

Keywords: AI tools, Digital humanities, Opportunities

[9] Haoyue Zhao currently attends the master's program in Digital Humanities at the department of ALM, at Uppsala University and holds a bachelor's degree in Chinese Language and Literature.
Email: haoyue.zhao.5647@student.uu.se

What AI does right and what it does wrong

In recent years, with the breakthrough development of AI tools, the ways in which AI will be applied to digital humanities is gradually becoming an inevitable topic in this discipline. Yet while there are critical flaws in the performance of artificial intelligence in humanities scholarship, it is clear that the general trend of its application to digital humanities research is irreversible. This has led to concerns, with some considering this trend as a challenge to humanities.

I give direct instructions to ChatGPT 4 to generate a 2000-word essay with headline, introduction, main body, conclusion, in-text citations and references, which defines the digital humanities and describes the field.

When I read the essay by ChatGPT, I found it did somethings right. First, it generally followed my directives, and the general layout of the paper is properly executed, which has title, main body, conclusion and references. Second, the statement of the essay is fluent, and it introduces digital humanities from aspects of definition and scope, historical overview and impact on scholarship, which makes readers understand digital humanities comprehensively.

However, although it has some good parts, it still has some crucial flaws which makes it impossible to be called a scientific text strictly. There are three main reasons for it. In the first place, there are some problems with the references of this article. The bibliography does not follow the standard citation format, such as MLA, APA, Harvard, etc. The format is chaotic, the information of the bibliography is not specific, for example, some of them have no number and date for issues, and all of them are not accurate regarding the number of pages. Also, the order of the writer, title, and publisher are completely corrupted. For another, the citations that ChatGPT use are not always correct. In this article, ChatGPT cites the work of Terras Melissa *What is Digital Humanities*?, but I found that she has not written this book, and although the third reference exists, the content it cites does not belong to the cited book. Moreover, the content of the essay lacks creativity, because more than half of this paper is cited from other literature. For example, the part of introduction and the first part of the main body are all cited from Wikipedia, which cannot happen in a manually written essay, because

the introduction of a human-written essay will contain mostly original content. In addition, although ChatGPT creates some content in other parts of this essay, it seems to simply summarize what it quotes, instead of giving its own opinions or evaluations, which are the cores of a scientific text. But despite this, there is an exception. In the part of Impact on Scholarship, ChatGPT evaluates the statement "It is not about using computers to do Humanities research for you" as "Digital Humanities is not just about using technology to make research easier but rather about using technology to ask new questions that were previously impossible to answer", which is constructive to some extent. But this opinion still has its prototype, which is the argument from *The Digital Humanities or a Digital Humanism*, by Parry Dave (2012, p. 433–434). Finally, each argument of the essay is not explained in a very complete and comprehensive manner. The number of words of this essay is well short of 2,000 words, with only 477 words, and its argumentation is very one-dimensional, only using citation arguments. As a result, each argument is merely mentioned, not argued for in the truest sense. For example, in the part of Historical Overview, it explains to the reader that the digital humanities were first called humanities computing, and then expanded further in scope to be called digital humanities, but it does not explain to the reader the detailed history of this change and the reasons for it. There are also no actual projects cited to demonstrate that there is indeed a gap in the definition of this discipline between the two ways of naming it.

Differences in the definition of Digital Humanities

In the essay written by ChatGPT 4, it defines the discipline of digital humanities mostly from Wikipedia, which said that Digital Humanities is a discipline at the intersection of computer science and humanities, and that it uses various kinds of digital methods to solve problems in humanities. It also invokes the metaphor of the "big tent" to illustrate the breadth of the Digital Humanities. From above, it could be summarized that GPT defines and describes digital humanities from three aspects respectively: the nature of the discipline, the way and purpose of research, and the scope of

the definition. As a result, I will also gradually elaborate my understanding of digital humanities as a discipline from these three aspects, which includes identification with the arguments of GPT, but more refinement and critical thinking.

First, it would be sensible that digital humanities is an interdisciplinary discipline in terms of the nature of the discipline. Jan Luhmann and Manuel Burghardt (2022, p. 148) argue that "DH is simultaneously a discipline in its own right and a highly interdisciplinary field", and Lev Manovich (2017, p. 56) defines that "Digital Humanities scholars use computers to analyse mostly historical artefacts created by professionals". As an independent discipline, digital humanities has clear disciplinary boundaries, however, we cannot simply define its disciplinary nature simply between the traditional humanities and computer science. According to the data of Jan Luhmann and Manuel Burghardt (2022, p. 161–162), the disciplines with the strongest relevance to digital humanities are computational linguistics and information science, while traditional humanities disciplines like art history, linguistics, literary studies and traditional computing disciplines like theoretical computer science, applied computer science, are somewhat different in terms of their research themes and research methods. Additionally, digital humanities is where literary studies, applied computer science, and statistics exist in transition zone, but cannot be called a crossover.

The argument of GPT that the digital humanities uses the method of social science and the computer as a tool to conduct and publish research might be reasonable. Just as N. Katherine Hayles (2012, p. 45) argues that "posit the digital humanities as a diverse field of practices associated with computational techniques and reaching beyond print in its modes of enquiry, research, publication, and dissemination". However, the aim of the digital humanities is not only about using digital tools to help doing research but rather about using technology to answer new questions that were previously impossible to answer. Because digital humanities uses both quantitative and qualitive methods, which combines strengths in humanities research and natural science research. For example, Drucker (2011, p. 2) defines the data that digital humanities uses as "capta", which is "taken" actively while data is assumed to be a "given", and it could answer the question of how humanists will confront the conceptual tools with humanistic principles while

keeping up with the times, along with the prevalence of quantitative research and the challenge to the authority of humanistic knowledge and the principles of critical thinking.

Finally, using the metaphor “meeting place” and “trading place” to replace “big tent” might be more appropriate. This disciplinary metaphor is suggested by Patrik Svensson (2012, p. *55*), which further expands the scope of digital humanities, and it would presume profound openness to several different epistemic traditions, and a facilitating role that is richly diverse and dynamic rather than purely instrumental or service minded.

Key factors AI do not have in research

Through analyzing the essay Digital Humanities: A Comprehensive Overview by ChatGPT 4, I find that there are two crucial instances where the writing of AI tools differ from writing performed by humans, in research and academic writing, and the instances further discussed in the first part of this review. In these instances it is clear that AI tools cannot completely take place of humans doing research in digital humanities.

Firstly, AI tools lack creativity, which could be attributed to the misuse of objectivity in the AI writing process. In the essay, more than half of it is cited from other literature, and its argumentation is very one-dimensional, only using citation arguments. It needs to be further explained that this objectivity of AI refers to quoting existing texts without critical thinking in a completely dispassionate manner. And Daston and Galison (2010, p. 17) suppose that this kind of objectivity is actually “blind sight”, they said “objectivity preserves the artifact or variation that would have been erased in the name of truth; it scruples to filter out the noise that undermines certainty…objectivity is blind sight, seeing without inference, interpretation, or intelligence”.

Moreover, AI tools lack academic ethics, which is mainly reflected in the data and references sections of this paper. If the misquoting of references and fabrication of data of AI is a low-level error that can be fixed, its direct citation of objective mass data without filtering will cause even more trouble for digital humanities research. Catherine D'Ignazio and Lauren F. Klein (2020, p. 149–172) call this kind of data “Big Dick Data”, and they suppose that data for digital humanities research needs ethical discipline and context.

The example of Clery Act data used in their book calls attention of digital humanists to the necessity of artificial interference with data, otherwise, the mass data would "run the risk not only of being arrogantly grandiose and empirically wrong, but also of doing real harm in their reinforcement of an unjust status quo".

Opportunities for digital humanities to use AI tools

Because of these problems and more of using AI tools, but the seemingly unstoppable trend of their application in humanities, some see this trend as challenges which could be summed up as "black-box" problem (Rockwell, 2016, p. 15). This mainly shows in two ways: "the first is the disappearance of the author… Ready-at-hand tools become transparent (unnoticed), and the creator's authorial responsibility for the instrument remains hidden" (Rockwell, 2016, p. 20). The second danger is that entanglement may cause the corruption of humanistic scholarship (Rockwell, 2016, p. 20–21). However, it seems wiser to regard it as opportunities, and there are some fields of digital humanities where AI tools could be applied and forge a broader future for digital humanities. Especially it can take part in the establishment of digital humanities laboratory, which is a new type of laboratory that combines the humanities with digital resources, services and tools (Pawlicka-Deger, 2020, p. 10). This is a new type of human-computer collaboration to some extent, which is very consistent with the discipline of digital humanities as "meeting place" and "trading place" (Svensson, 2012). Moreover, this collaboration between humans and AI can to some extent alleviate the inequality of collaboration between humanities scholars and science scholars, and reconcile the contradictions that arise between the two in terms of research culture, by keeping the two out of direct contact and using AI as an intermediary (Snow, 2012).

It could be possible for digital humanists to use AI to assist themselves to complete the experimental data output. As a matter of fact, most humanists are not as well versed in programming and data manipulation as professionals are, but AI can do most of the complex programming work in

digital humanities research projects on command or help humanists in checking the code they write themselves.

Secondly, AI can quickly retrieve, organize and summarize the experimental materials for digital humanists, including professional papers, experimental data, graphs, etc. It can even summarize this material and generate outlines to provide research direction and inspiration for digital humanists. This approach greatly improves the efficiency of humanities academic research and can quickly sift and organize the research materials needed by humanists, largely solving the problem of studying a large amount of literature in the early stages of their research.

Last but not least, AI can help digital humanists to learn and check their research materials and experimental results, like translate or explain terminology more accurately, and check grammar in academic writing. Since digital humanities research approaches are comprehensive and its research topics are diverse, digital humanists often take on a greater workload than traditional scholars when conducting research for their projects. For example, digital humanists are often exposed to a lot of professional terms outside of their own specialty but relevant to their research projects, and AI tools can quickly demystify these abstract concepts and collect the literature on them. Digital humanists usually also need to understand, learn and master some kind of data model in a short period of time, which can be difficult for non-computing academics, and AI tools can provide digital humanities scholars with some of the necessary learning tutorials and demonstrate algorithmic processes. Moreover, the use of AI as a tool for checking the grammar or diction of scientific texts is also becoming a common practice within digital humanities laboratories.

Conclusion

Although ChatGPT is able to write a generally complete essay following the instructions, there were still many shortcomings and the essay on digital humanities that ChatGPT produced could not be considered a real scientific text.

I reflect on the proposed definition of ChatGPT of digital humanities and amend it from the aspects of the nature of the discipline, the way and purpose of research, and the scope of the definition. Digital humanities is an

interdisciplinary discipline and has its clear disciplinary boundaries, but emerging disciplines like computational linguistics and information science are more relevant to it. Digital humanists use various kinds of digital tools and combine qualitative and quantitative methods, and they aim to find new research methods for the humanities. There could be two decisive reasons that the authority of humans is irreplaceable in digital humanities: one is creativity and the other is academic ethics. The existence of the former proves that the attitude of research is not only blindly objective, but also requires the inclusion of subjective and judgmental elements. The loss of the later might cause academic plagiarism. It can even lead to unethical data and findings, which has negative influence on society and academics. Although using AI tools completely to do digital humanities research has risks, application of AI tools in digital humanities should not be seen as a challenge. In fact, it is an opportunity, especially in the fields of establishing digital humanities laboratory, including providing technology assistance of experimental data output to digital humanists, organizing experimental material and working as a tool to learn or check.

In conclusion, it is inevitable that the digital humanities will apply AI tools to research, as dictated by its open, inclusive and diverse disciplinary nature and scope of research. The future is full of uncertainty, but humans must dominate the field of research in the digital humanities and give humanistic meaning to digital projects. In this context, AI tools are providing opportunities for the future development of the discipline.

References

Daston, Lorraine (2010). Objectivity. New York: Zone Books.

D'Ignazio, Catherine and Klein, Lauren F. (2020). Data Feminism. The MIT Press.

Drucker, Johanna (2011). Humanities Approaches to Graphical Display.

Hayles, N. Katherine (2012). How We Think: Transforming Power and Digital Technologies. In: Berry, David M. (ed.) Understanding Digital Humanities. London: Palgrave Macmillan UK, p. 42–66.

Luhmann, Jan and Burghardt, Manuel (2022). Digital humanities—A discipline in its own right? An analysis of the role and position of digital humanities in the academic landscape. Journal of the Association for Information Science and Technology, 73(2), p. 148–171.

Manovich, Lev (2017). 3. Cultural Analytics, Social Computing and Digital Humanities. In: Schäfer, Mirko Tobias and Van Es, Karin (eds.) The Datafied Society. Amsterdam University Press, p. 55–68.

Parry, Dave (2012). The Digital Humanities or a Digital Humanism. In: Gold, Matthew K. (ed.) Debates in the Digital Humanities. University of Minnesota Press, p. 429-437.

Pawlicka-Deger, Urszula (2020). The Laboratory Turn: Exploring Discourses, Landscapes, and Models of Humanities Labs. Digital humanities quarterly, 14(3).

Rockwell, Geoffrey (2016). Hermeneutica: computer-assisted interpretation in the humanities. Cambridge, MA: The MIT Press.

Snow, C. P. (2012). The two cultures. Cambridge: University Press.

Svensson, Patrik (2012). The digital humanities as a humanities project. Arts and Humanities in Higher Education, 11(1–2), p. 42–60.

Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

INA-MARIA JANSSON[10]

A Past for Everyone? Digital Participation as Method for Democratic Cultural Heritage Production

This article is a summary of the doctoral dissertation *A Past for Everyone? Digital Participation as Method for Democratic Cultural Heritage Production*, published at the Department of ALM and defended at Uppsala University in June 2023. The comprehensive summary of the dissertation is in Swedish and is publicly available on DiVA (http://www.diva-portal.org).

Keywords: digital participation, inclusive heritage, democratisation, heritagisation, knowledge organisation, information infrastructures

10 Ina-Maria Jansson is an archivist at Uppsala University Library, Special Collections Division. She works with personal papers and digital born collections. Her research interests are cultural heritage formation processes, cultural memory, information structures and cataloguing practices at cultural heritage institutions.
Email: ina-maria.jansson@ub.uu.se.

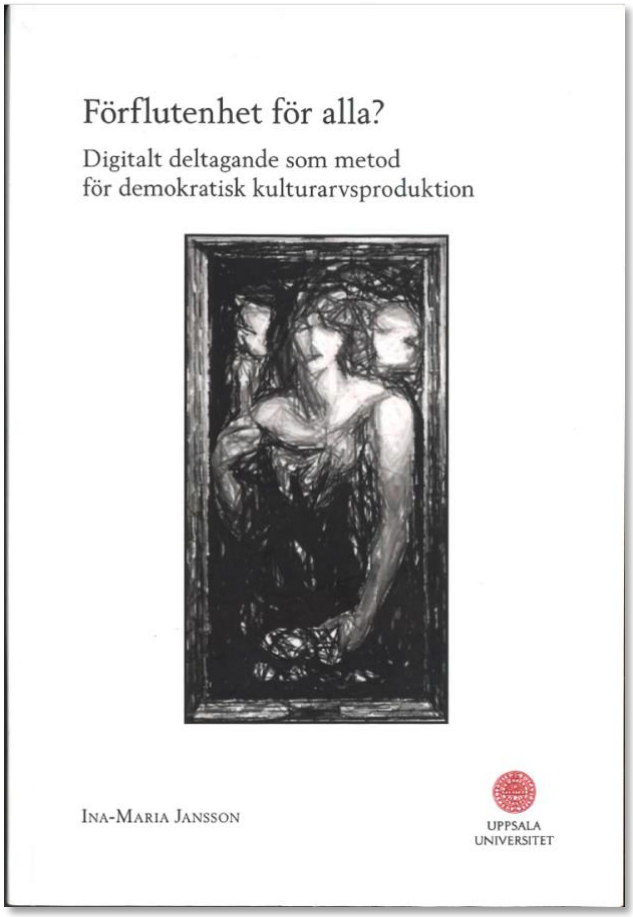

Image 1. Front cover of the printed dissertation.

Introduction

This thesis inquires into the conditions for how digital participation can be a method to produce cultural heritage in relation to collections at cultural heritage institutions such as archives, libraries and museums.

Ever since the digitisation and online publication of heritage collections items and catalogues became possible, heritage experts have debated whether increased accessibility and possibilities for digital participation also would lead to a more democratic and inclusive cultural heritage. The idea that institutional heritage is in need of being democratised is based on the notion of how these institutions represent an authoritative type of heritage that is defined hegemonically by a social and intellectual elite, excluding the voices of many other groups of people in the society.

In this thesis, heritage is understood as being dependent of specific institutional information practices, and defined by the way institutions collect, classify, contextualise and present cultural heritage. Such institutional processes of heritagisation are exemplified by archivalisation and musealisation and constitutes a special heritagisation process of objects and information. These processes are, together with professional expertise and authority, supported by institutional information structures.

Research problem

One of the limitations with earlier research is that participatory production rarely have been studied as a specific mode of producing cultural heritage. Consequently, the differences and relations between institutional heritage production and participatory production have not been discussed in terms of the apparent clash between the two different modes of heritage production. This is problematic given that such a clash ought to have consequences for how well those modes of production can operate together.

This thesis aims to explain how digital participation in the activities of cultural heritage institutions can produce cultural heritage in contexts of institutional production of heritage and the conditions for how participants can contribute to a more inclusive production of heritage.

The thesis combines perspectives from three different scholarly fields building on conceptual and theoretical understanding from information studies, cultural heritage studies and memory studies.

Aim and research questions

The three research questions of the thesis are answered through four studies. They are independent from each other and three of them are published in different scholarly journals. The first three (I, II and III) are empirical studies in which practical examples of participatory project in a cultural heritage institutional setting is studied. They are investigating the connection between information and cultural heritage and how information structures and participants influence each other. The fourth study takes, in

contrast to the first three ones, a conceptual perspective to participation and focuses on how participation have been problematised as a democratising practice in the academic literature.

Table 4 below gives an overview of the research questions and the studies through which they are answered.

Main research questions	Sub research questions	Study through which the question is answered
1. How is cultural heritage produced through digital participation?	a) How does participatory created information influence existing cultural heritage collections? b) How is participatory created information influenced by being incorporated in heritage collections?	I, III
2. Under what conditions are participants able to change the institutional production of cultural heritage?	a) How do information structures shape the conditions and valuation mechanisms of heritagisation? b) How do participants question and negotiate those conditions?	II
3. Under what conditions can institutional production of cultural heritage become more inclusive?	a) How is digital participation problematized as an inclusive phenomenon in the literature? b) How can alternative understandings of digital participation as a problem incite change?	IV

Table 1. Overview of main research questions, sub-research questions and empirical studies.

Studies I–IV

Paper I. Organization of user-generated information in image collections and impact of rhetorical mechanisms

The first paper answers the question of how cultural heritage is produced through digital participation by studying cases of incorporation of participatory created heritage metadata

into institutional collections. It questions participation as an open and democratic practice by focusing on the limiting effects of institutional information structures on participatorily created information. The background of the study lies in how participatorily created information such as user comments in free-text format are especially propagated as empowering users and their influence on cultural heritage. However, as the study shows, in adjusting user-created information to suit the collection management system in use, rhetorical mechanisms of institutional information structures dictate the conditions for how the participatorily generated information is incorporated into the institutional information structures.

The empirical material of the study was collected in 12 interviews with professionals, who reflected about their work with administration of user-comments to be added to image collections as metadata. The experience of the informants spans over six different image database-systems. By analysing the professionals' statements about how the information structures affected the decisions they made in the moderation process of participatorily created information, several rhetorical mechanisms were identified in the systems. This article shows that the design of collection management systems can cause user-generated information to be discriminated and lead to decreased data reliability, searchability and even loss of crowdsourced data. In particular, personal memories and perspectives are among the types of information that are most negatively affected. To conclude, collecting user comments is a problematic method to use for adding multiple perspectives to cultural heritage collections. It requires carefully designed collection management systems in order to avoid distortion of user-created information.

Paper II. Creating value of the past through negotiations in the present: balancing professional authority with influence of participants

The second paper studies the conditions of how participants are able to influence the institutional production of cultural heritage. The paper investigates how participatory influence in an archive transcription project hosted by the Copenhagen City Archive is made possible by interaction between participants and archivists on an online platform. By studying what possibilities participants have to affect the

processes, structures, and end usability of the information resource, their space for agency in relation to the archive institution is identified.

To be able to separate different levels of user influence, the concepts of maximalist and minimalist participation are used for theoretical framing of the study. They also offer a backdrop for understanding of the difficulties of automatically claiming participation as a democratic phenomenon, and help in clarifying that participation without possibilities of influencing its basic conditions cannot be counted as highly democratic.

The most of the material for paper II was collected from a discussion forum that is connected to the online transcription project *Funerals 1861–1912 (Begravelser 1861–1912)*. The project aims to transcribe handwritten funeral records from the city of Copenhagen in order to construct a database of research data containing Copenhageners' causes of death, professions and other living conditions at the time of creation of the records. Participating transcribers can communicate with the archive institution by writing in the discussion forum – a communication that elicits questions, suggestions for improving the database and participants' perspectives of what they hope to gain from participating in the project.

To analyse this communication, a framework inspired by Community-Based Participatory Research (CBPR) was used. The results show how participants used the forum to gain influence in the project, and how several elements of CBPR—shared influence, mutual development, and mutual use—were manifested in the forum. On the other hand, the results also point out that information structures pose several limitations to how institutions can adjust participatory projects in hindsight. The final decision to implement changes to institutional information structures also always rests with the professionals, not the participants. A conclusion from this study is that even though institutions aim for participatory influence, this aim can only be reached to the degree the institution has prepared their information structures to be flexible and welcoming of the changes that participants might suggest.

PAPER III. "I SHOT THIS PICTURE BECAUSE IT WAS VERY REAL": STRATEGIES FOR FRAMING AUTHENTIC PERSONAL EXPERIENCE IN A CULTURAL HERITAGE COLLECTION

This paper aims to increase understanding of how digital participatory collection works to produce cultural heritage, and specifically how it works to introduce dissonant perspectives to heritage into institutional collections. The study sets out by introducing heritage dissonance as an inherent mechanism in all types of heritage, as opposed to the idea that it is present only in specific contested types of heritage. By separating between meaning-making based on two different rationalities; institutional and non-institutional, the paper suggests how dissonances in meaning-making of cultural heritage can be founded in institutional structures, for example, structures for managing heritage information.

The empirical material for paper III was collected in *Samtidsbild,* a Swedish digital museum platform created by Stockholm County Museum for collecting contemporary photographs from the public.

Captions of contributed photographs were analysed in search for strategies of how participants bestow meaning and motivate their contributions as warrantable heritage in an institutional setting. The paper discusses also if these strategies are successful for introducing dissonant heritage interpretations to the institutional collections. The analysis is supported by a framework which rests on the theoretical premises of personal, emotional dimensions of meaning-making in heritage contexts – in other words, a framework to analyse non-institutional meaning-making. Constructing personal meaning of certain places or events are examples of how such emotional meanings can be generated. The results show that contributors' strategies for meaning-making of their contributions to *Samtidsbild,* despite the seemingly good opportunities of personal meaning-making remain within the bounds of an authorised discourse of heritage, without introducing dissonant perspectives. The participants' strategies of heritage interpretation are, however, based on personal experiences and emotive responses and emerges as an alternative to institutional rationalities of interpretation, displaying a more personal and authentic perspective of heritage.

The conclusion is that the most important function of *Samtidsbild* and similar initiatives might not be to collect dissonant heritage interpretations but rather to provide means to broaden the perspectives of institutional heritage by including input from individuals whose heritage values

are already, to some extent, aligned with those of the institutions.

PAPER IV. CHALLENGING THE PROBLEM OF UN-DEMOCRATIC PARTICIPATION. FROM DESTRUCTION TO RE-CONSTRUCTION OF HERITAGE

Paper IV was spurred by the aim to contribute to the understanding of digital participation in heritage collections as a democratising practice. In a considerable part of the scholarly literature, digital participation is described as a problem because of the lack of participatory influence. Thus, it is suggested to fail to work as a method to shift cultural heritage processes to a more open and inclusive direction. The problem is traced back to various factors and conditions, of which the core seems to be the balance of institutional and professional openness versus the duty to preserve and safe-guard cultural heritage. The paper IV sets out to identify and question silent assumptions concerning how the insufficient influence of participants is conceived of as a problem.

The methodological approach for the study draws from Carol Bacchi's method for studying problematisations. Termed "what's the problem represented to be?" (WPR), this method puts emphasis on the analysis of different ontological elements of problematisations, including key concepts, binaries and categories. Three carefully selected scholarly texts made up the material for analysis, all of which incorporated a problematisation of insufficient participatory agency.

The results of the study show that participation is problematised based on the assumption that participatory agency risks jeopardising the protection of heritage and thus leads to parts of the public memory to become forgotten. To challenge the idea that participatory agency is destructive, the paper IV argues for elaborating an understanding of what forgetting entails for heritage. Framing forgetting as a potentially both harmful and generative concept enables a separation of destructive forgetting (e.g., destruction of historical evidence) and constructive forgetting (re-contextualisation).

The main practical implication of the paper IV is that by understanding forgetting as a potentially beneficial activity for the representation and construction of heritage, it provides a conceptual rationale for facilitating re-contextualisation in the design of multi-layered information structures for heritage collections.

Results and discussion

The thesis found that the possibilities to create inclusive institutional heritage collections through digital participation is limited and that the diversity of participants and their motivation to participate are dependent on cultural frameworks. If participation is a successful method to use for inclusion can only be decided case by case. It is dependent on who is to be included, why, and how well they identify with the cultural frameworks, which direct the cultural heritage definitions of the institution. However, the participatory frameworks do not necessarily need to be culturally or socially different for participation to have a diversifying and complementing effect on the collections. Even participation of participants whose majority share the frames of reference with cultural heritage experts at the institutions can lead to a diversification of expertise and negotiation of such institutional heritagisation practices as archivalisation and musealisation.

The findings also direct attention to that in order to facilitate diversification, institutions need to implement flexible information structures that provide support for user-generated information. This might require institutions to re-evaluate their approach to preservation and to open up for a more flexible method which recognises constructive forgetting and negotiation of heritage as part of their responsibility and day-to-day work with heritage collections. One of the biggest challenges for institutions is therefore how to tune in to such a more dynamic approach to heritage and how to prepare for participatorily generated suggestions and additions, while at the same time, keeping up with caring for integrity and preservation of heritage and heritage metadata.

One of the solutions proposed in the paper IV to this dilemma is to utilise the possibilities for incorporating complexity in digital structures. By increasingly working with more multi-layered structures that are able to document metadata-about-metadata (for example, to state provenance or time of edition) or extended possibilities to add contextual metadata, such as personal narratives, institutional information structures can be made to be more inclusive for participatory contributions. However, such complexity can also be expected to increase the demand for a more effective and flexible presentation of cultural

heritage information to avoid exposing users, may it be experts or non-experts, to information overload.

Taken together, the described flexibility of institutional information structures can lead to a more *sustainable cultural heritage* in the sense that it can provide a solution that helps to create more inclusive cultural heritage for a more inclusive, democratic and conflict-free society. This could mean building an increased tolerance for contesting heritage meanings and accepting heritage negotiations. Acceptance of a messier heritage approach also entails the acceptance of change and even controlled and constructive destruction as a natural part of the heritage life cycle.

References

For references, see the original edition of the text:
urn:nbn:se:uu:diva-499608

Tidskrift för ABM | Vol. 8. No. 1, 2023.
Institutionen för ABM, Uppsala Universitet.

AMALIA JUNESTRÖM[11]

Content Moderation and Fact-Checking: A Study of Journalists' Information Practices in the Contemporary News Media Landscape – Svensk sammanfattning

Det här bidraget består av den svenska sammanfattning som ingår i avhandlingen *Content Moderation and Fact-Checking: A Study of Journalists' Information Practices in the Contemporary News Media Landscape*. Avhandlingen försvarades den 27 maj 2022 och är publicerad i serien *Skrifter utgivna av Inst. för ABM* vid Uppsala universitet. Avhandlingskappan är fritt tillgänglig på DiVa (www.diva-portal.org).

Nyckelord: informationspraktiker, informationslandskap, genre, informationsinfrastrukturer, diskurser, moderering, faktagranskning

---

11 Amalia Juneström vikarierar som universitetslektor vid Institutionen för ABM, Uppsala universitet. Där undervisar hon på masterprogrammet i ABM och masterprogrammet i digital humaniora och är även verksam inom forskningsprojektet CAPTURE. I sin forskning har hon undersökt journalistiska informationspraktiker. Email: amalia.junestrom@abm.uu.se.

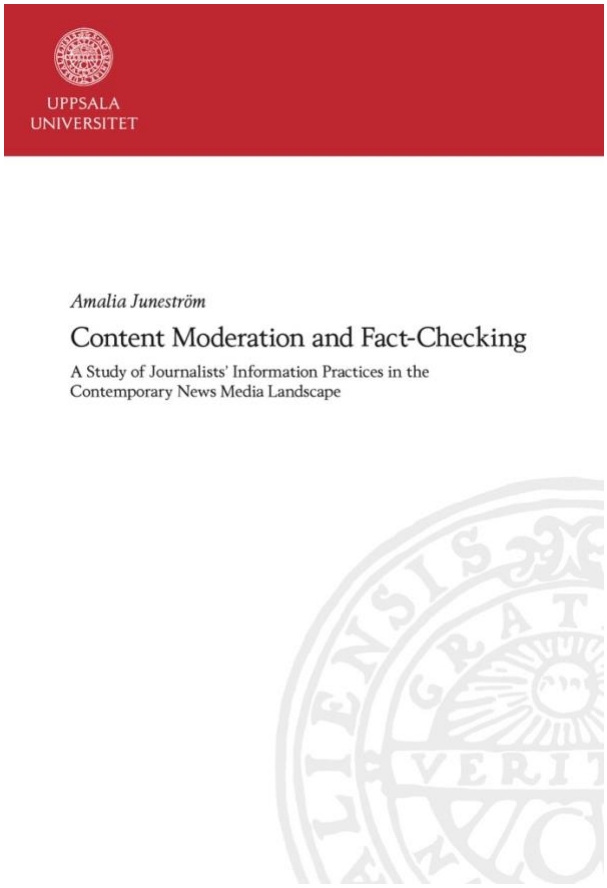

Bild 1. Avhandlingens tryckta framsida

Svensk sammanfattning

INLEDNING

Denna avhandling utforskar hur journalister utövar två specifika informationspraktiker som båda har växt fram parallellt med att journalistiken i allt högre grad har digitaliserats. Praktikerna – moderering av kommentarsfält på nätet samt faktagranskning av påståenden som i hög grad sprids på internet – har tagit form inom dagens medielandskap och blivit en del av detta.

Avhandlingens syfte är att belysa hur journalisters sätt att hantera kommentarsfält på nätet och att faktagranska har kommit att bli en del av ett digitalt journalistiskt informationslandskap. Mer specifikt avser avhandlingen att öka kunskapen om hur stötande användarkommentarer till

nyhetsartiklar och problemet med desinformation har bidragit till att skapa nya informationspraktiker.

Två av avhandlingens delstudier fokuserar på journalisters moderering av kommentarsfält till nyhetsartiklar på nätet och två på faktagranskning.

ARTIKEL I

Avhandlingens första delstudie syftar till att belysa hur professionella praktiker, som uppstått för att hantera stötande inlägg i kommentarsfält till nyhetsartiklar på nätet, har utvecklats i ett digitalt nyhetsjournalistiskt informationslandskap.

I studien genomfördes tio semi-strukturerade intervjuer med journalister från fyra olika länder – Sverige, Danmark, Tyskland och Kanada – om deras arbete med att moderera kommentarsfält under nyhetsartiklar som publicerats på nätet. I undersökningen tillämpades en variant av praktikteori för att förklara hur journalister resonerar kring sitt arbete med att bemöta användare i kommentarsfält. Det teoretiska ramverket är inspirerat av Lloyds informationslandskapsbegrepp, vilket är ett begrepp som teoretiserar den miljö som journalisterna verkar inom. I den här studien användes informationslandskapet som en förklaringsmodell för hur journalisters upplevelser och erfarenheter medverkar till att informationspraktiker förändras och utvecklas inom en specifik kontext.

Artikeln belyser betydelsen av hur normer, värderingar och ideologier, liksom regler och känslor, är avgörande för hur journalister uppfattar vad det är de gör när de arbetar med kommentarsfält.

Artikeln visar också att journalister i hög grad stödjer sig på samma ideologiska resonemang när de förklarar helt olika sätt att hantera användarkommentarer. Exempelvis kunde de intervjuade journalisterna använda främjandet av demokratiska principer såsom yttrandefrihet som argument både för och emot en strikt moderering av kommentarsfält. Det framgår även att den gemenskap som journalister uppfattar att de är en del av och som påverkar hur de förstår sina arbetsuppgifter inte är särskilt beroende av utövarnas geografiska hemvist eller specifika organisationstillhörighet. Istället har det nationsöverskridande informationslandskapet, som journalisterna är en del av, större betydelse för hur de uppfattar sitt uppdrag.

ARTIKEL II

Avhandlingens andra delstudie syftar till att skapa förståelse för hur vitt skilda praktiker för att bemöta problemet med stötande användarbeteenden i kommentarsfält till nyhetsartiklar är sammanvävda i en dynamisk mediestruktur.

Studiens empiriska material utgjordes av samma semi-strukturerade intervjuer som låg till grund för avhandlingens första delstudie. I studien analyserades tio intervjuer med journalister från fyra olika länder. Analysens fokus låg på journalisternas berättelser om hur deras arbetsmetoder påverkar andra aktörer, som också verkar inom kommentarsfälten, bemöter användarkommentarer som av dem uppfattas som stötande.

I artikeln används informationsinfrastruktur som teoretiskt begrepp för att ringa in hur praktiker bidrar till nya arbetssätt och mediestrukturer.

Några av de praktiker som ingår i strukturen, t.ex. moderering av användarkommentarer och hashtaggning, förefaller vid ett första ögonkast att ha mycket lite gemensamt. Drivkrafterna till varför olika aktörer engagerar sig i kommentarsfält skiljer sig också åt och i kommentarsfälten möts vitt skilda aktörer och praktiker som har en sak gemensamt, nämligen att de uppkommit som en reaktion på att vissa användarkommentarer ibland uppfattas som stötande.

Snarare än att förändring skapas av enskilda aktörer belyser artikeln att en mångfasetterad struktur av olika praktiker bidrar till förändring genom att de verkar i relation till varandra.

ARTIKEL III

Avhandlingens tredje delstudie syftar till att skapa förståelse för faktagranskning som en samtida journalistisk genre. I undersökningen bidrog genrebegreppet till att förklara hur tre amerikanska faktagransknings-organisationer arbetar med sina webbsidor.

I studien analyserades innehåll från Snopes, FactChecks och PolitiFacts webbsidor. Materialet bestod av ett urval av faktagranskningar och visuella element som publicerats på organisationernas webbsidor under två olika perioder. Materialinsamlingen ägde rum vid två tillfällen med sex

månaders mellanrum under corona-pandemin (den 18 till 23 mars respektive 22 till 28 september 2020). På så sätt utgjordes en stor del av materialet av rapporteringar relaterade till covid-19. Analysen av materialet fokuserade på hur faktagranskarna kommunicerade sina utlåtanden visuellt och textuellt.

Studien fann att faktagranskarnas arbetsprocesser, metoder och sätt att kommunicera i hög grad har kommit att bli typifierade. Ett exempel på ett typifierat arbetssätt är de värderingssystem som uppkommit för att betygssätta sanningshalten i de påståenden som faktagranskas. Ofta använder sig faktagranskare av färgglada symboler för att utvärdera och visualisera sina utlåtanden. Andra exempel är de tydliga och vägledande rubriker som informerar läsarna om utslaget av en faktagranskning innan de läst artikeln samt de lättillgängliga artikelstrukturer som kännetecknar faktagranskningarna.

Artikeln belyser hur de uttryck som faktagranskningsgenren tar sig har formats av förekomsten av olika behov inom dagens informationslandskap. Exempelvis har uppfattningen om att det finns ett behov av att återskapa ett förlorat eller skadat förtroende för nyhetsmedier hos allmänheten bidragit till att utforma nya metoder för att på ett transparent sätt kommunicera vad det är faktagranskare gör. Även om de värderingar och föreställningar som dominerar traditionella nyhetsmediediskurser återfinns i ett faktagranskningssammanhang har också föreställningar, som annars är kännetecknande för ett akademiskt arbetssätt, fått betydelse för hur faktagranskarna arbetar. Tillsammans har dessa värderingar bidragit till att ge upphov till en ny journalistisk genre som skiljer sig åt från andra nyhetsgenrer. Denna nya genre har införlivats i en samtida nyhetskultur och bidrar i sin tur till att påverka och förändra det journalistiska informationslandskapet.

Artikel IV

Avhandlingens fjärde delstudie syftar till att belysa hur en samtida faktagranskningspraktik konstrueras diskursivt i svenska nyhetsmedier och till att åskådliggöra ideologiska föreställningar och värderingar om faktagranskning i samhället. I undersökningen analyserades ett urval texter som publicerats i två av Sveriges största morgontidningar och olika diskurser som kännetecknar hur faktagranskning

framställs i svensk nyhetsmedia identifierades. Studien utgick från Faircloughs kritiska diskursteori (CDA). Genom tillämpning av ett diskursteoretiskt tillvägagångssätt gjorde studien vissa teoretiska ställningstaganden som har att göra med språkets sociala betydelse. En utgångspunkt i studien är till exempel att språket är meningsskapande och bidrar till att forma hur människor handlar och förhåller sig till världen omkring dem.

I studien identifierades tre övergripande diskurser som kontrasterar och motsäger varandra. Den första förhåller sig positiv till faktagranskning och bejakar de arbetssätt som ingår i praktiken. Den andra diskursen förhåller sig negativ till faktagranskning och bestrider bland annat att det skulle vara en effektiv arbetsmetod för att bekämpa desinformation. Slutligen identifierades en tredje diskurs som på engelska benämndes ”the agency discourse”. Denna diskurs reflekterar olika uppfattningar om aktörernas roll i faktagranskningen och föreställningar om vem som bör ta på sig ansvaret för att bekämpa spridning av felaktig och missvisande information.

Samtliga diskurser rymmer flera underdiskurser. Den första diskursen rymmer exempelvis två underdiskurser varav den ena reflekterar uppfattningar om faktagranskning som ett botemedel mot desinformation och den andra uppfattningar om faktagranskning som en metod att fastslå vad som är sant och ett sätt att upprätthålla objektivitet. Inom den andra diskursen identifierades tre stycken underdiskurser. Den första ifrågasätter faktagranskning som ett effektivt verktyg i kampen mot desinformation. Den andra underdiskursen reflekterarar föreställningar om att faktagranskning skulle kunna vara skadligt för demokratin och den tredje en oro över att praktiken skulle kunna bidra till att legitimera och ytterligare befästa konspirationsteorier. Inom den tredje övergripande diskursen identifierades två underdiskurser. Den ena etablerar faktagranskning som ett journalistiskt åtagande och den andra reflekterar föreställningar om den roll sociala medier spelar eller bör spela i arbetet med att faktagranska information som sprids på nätet.

Studien fann att olika diskurser överlappar varandra och att flera av dem står i direkt konflikt med varandra. Artikeln visar även att faktagranskning i mycket hög grad är politiserad. Det framgår också att uppfattningarna om faktagranskning är betydligt mer komplexa än vad de ger

intryck av att vara i den växande forskningslitteratur som undersöker ämnet.

SLUTSATSER

Avhandlingen visar att de två undersökta informationspraktikerna – moderering av kommentarsfält och faktagranskning – är sammanflätade med journalistiska normer och värderingar samt med de sätt på vilka journalister uppfattar sitt uppdrag. Efterhand som praktikerna har tagit form har de vävts in i journalisternas uppdrag på ett sådant sätt att de blivit en del av dagens journalistiska informationslandskap. Avhandlingen visar också att de sätt på vilka praktikerna utövas har uppstått genom en förhandlingsprocess där olika journalistiska principer ställs mot varandra. Exempelvis vägs demokratiska principer om yttrandefrihet mot journalistisk kvalitet i en diskussion om hur användarkommentarer och desinformation ska hanteras och bemötas. De motiv som driver journalister att engagera sig i moderering av användar-kommentarer och faktagranskning kännetecknas av en spänning mellan olika handlingsalternativ. Det framkommer exempelvis att motiven bakom journalisternas val att inlemma de två praktikerna i sin yrkesutövning är mycket mer komplexa än vad de vid en första anblick ger sken av att vara. De sätt på vilka problemen med aggressivt användarbeteende och spridning av desinformation hanteras är långt ifrån givna och de är i högsta grad förhandlingsbara.

Referenser

For references, see the original edition of the text: urn:nbn:se:uu:diva-471068