Special Issue: Data Reflectivity: New Pathways in Bridging Datafication and User Studies

CONVERGENCE

Repairing what's not broken – Algorithm repair manual as reflexivity device

Convergence: The International Journal of Research into New Media Technologies
2024, Vol. 30(6) 1871–1892

DOI: 10.1177/13548565241261978
journals.sagepub.com/home/con
Sage

Ana Pop Stefanija and Jo Pierson
Vrije Universiteit Brussel, Belgium

Abstract

In this article, we outline an innovative participatory method for reflexive engagement with algorithmic systems and the underlying processes of datafication that accompany them. Faced with the challenges of thinking critically and reflexively about algorithmic systems, both as non-expert individuals and expert researchers, we develop and elaborate on an approach for engaging participants in *thinking with*, – *through* and – *about* algorithmic artifacts. In developing our approach, we start from the premise that algorithms are always broken, and we devise Repair Manuals as productive reflexivity devices that will enable for reflective and reflexive encounters with algorithmic artifacts. Borrowing from the approaches developed by Shannon Mattern and Joseph Dumit, we take algorithmic data artifacts as entry points to embark on an investigative, self-learning and sense-making journey of the inevitable entanglement between the individuals and the all-encompassing algorithmic systems. The results from our study show that this approach offers valuable opportunities and insights both for educators and for researchers, and can be used equally for empowerment and educational goals.

Keywords

algorithm repair, algorithm repair manual, algorithmic artifact analysis, data reflexivity, critical pedagogy, methodology, human agency, participatory research, data reflectivity

Introduction

'*Become an egg donor. Give a beautiful gift to an unconditionally loving family… US$200.000 in compensation…*' – this is the ad that Irene,[1] a young female participant in our study, received 1 day on her Instagram feed (see image 1). Irene asks herself:

Corresponding author:
Ana Pop Stefanija, imec-SMIT, Vrije Universiteit Brussel, Pleainlaan 9, Brussel 1050, Belgium.
Email: ana.popstefanija@vub.be

> '*Why did I get this ad? When clicking on the "why am I seeing this ad" button, I am told that it's because I am a female aged between 18 and 25 living in the San Francisco Bay Area. There is no additional information* […] *I first noticed this type when I was 19 and living in San Francisco, the most expensive city in the world at that time. I was a first-year bachelor's student, and my financial situation was quite severe. I took notice of the ads because they appealed to my specific situation* […] *My perception is that these ads prey on the vulnerability of young women (especially students and/or recent graduates) who need quick cash to pay off loans and get through school'*.

Thinking critically and reflexively about algorithms and their data outputs is challenging for experts and academics, and even more so for non-expert individuals. Characterized by opacity due to their complex nature and the intentional corporate obfuscation, algorithmic systems pose significant challenges for their investigation and for studying their effects on individuals and societies. Algorithms are not *study-ready* and in the last decade a number of critical data and algorithmic studies (Iliadis and Russo, 2016; Lindgren, 2023; Kitchin, 2017) approaches have been developed in the attempts to find ways to investigate them (see Pop Stefanija and Pierson, 2020). The challenges are even greater when individuals, as non-experts, are trying to learn, make-sense of, and understand algorithms and their impacts. Lacking technical knowledge and tools, individuals rely on own theories, ways of thinking and ideas of 'what algorithms are, what they should be, how they function and what these imaginations in turn make possible' (Bucher, 2017: p. 40).

The research we present in this article is a result of a pedagogical and a research dilemma. The pedagogical emerges from our work as educators: how to teach students in media and communication studies about datafication and algorithmic systems and their societal impact. Preparing for a new semester in 2022, we wanted to devise a learning path that will enable the students to critically reflect on data and algorithmic systems. We envisioned this as not just *learning* in an abstract, theoretical manner, but also as *acquiring* critical knowledge in an experiential manner by *thinking with*, – *through* and – *about* algorithmic artifacts. Considering

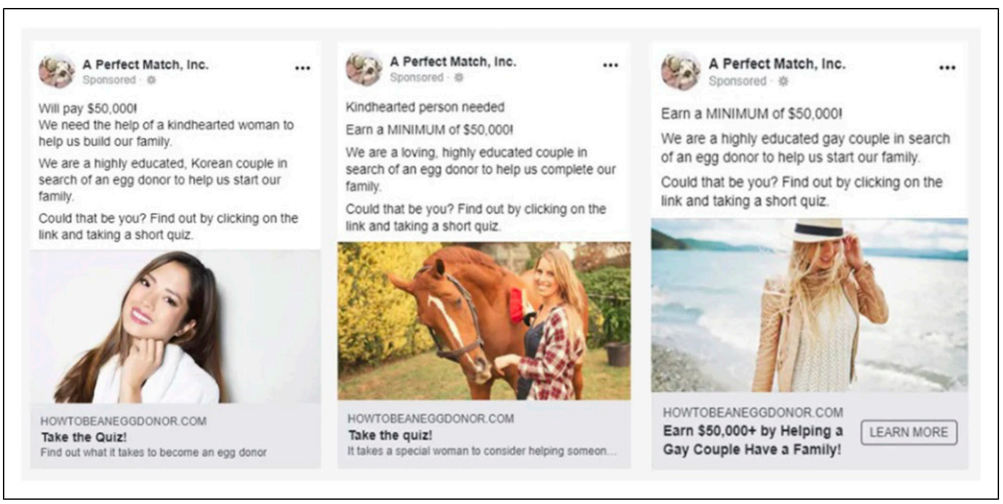

Image 1. Screenshots of the algorithmic artifact of Irene.

the students as potential *researchers-in-the-making*, we wanted to teach them to think, experiment, and be open to non-traditional modes of data-related scientific inquiry, applying in practice what they learn about in class.

As researchers, we were also challenged by the question of how to enable a learning and reflexive process for non-expert, common citizens? At the beginning of the semester our students were non-expert individuals – their ability to understand and interact with AI systems and their outputs, in a (self-)reflective manner, was diminished by the opaque nature and the tendency of these systems to (seamlessly) hide in the background. We wanted to devise, test, and evaluate a method that 1) could be used to enable citizens to critically understand these systems; 2) that will transform our non-expert students into qualitative researchers, and 3) that will also enable us, as researchers, to collect insights on their learning and sense-making processes. How to design an approach that could also be used by other researchers and educators, that will enable critical self-reflection based on understanding, reflection, and sense-making, both for non-expert and expert citizens?

We found methodological inspiration in the courses of Shannon Matter (2022) and Joseph Dumit (2014). We developed these approaches further, taking the concept of Repair Manual and devising it into a reflectivity device for studying and understanding algorithmic systems. Combining this tool with a critical pedagogy approach (Freire 2000; Markham, 2019) proved a promising and productive method to teach about and study these systems.

In our approach, we take algorithmic data outputs as an entry point to investigate, reflect, and make sense about both algorithmic systems and the underlying datafication processes they rely on. We start from the premise that algorithms are always broken, in a way. In the sections that follow, we will elaborate on why we consider algorithms as broken and why we see this standpoint as a productive prism for reflexively studying and investigating them. We further briefly present the pedagogical sources serving as inspiration. We later elaborate in greater detail the research exercise we did with our students, their learning journey, and the valuable insights, both theirs and ours. At the end, we offer our own reflection on the process, the suitability of the approach for critical thinking and learning, for teaching and for research, and we point to areas that require particular attention.

'Broken' algorithms as reflexivity devices

By considering algorithms as 'always broken', we mean that algorithmic systems, as socio-technical assemblages relying on ubiquitous datafication processes (Van Dijck, 2014) and inferential analytics (Wachter and Mittelstadt, 2018), always embed power asymmetries. Algorithmic systems always work for some, but not for all. They have social power (Beer, 2017) that comes through both their outputs and their effects. Their outputs affect individuals to various degrees, and they mirror, perpetuate, and reinforce societal inequalities that affect people in numerous ways (Eubanks, 2018; McQuillan, 2022; O'Neil, 2016). Part of a socio-technical assemblage, algorithmic systems are often not immune to inscribed biases and established discriminative practices. Algorithmic systems are often weaponized for political and corporate interests and for keeping the societal status quo, by producing and reproducing particular values and perpetuating (unjust) practices.

As a part of a broader socio-technical assemblage, algorithms can be broken in various components and stages of their entire life-cycle – from their inception to their deployment, from the technical design work to their societal impact. Algorithms are broken in the sense that they often offer technical solutions for (purely) societal issues. The very recognition of something as an issue often implies a power position of recognizing some aspects and ignoring others (Rieder et al., 2022). Data used both for training the algorithms and as computational input for algorithmic outputs is

often labeled based on essentialist choices, prioritizing one meaning of the data over other (Crawford and Paglen, 2019, 2021). This data is often mistaken as being the same as what it represents, traces, or signals (Gitelman, 2013; Hildebrandt, 2022). Additionally, there is always a choice architecture about what (data) will be considered a proxy for what. Understood as 'the machine-readable variables that stand for relevant features and targets' (Hildebrandt, 2022: p. 2), often these choices and proxies minimize the complexity of humans and societies into a single, essentialist understanding, where the proxy is mistaken for what it stands for (ibid.). The brokenness of algorithms might emerge from the fact that technology designers are constricted by socio-technical normativities and their practices 'come to stabilize and embed norms within certain systems by setting up plans or programs for learning machines to execute (metrics, ground truth dataset, optimization function)' (Grosman and Reigeluth, 2019, p. 9). Technology designers possess an 'institutionalized authority to construct reality' (Meuser and Nagel, 2009: p. 19). As Barabas et al. (2020) outline, they tend to uncritically inherit dominant modes of seeing and understanding the world and of accepting the default assumptions, leading to discriminatory design. Algorithmic technologies are not produced in a societal vacuum. They are influenced by, and inscribe the needs and demands, the goals and the priorities, and the ideologies of the actors requesting and using these technologies (Miceli et al., 2022; Suchman, 2006). These business decisions do have societal impact.

Both academic research and investigative journalists have documented a number of these algorithmic harms. From Sweeney's (2013) racial discrimination in ad delivery, to the racial disparities in risk assessment algorithms (Angwin et al., 2016), to the inscribed biases in image recognition databases (Buolamwini and Gebru, 2018; Crawford and Paglen, 2019), to 'fraud prediction' algorithms automating injustice (Constantaras et al., 2023), to racial biases in ad portals equating black girls with porn (Yin and Sankin, 2020), to replicating location-based and class-based historical inequalities for university entry requirements (Benjamin, 2022) – these are just a handful of the numerous real-life algorithmic harms.

In that sense, when we talk about algorithms being broken, we do not mean a breakdown in the sense of not functioning well, or at all. On the contrary, these are algorithmic systems that function perfectly per their goals. They provide personalized and targeted outputs, sort (people and things), recommend actions, follow the established logic and metrics for optimal functioning. We also borrow partially the concept of 'broken world thinking' from Jackson (2014), who understands this position as both normative and ontological – 'it makes claims about the nature of technology and its relationship to broader social worlds' (p. 221), and as empirical and methodological in the sense that it invites new and different kind of research.

The premise that algorithms are always broken, implies that algorithms could and should be repaired. We follow Markham's approach (2019) that taking a critical stance 'begins with the idea that there's something wrong and works to investigate the who, what, where, when, and how of this wrongness'. (p. 757). We understand repair not as an intervention on a technically broken, malfunctioning algorithm, but as a corrective work on a perceived brokenness of algorithms (see also Velkova and Kaun, 2019). As such, repair does not (always) refer to a state in which the algorithm is 'fixed' once and for all, but it is an ongoing critical process of negotiation between the socio-materiality of the algorithm and its users. As Graham and Thrift (2007) outline – 'fault-finding and repair is a process of ongoing, situated inquiry' (p. 4).

Imagining how algorithms could and should be repaired necessitates first their inspection. We envision this examination by taking an algorithmic artifact as an entry point, and undergoing two distinct, but related processes – the process of reflectivity, followed by a process of reflexivity. We understand reflectivity in the sense of the Call for papers of this special issue – algorithmic outputs, according to the dominant algorithmic paradigm and ideology of dataism (Van Dijck, 2014), reflect

users' behaviors. Based on troves of captured and approximated data taken as real representation of individuals' 'qualities, preferences, characteristics, intentions, needs and wants of users' (Lehtiniemi, 2016: p. 4), each algorithmic output is a result of the user's produced and assigned algorithmic identity – 'an identity formation that works through mathematical algorithms to infer categories of identity on otherwise anonymous beings' (Cheney-Lippold, 2011: p. 165). In that sense, when encountering a single algorithmic output, we can say that the individuals are encountering their algorithmic reflection – as if 'You are the algorithmic outputs you receive!'.

The process of reflectivity is the first, essential encounter with the algorithmic artifact. Once this process of reflectivity, of meeting one's algorithmic identity is done, the process of reflexivity starts. This is a process of sense and knowledge-making. It includes distinct stages, from the simple noticing and observation of the algorithmic artifacts, to finding hidden aspects and investigating further, by engaging existing experience, practice, and tacit and affective knowledge. Reflexivity should enable the production of new knowledge and often new skills, providing essential insights that will further inform and develop the strategies for repair.

We describe, elaborate, and reflect on how these processes were unfolding in practice.

The search for pedagogical and scientific approach

Making algorithms and the underlying datafication systems knowable and study-ready to individuals as common users, requires tactics and approaches that overcome and/or maneuver around their opacity. Often, these require innovative and creative methods to achieve that. A number of studies have already experimented with various approaches. These range from using data rights as tools and using platforms' transparency tools (Pop Stefanija and Pierson, 2023), to looking for unusual places for finding algorithmic encounters (Bucher, 2017), to creating creative ways of sense-making (Bishop and Kant, 2023) and soliciting experiences (Büchi et al., 2023; Kapsch, 2022; Lomborg and Kapsch, 2019; Lupton 2021; Swart, 2021), and to making data visible (Lupton, 2017), among others. The approaches differ in their 'sites of encounters', mechanisms of openings (Weiskopf and Hansen, 2022) and objects of study. For example, Ruckenstein and Granroth (2019) focus on emotional reactions, while Bucher (2017) focuses on sites of affect. Their shared aim is to make the datafication processes and their algorithmic outputs visible to individuals, and even more, study-ready to enable a critical understanding of these processes.

Already having an experience with innovative ways of studying algorithms (see Pop Stefanija and Pierson, 2020, 2023), we were looking for a tool which will enable individuals to learn and investigate datafication and algorithmic systems in a critical and reflexive manner. For this there were a number of self-imposed requirements. First, our goal was to teach our students how to do critical research in practice, not only in theory. Second, the chosen approach should enable our students to investigate these systems while occupying two simultaneous positions: as non-expert individuals affected by an algorithmic decision-making, and as researchers-in-the-making critically dissecting and analyzing an algorithmic artifact. Finally, the approach should not require technical knowledge or opening of the 'black box'. As a critical research practice, the critical method should 'reveal the tensions between institutional, automated, and economic systems, to surface and explain societal inequalities and relations of power, and to make visible people's actual experiences, actions, and hopes' (Pink, 2022: p. 751). Conceptualized like that, it should facilitate critical encounters with data and algorithmic systems that go beyond simple noticing, but enable an engaging and productive interaction, making algorithms reveal themselves and opening them up for inspection and reflection.

The approaches we found inspiration in and developed further, are two educational approaches by fellow academics – Shannon Mattern's *Redesigning the Academy* course (2022) and Joseph

Dumit's (2014) *Implosion project*. For her course, Professor Mattern developed an exercise that, amongst others, also borrows from Dumit's (2104) approach. Inviting her students to think how Academia can, through its various institutions and objects, be imagined differently, she is guiding them to re-think Academia's very foundations and produce a Repair Manual (*R. the Academy*, 2022). Through a critical process of *questioning* the wider contexts, assumptions, and localized applications, the manual aims to address (and later redress) 'failures and inequities in the larger academic system' (Mattern, 2022) by investigating aspects such as implied users, affordances, ideologies and power relations, and by imagining how the various academic objects can function as 'a tool of institutional and social repair' (ibid.). To do this, Mattern borrows from Dumit's pedagogical approach.

Elaborated in detail in his 'Writing the Implosion' article (2014), Dumit's main premise is that everything in the world, both man-made and natural, contains a multitude of co-existing aspects, actors, and processes, all coming together in a particular artifact and point in time. Through a 'violent collapse inward', it is possible to investigate artifacts by starting one's investigation at the surface of the artifact and from there unravel the different 'threads' making its tissue: 'labor, professional, material, technological, political, economic, symbolic, textual, bodily, historical, educational' (Dumit, 2014: p. 350). Uncovering their embeddedness, formed relations, and coming together will enable us to learn more about power relations, politics, traditions, interests, normativities – things existing outside the artifact that affect its use and the people around it.

We find Dumit's and Mattern's approach to artifacts valuable because it resembles the understanding of data and algorithmic artifacts as socio-technical assemblages – an ongoing organizing of multiplicities (Bucher, 2018). These multiplicities are made up of heterogenous objects, aspects, actors, and dimensions of people, knowledge, institutions, norms, traditions, etc., in a perpetual and constant shifting of relations between them (Bucher, 2018). Data and algorithmic assemblages may include 'systems of thought, forms of knowledge, finance, political economy, governmentalities and legalities, materialities and infrastructures, practices, organizations and institutions, subjectivities and communities, places and the marketplace where data are constituted' (Iliadis and Russo, 2016, p.3). In each algorithmic artifact, we can uncover the data feeding it, the algorithmic logic applied, the designers' normativities, the companies' business interests, and the societal values, norms, and structures. Understanding how these work within a larger system – an assemblage of different actors, they need to be understood and investigated *across*, not *within* a particular system (Ananny and Crawford, 2018, p.11). An analysis of this kind can, itself, be seen as a process of assembling the various aspects together, making them visible, foregrounding and bringing them to the surface.

Exercise setup

A repair manual first identifies a problem. Each repair work starts with a diagnostic phase, consisting of the creation of 'a consistent account of the troubled state of the machine from available pieces of unintegrated information' (Colmellere, 2016: p. 87) – a record of what is wrong, what does not work, why and how. One needs to acquire knowledge of how distant parts are interconnected, and how one (broken) part affects the others. Then, a manual is assembled that offers a know-how for repair – specific instructions and step-by-step guidelines. A repair manual is specific enough, applying to the problem at hand, but also general enough to cover the potential issues that might arise in the future. It takes into consideration its various potential uses by different users, acknowledging and foreseeing their specifics, needs, and contexts, so it is adjustable and robust enough.

The same principles apply to an Algorithm Repair Manual. The repairing process is preceded by an investigation and diagnostics process – what we called in our exercise the Algorithmic Artifact Analysis. Dumit's (2014) article outlines all the necessary steps to do one, so it served as a guideline and a mandatory text for the students.

We instructed the students to look for a personalized algorithmic artifact, which we define as an algorithmic output which is a result of automated decision-making based on datafication processes. In that sense, we asked our students to look into the particular workings of platform algorithms (Cotter, 2020; Fisher, 2021) – algorithms integrated into online decision-making systems (social media, recommender, or personalization systems), that they encounter and use in their daily lives. Some examples of algorithmic output were provided: a friend suggestion on Facebook; an ad shown on Facebook, Twitter, or Instagram; a 'Why am I seeing this ad' explanation (e.g., Twitter, Facebook); a suggested video on TikTok; a suggested movie on Netflix; a 'Made for you' playlist on Spotify; the Search/Explore posts on Instagram; or LinkedIn related outputs.

The students were prompted to start from the micro, the individual, from their own daily encounters with algorithmic outputs. This was the first step of the analysis – finding an artifact that tickles their curiosity. Once chosen, an iterative process of inspecting, studying, and untangling the artifact starts. Departing from the surface, that is, the easily perceivable, they gradually need to dive deeper. After spending some time with their artifact, the students were encouraged to write down their thoughts on a number of questions that might apply to their context: why did I get this recommendation/content/personalization? How does this artifact work? How does it affect me, diverse groups of people, or even society? What are the implications of the algorithmic output for me, and for others? What are the benefits, risks, harms? Does it work equally for everyone? If not – why not, and how? We encouraged the students to approach these questions by thinking about them from various aspects: power issues, labor issues, surveillance, personalization, privacy, autonomy, control, identity, black-boxing, technological power, etc.

Once they identified what needs repair and why, the students could start with the Repair process. Revealing the 'brokenness' of algorithms through the process of what can be seen as a 'forced breakdown' – and as such as a process of 'unblackboxing' (Graham and Thrift, 2007, p. 8), enabled the students to learn to see and engage with the technology in new ways (Colmellere, 2016). In the spirit of real-world manuals, they should offer specific instructions and step-by-step guidelines. Their reflections should give answers to (some of) the following questions: How would you repair this artifact? What is needed to repair it? Why do you care about these repairs? What should they achieve? Why are they important for you, for other groups and for society in particular? What would be the individual, societal, political, labor, technological, etc. implications of such repair? If power imbalances issues are identified – how does the repair tackle them? The aim is to think how to improve, redesign, or repair the artifact so to redress potential failures, harms, and inequities. Or how to be deployed differently and to function as a tool of technical and societal repair. If unrepairable – to even completely abolish the algorithmic artifact.

Since the aim of this research exercise was to reflect from a critical perspective, the students were instructed to approach artifacts both as individuals affected by algorithmic decision-making, and as researchers critically dissecting and analyzing an algorithmic artifact. Considering this was also an educational task, they were expected to use academic literature to justify and elaborate on the issues that required repair. Our aim was to see how well the students can identify the interrelatedness of technology practices with societal issues and to what extent they were able to substantiate their interventions by theoretical concepts, while considering both the individual and the wider socio-technological context. From our part, we adopted the approach of 'critical companionship' (Ziewitz and Singh, 2021). We held a workshop before the start of the exercise, and continuous support and

feedback were provided, thus facilitating and supporting the process of learning and critical reflexivity.

A total of 28 students participated in the research exercise and produced 19 repair manuals. They were all at the time students at an international postgraduate program in the researchers' University, following the course *Data, Privacy and Society*. The research was part of their mandatory written exercise, as being one of the two assignments they could opt for. The students could choose if they want to work individually or in groups: nine worked in pairs of two, the others individually. No information was collected about pre-existing knowledge on algorithms, but we expected that it varied among the students, having in common a shared interest to learn more.

Trip down the algorithmic rabbit hole

First encounters of a different kind

Because of the tendency to blend into the background, algorithmic systems become increasingly unnoticeable for individuals. This makes the user experience seamless and the presence of algorithmically produced outputs not (at once) apparent, often not even being recognized as outputs of algorithmic workings (Couldry and Mejias, 2019; Veale and Delacroix, 2020). As one of our participants, Irene, notes in her artifact analysis – '…When I saw this specific advert, it was displayed in the Instagram story function, wedged between posts from family and friends'. As Lomborg and Kapsch's (2019) research shows, when algorithmic outputs are perceived as smart, effective, and convenient, users stop noticing them. Participants Alexandra and Sebastian noted: 'Another thing that is of our knowledge is that the artifact was subtle, subtle enough that it could have been either used or discarded by the data subject without questioning its appearance in their feed'. Consequently, one of the first challenges of data reflectivity was to make these algorithmic outputs visible and recognizable. We guided the students to look for them in a few diverse ways.

Algorithms become the most noticeable when they malfunction (Ruckenstein and Granroth, 2019), surprise (Lomborg and Kapsch, 2019), irritate (Ytre-Arne and Moe, 2020), or in moments of perceived breakdown – when they don't behave as people expect (Bucher, 2017), making their failures and errors perceivable. That was the case for our participants Rishabh and Tine: 'While considering the app's target audience, it appears that the users are astrology believers. But the ad was directed towards a non-believer. In such cases, the algorithm did not recommend the ad properly and it went to the general population' (Rishabh & Tine). Aradau and Blanke (2022) use the concept of algorithmic trouble or 'glitches' to refer to these moments. Building on Donna Haraway's invitation to 'stay with the trouble' – understood as intentional stirring up and messing in a plethora of configurations of matters and meanings (ibid., p. 216.), we invited our participants to 'look for the trouble'. Ytre-Arne and Moe (2021) see a potentiality for user agency when noticing actively algorithms' imperfections instead of just blindly accepting them, thus opening the space for criticizing them (p. 820).

We also encouraged our participants to be 'on the lookout for' algorithmic artifacts that in fact function very well, too – to look for what Bucher (2017) calls 'whoa' moment, when we are faced with these very well-functioning personalization algorithms. As Seberger and Bowker (2021) say, infrastructures (data and algorithmic) can often become visible and known only through the qualities of their functioning (p. 1714). In those moments, they 'bring about the experience of absurd alienation', as if feeling out of place (ibid.) precisely because they work too well. As Kati asks herself: 'I am deeply interested **why I do not** [*our emphasis*] see any Russian propaganda

videos in my recommendations on YouTube?' – 'her' YouTube algorithm works so well, it is not providing her with content she is absolutely not interested in.

We advised the students to also look for artifacts that trigger emotional responses. Affective encounters (Bucher, 2017) are also productive to start reflecting on our interactions and entanglements with algorithmic systems. Ruckenstein and Granroth (2019) call this 'the emotional as methodological entry point' (p. 5) and see this as the opening-up of the possibility for a critical commentary, and for a process of reflection and sense-making. The affects and emotions can be negative, as is the case of Erika – 'Am I seeing this video because ASMR and organizing/restocking videos are often associated with femininity or being a woman? If that is the case, is there an assumption that I identify myself as a woman based on my personal data?'. Or they can cause discomfort, feel eerie or concerning, like the case of Alexandra and Sebastian, who received a Facebook suggestion for a group offering accommodation in Brussels, although Alexandra only discussed this information off Facebook. They can also be perceived as irritating and offending, as it was for Marcel and Milagros: 'Not only I don't use that brand, I also don't usually shave "smooth cuts" […] What I didn't like about this ad was the only relation between me and the commercialized product is my gender'.

The exercise of looking around and identifying algorithmic artifacts required from the students an *attentive engagement* with their algorithmic devices. Against the flow of mindlessly *scrolling around*, it necessitated a mindful and more careful interaction, an awareness of the algorithmically mediated surroundings, and incentives to stop and think. As Kati writes down in her notes: 'I haven't questioned why I see what I see in my recommendations, nevertheless, I cannot deny the importance of questioning, why I see this content as a recommendation on such a political matter as war'. This allowed them to notice the 'blind spots', in the sense of taking curated feeds for granted, but as we discuss later, also their blind spots as researchers – what *is* and *remains* a knowledge 'blind spot'.

The encounter with the algorithmic artifacts for the students was also an encounter with their algorithmic reflection – how the algorithm sees them and gives them recommendation based on that image. This opens the process of reflexivity.

First reflections of a different kind

After the first encounters, a process of investigation, reflection, interpretation, and sense-making begins. For our students, that also meant first looking for algorithmic openings, for entry points from where the algorithms can start unfolding and reveal their hidden aspects: data, creators, marketers, structures, or algorithmic logic. To do that, they often were using the very affordances of the platforms, such as 'Why am I seeing this ad' options (Image 2), looking into Privacy Policies and Terms of Service, or looking into their own app/platform settings (Image 3).

That also required the employment of theoretical concepts, discourses, and traditions of thought that the students themselves brought with them, a kind of an intellectual and academic baggage, for initial sense-making. Since the research assignment was unfolding throughout the entire semester, they had the chance to continuously encounter new theories and concepts, and thus also update their interpretative arsenal and conceptual vocabulary (Swart, 2021). This iterative and 'incremental learning' (D'Ignazio and Bhargava, 2016) provided them with more time and space to reflect, giving them the opportunity to explore and think-through gradually the more complex aspects. This is a result of the increase in their abilities, insights, and knowledge.

We see this in the separate assignment parts that we received from our students. The exercise was structured with two feedback sessions, so the students had the possibility to go through at least three reflection phases, detailing them in three separate documents. The first one was the choice of the

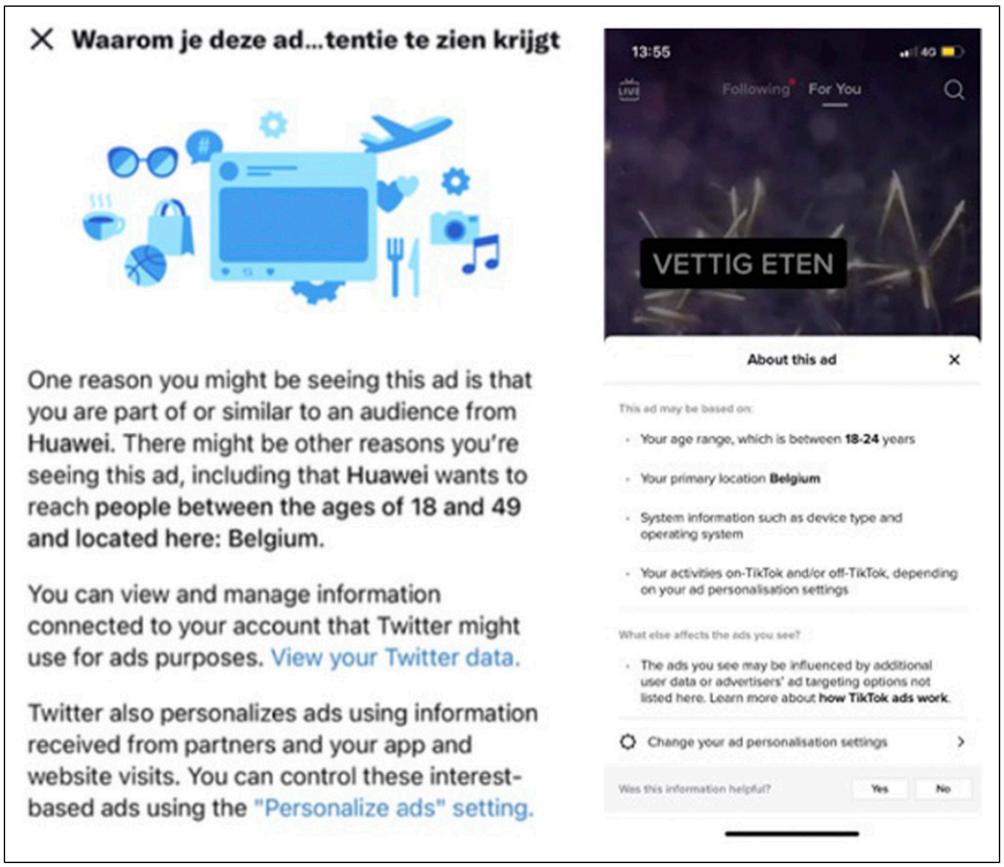

Image 2. Screenshot of 'Why am I seeing this ad' on TikTok (right, by Marcel and Milagros) and Twitter (left, by Ismail and Guojun.

artifact, the second one was the initial artifact analysis and the third one was the final assignment – the artifact analysis + repair manual. We use all these separate documents, containing data as texts, brainstorming notes (e.g., Image 4 and 5), screenshots (e.g., Image 1), sketches (Image 9), and designs, for our analysis. This also allowed us, as educators, to witness the transformation, development, enrichment and deepening of their reflections, analysis, and sense-making. As Markham (2019) says, 'lived experience is an ongoing negotiation of meaning, always within interactions' (p. 755). The same applies for the learning process too.

When it comes to sense-making, this is not a question of (mere) technical understanding, but it is an interpretative process (Ytre-Arne and Moe, 2020: p. 811). We can compare this process with the one of decoding, as a communicative practice (Lomborg and Kapsch, 2019): feeling, evaluating, knowing algorithms in daily life (p. 746). The process of decoding also could be understood as a process of spiraling from one question to another once we start to uncover and recognize the connection between the different layers of a one simple question. For example, in the case of Karina, she goes from being discontent of Spotify showing her content she does not identify with 'Why did Spotify offer me astrology themed content? (I don't believe in it)' to immediately asking herself 'How personal and artificial curation work on Spotify? (I believe the playlists are made by real

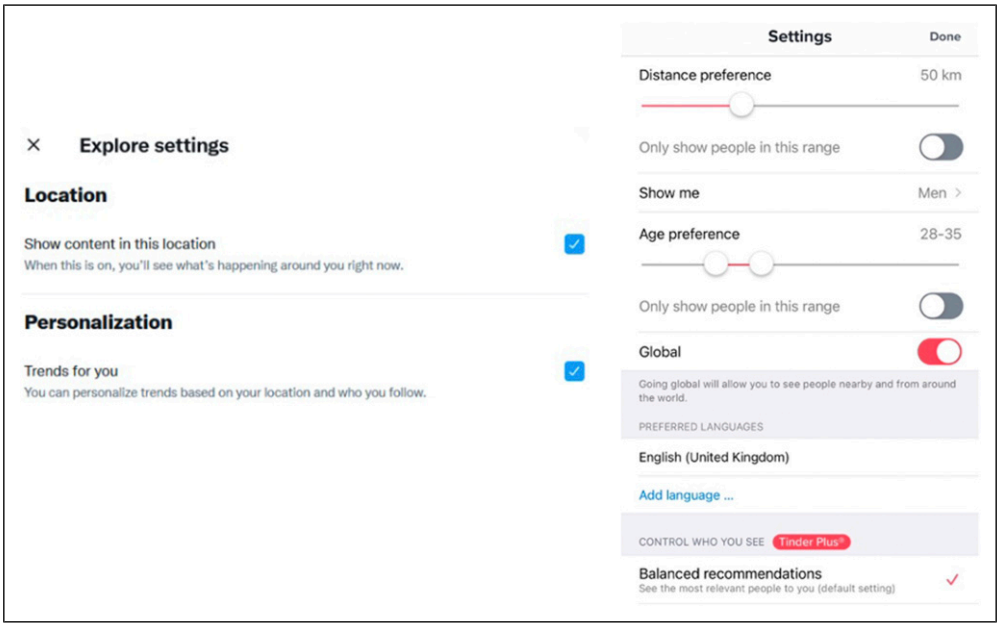

Image 3. Screenshot of own Twitter settings page (Gert, left) and Tinder (Karina, right).

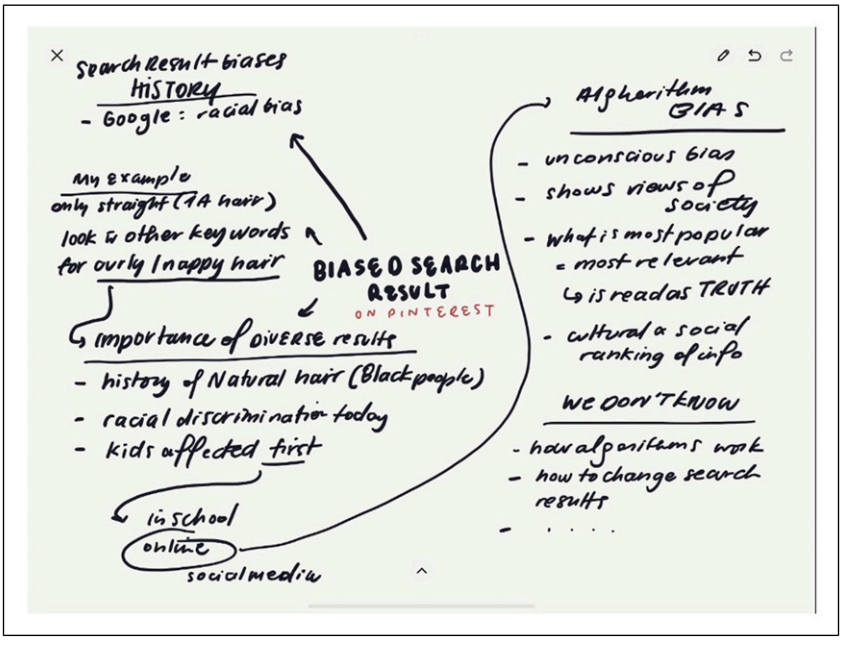

Image 4. Brainstorming process notes of Diana.

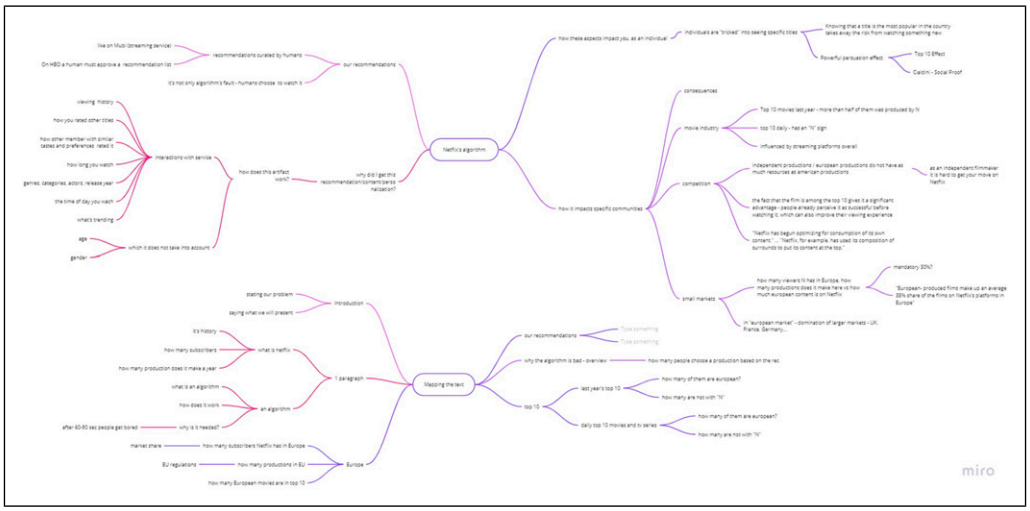

Image 5. A mind map of Grace and Zhenzhen to map all the things they know and do not know.

people but it is an algorithm deciding to serve it to me)', where she also already made some assumptions and initial sense-making.

In their journey for sense-making, they are also reflecting and questioning not just the algorithmic and datafication systems, but also themselves, often expressing emotional reactions: what did I do to receive this? Would I have been targeted differently if I did something differently, what do other people then receive as an algorithmic output? Why am I categorized as this? These emotional reactions are part of the sense-making process, and they merit attention, as Ytre-Arne and Moe (2020) outline, because they allow us to uncover the ambivalent feelings and contradicting experiences with datafication and algorithmic media (p. 811). We have noticed also that as the time passed and the students were getting more in more immersed into their investigative process, some of them were growing noticeably angry, discontent, and irritated – 'Let's have a closer look at how Instagram gets its money' (Marcel & Milagros).

From the initial artifact, through the process of uncovering the related and hidden aspects (see Images 4 and 5), the students arrived to various aspects they think merit attention: from the extent of the Chinese Communist Party's involvement is Huawei's working as a private company (Katarzyna & Hanna), to tracking the hidden third-party sharing connections between different actors in the datafication ecology (Image 6). Alex and Sara, for example, tried to track this by matching the contents published by @my.body.match with the apps installed on the same device (Image 7). They even went to look at the company's annual revenues (Alex & Sara), inspecting all of their previously published advertisements (Alex & Sara), to estimate how much the particular shown ad would cost (e.g., 'For our artifact that would mean the cost would be around US$3525 or €3324' – Katarzyna & Hanna).

While investigating their artifacts, they uncover and touch upon a variety of aspects and issues, like LGBTQ + rights, patriarchy and cultural traditions, workers' conditions in China, state-run propaganda, excessive datafication and third-party data sharing networks and actors, monopoly over markets and (content) producers (e.g., Facebook vs news and traditional advertising industry, Spotify vs independent artists), decolonization of beauty standards.

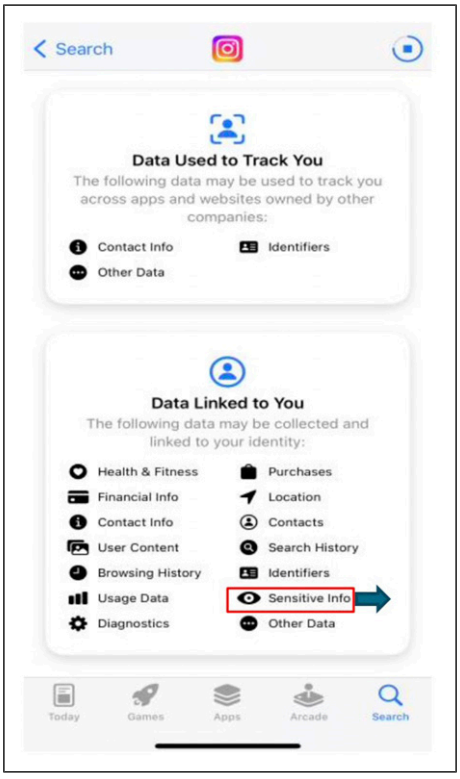

Image 6. Alex and Sara's screenshot of Instagram's data collection practices (iPhone settings).

Since they were learning theoretical and analytical concept throughout the semester, the students were also borrowing from and applying different theoretical concepts. They learned to recognize which framings, conceptualizations and discourses apply to what they are investigating. On the other hand, this helped them to improve and expand the sense-making and reflexivity processes, to start seeing the artifacts and events they are investigating from a different light. It also enabled them to better recognize and express what they are noticing and feeling. Most of the concepts were coming from media and communication studies, but also from critical data and algorithmic studies, including algorithmic justice, but also behavioral psychology, HCI, technology design and even political economy ('Why are homeless or poor people unable to find jobs? I think Marx and Bourdieu can explain these questions very well' – Erika).

Repairing what's not broken

Translating these various aspects into issues to be repaired, the students' Repair Manuals identified numerous issues that they see as problematic and that require an urgent intervention. These range from algorithmic amplification of polarizing content and fake news to recommender systems harms, polarizing hashtags, cross-platform tracking, social biases inscribed in algorithmic systems, to the politics of content personalization, Spotify's 30 seconds algorithms, racists and discriminatory beauty standards in search results, commodification of users, virality, echo chambers and filter bubbles, third-parties data sharing and processing, commodification and immaterial labor, Tinder's

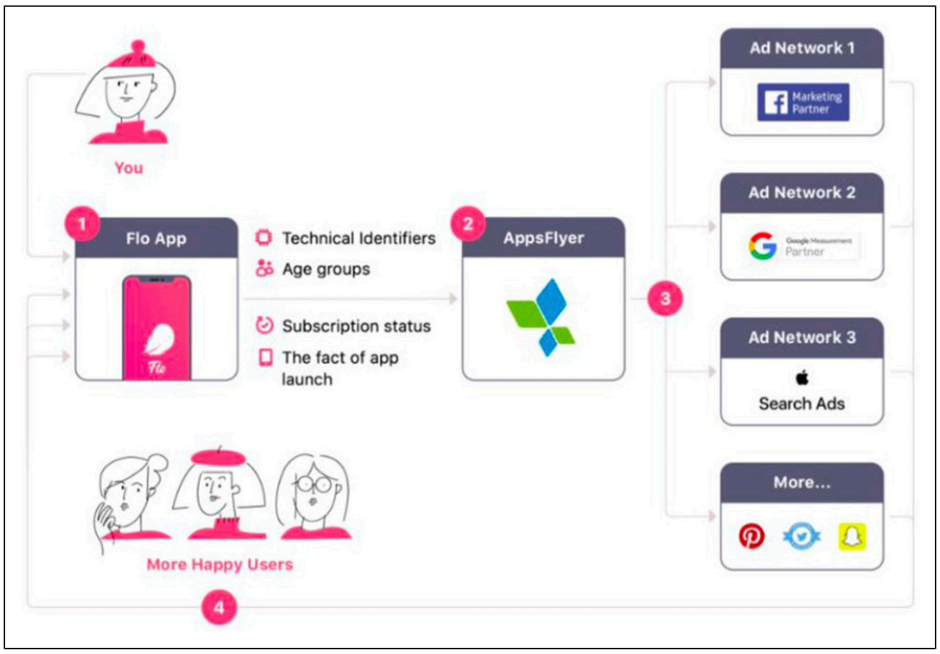

Image 7. Alex and Sara's investigation of Flo app's data collection and sharing policy (as in 2022).

Image 8. User-interface design repair solution for making sponsored ads more seamful (Katarzyna and Hanna).

Image 9. Alex and Sara's manually drawn algorithm repair solution.

monetization of desires and intimacy, homogenization and discrimination in personalized recommendations, propaganda by foreign actors, and others.

Their repair solutions were related equally to the algorithmic (the technical) and the societal (see Image 8 and Image 9). This might be an outcome of our instructions to look of how the technical is in the social and how the social is in the technical, and to look for connection between the two. They were also encouraged to find solutions that would be robust and consider the various actors and contexts that would be subjects of these algorithmic workings, but that also have power and authority to demand and execute the needed repairs. This is mirrored in the repair requests: repairs would have to be done to the algorithmic systems themselves, but also interventions would have to be undertaken by governments and other societal actors (e.g., regulatory and educational institutions). The manners in which these repairs could be achieved also vary: from building skills and literacy (e.g., foster skills like 'critical ignoring', creating civic education), to intervening in apps/platforms interface directly, to removing problematic content, to prioritization of reliable sources, to improvement of built-in feedback systems, to changing of the business model (e.g., to make Instagram a paid platform or to decommodify the platforms) and to the use of tactics like 'self-nudging'.

However, a noteworthy insight is that a number of students proposed interventions that would require interventions at individual level, often in terms of behavior change, skills and literacy building and taking privacy preserving actions (e.g., browsing anonymously). Often these repairs uncritically might lead to responsibilization – shifting the burden of responsibility to individuals. As the following examples illustrate, many of the repairs should repair user behavior – 'part of the solution to potential challenges lies with the user […] we can hope that the users themselves choose to act on their awareness' (Alexandra & Sebastian). Critical ignoring and critical data literacy are seen as user's responsibility and a repair solution – 'Ignore the advert and continue scrolling' (Irene). Or as Chira proposes – 'To achieve sustainable data protection, users remain the most

important actors to make this happen. Although platforms have the power to shape user activity, it is the users themselves who determine and influence what kind of content they upload, share, and are exposed to'.

This is a similar finding to the one of Ytre-Arne and Moe (2020) – although the students were instructed that their repair could be even a complete dismissal of the algorithm, not a single participant proposed this as a solution, and this offers interesting insights. As Lomborg and Kapsch's research shows (2019), they seem to experience algorithms as a 'necessary evil'. Markham (2021) borrows the notion of discursive closure to describe the situations in which 'people shut down alternatives themselves, naturalizing problems as "just the way things are"' (p. 392). This effectively discourages further change and 'close off options for thinking otherwise' (ibid.) and with it, removes human agency. Some of the repairs even propose improvement of the algorithmic processes – 'As such, this slight change will grant users who listen with greater attention to a song, a greater chance of being profiled more accurately and enjoy better services' (Andres & Elia).

Critical pedagogy and critical companionship

In conceptualizing our approach, we are guided by what Markham (2019, 2021) calls critical pedagogy – 'using methods of good teaching and learning (pedagogy) to raise critical consciousness about something' (p. 755), building further on Paolo Freire's (2000) work on pedagogy of the oppressed. This approach connects two understandings: the first one is that one learns best by foregrounding knowledge-making on personal experience (Markham, 2019: p. 756). The second is that the role of the educator, the pedagogue is to guide the students, to steer them towards sense- and knowledge-making, not by lecturing, but by letting them investigate, come to conclusions and form knowledge themselves, always within their own context and situatedness, on the basis of what they deem important. An approach like this acknowledges and nurtures the validity of non-academic knowledge, the lived experiences, and the practice-based knowledge. In the particular implementation developed by Markham (2019) for the field of datafication and algorithmic studies, it enables individuals to 'better understand their conditions and the hegemonic structures that enact them' (Pronzato and Markham, 2023: p. 101).

Critical pedagogy organizes the learning process in a manner that will allow individuals to come to the same insights and knowledge as researchers would, just their path for getting there would be different (Markham, 2019). This is knowledge-making that relies on incremental learning. That was also our aim with how we structured our assignment exercise. In the first phase, lacking the academic rigor, technological vocabularies (Swart, 2021), knowledge, and skills to properly execute a research task, our students were relying on their own experience and their folk theories. We call this first phase the 'guesswork' phase – where participants have more questions than answers, where knowledge and tools to look for that knowledge lack, and where we can see more the algorithmic imaginaries and folk theories at work. This is a messy, chaotic sense-making phase, driven more by reflection based on lived experience. That is visible also from the brainstorming notes our student participants provided us with (Image 4 and Image 5).

In the consecutive phases, the academic knowledge-making gradually takes over. As students are provided with more theoretical frameworks and concepts, and tools to investigate the issues at hand, their skills and capacities increase too. They look for more ways to find information and start a process of 'informed sense-making'. The reflexivity processes taking over here are more structured, guided by theoretical concepts and earlier academic research. This enables for a more critical and in-depth investigation and is grounded on recognizing one's own experiences in theoretical and more formal discourses.

At the end of the assignment, in the Repair manual, the individual, the lived, and the experiential meet – the proposed solutions are 'joint work' of what they have identified as individuals and non-experts, but now they are better equipped to tackle them in a structured and informed manner. Their own experience, thoughts, emotions, and affect are still prominent – after all, they started the investigations based on what touched them personally, but as Markham (2019) says 'they have discovered and trained within themselves the capacity to ask specific kinds of questions' (p. 760). To this we would add – and to answer them, too, now that they own some intellectual frameworks and research skills to do so.

Our role, as pedagogues, was to help our students become qualitative researchers while providing them with the theoretical and practical tools and skills sets, by giving them the opportunity to work on something they find important for themselves and guiding them along the way. We helped them make the transition from distant theoretical ideas and abstract black-boxes of algorithmic systems to the tangible and investigable objects of studies. In that regard, we could define our research and pedagogical strategy as one of critical companionship, too (Ziewitz and Singh, 2021).

A distinctive trait of the critical companionship approach is that an experienced practitioner (educators, in our case) guides and accompanies a less experienced one on an experiential learning journey. This learning journey, however, is a (joint) learning journey for the educators too (Ziewitz and Singh, 2021) – both us and our students are trying to make sense. Our students of the datafication processes they are entangled it and the algorithmic artifacts they are encountering, and us, the educators, of the experiences, questions, and guidance needed by our participants. We are skilled and equipped to offer knowledge and skills transfer and support, but we also learned along the way.

Another significant aspect of critical companionship is that – along the process – it blurs the boundaries between the researcher and participants. This is not just because both go through learning and transformative processes, but also because our participants, also became (non-expert) researchers (Ziewitz and Singh, 2021). Our students underwent a process of observation, inquiries, of conducting their own investigations, of acquiring expertise through incremental learning, and through different phases of attempting to make sense and develop (own) theories.

We believe that this exercise, and the reflections, sense-making and knowledge-making processes contributed also to empowerment of our participants not just as future researchers, but as individuals too. We are aware that, as Ytre-Arne and Moe (2020) say, that 'to interpret is not the same as acting against or escaping datafication' (p.820) and algorithmic workings, but it is still not just a form of critique, but also productive force to build and adopt resistance and repair practices. In that regard, the radical uncertainty of knowing, with which our students started this exercise, has not been fully diminished. There are limitations to these reflexive practices, and that is not only because our students lack knowledge and skills. The way algorithmic systems are being built and deployed, the maneuvering space between knowing and being able to act, between agency, resistance, and repair, is shrinking.

Reflections

In this section, we will discuss a number of insights and issues that we have identified during the course of the study, and after. What we did not do in this article, however, is analyze the particular issues or the repair solutions proposed. That would be a material for a further article. For us, this was a research exercise and a pedagogical process. We will reflect on these two separately, discussing the insights, the potentials, and the limitations of an approach like this.

When it comes to assessing the suitability of an approach like this for research, we believe it could be a productive one for gathering insights on individuals' experiences and sense-making

processes. It could be used also as a research tool to think about, reflect on, and investigate what individuals find both important and problematic, even after (academic) insights and knowledge have been acquired. This approach could be also used to see how existing datafication and algorithmic systems could be improved to work better for individual's interests, public interest values and human rights.

The repair manuals, on the other hand, offer valuable insights into how the individuals would like these systems to be. They could give specific points for improvement, identify the critical points, provide insights how user agency is envisioned and what forms it can take. Both in terms of design, business practices, but also transparency, explainability, and accountability. In that sense, it could be useful for technology designers and practitioners, and for regulators for policymaking, evaluation, and enforcement.

When it comes to further use of this approach, for either research or pedagogical purposes, it needs to be mentioned that the execution, the outputs, and the outcomes of an exercise like this will significantly differ depending on who the participants are – students or common users. Our participants were students, and the exercise was done in an educational setting, with clear expectations for a learning outcome. They had the time (one semester) to think, reflect, investigate, go through iterations, receive feedback, be presented with concepts and theories, so they could form subject-specific knowledge. As students, they held two positions simultaneously: as non-expert individuals, and as researchers-in-the-making. They were non-expert users in the sense that they followed this course and did this exercise at the beginning of their studies. Choosing this subject and assignment, it is probable that they were interested in the topic and potentially had some prior knowledge but were lacking proper research skills and tools to investigate datafication and algorithmic systems. As researchers-in-the-making, they gradually arrived there – the exercise (along with the course itself) was aiming at teaching them how to approach, think, and critically investigate these systems, and to provide with a toolset to do that. The process- and hands on-oriented teaching was reinforced by the iterative nature of the process. This also allowed our participants to become co-researchers too.

There are some constraints and practical and conceptual challenges if this approach is done with non-experts that should be taken into consideration. The first challenge is related with time – unlike students, it might be hard to recruit and mobilize participants for longer periods of time. This could be solved by organizing a smaller scale, one day setup. However, this leads to another challenge – the quality of the participants' engagement and thus of the expected outputs. Since there will be no time for thorough reflection and investigation, and for skills-, technical vocabulary- and knowledge-making, the scope and depth of the insights (for participants and researchers) might be diminished.

This means that the approach should be downscaled if used in a non-educational setting. This could be done, for example, by letting participants focus on only one aspect they find 'problematic'. Starting from the individual in the artifact analysis phase, they could do the repair phase working in groups. Each group could be guided by the researchers who should provide and facilitate knowledge (-making) and vocabulary, taking on the role of both educators and (workshop) facilitators. While knowledge and sense-making processes do take time, equipping the participants with critical knowledge might be challenging, but not impossible (see e.g., Bishop and Kant, 2023).

The benefit of an exercise like this for the non-expert participants lies in the potential for their empowerment through fostering of critical knowledge, in the sense of giving them a toolkit to conduct independent critical investigations into algorithmic artifacts and their underlying datafication nature. It would also encourage them to further look critically and dissect the artifacts they encounter along the way. As such, it could also lead to a critical consciousness-rising, that would

enable them to participate more equally in public discourses on algorithms and their societal impacts.

If the results of an exercise like this are used for research purposes, we cannot expect that all the outputs from the students would be of same quality and will be equally useful and valuable. Students, of course, come with different drives, capacities, time, effort, and dedication put into the assignments. If insights are expected that should inform later research, additional guidance will be required. For us, since this was a graded assignment, we were balancing between guiding them supportively and letting them work individually.

Conclusion

In this article we have developed a new methodological venue and a pedagogical approach for reflectively engaging with algorithmic data systems. We were guided by the attempts to overcome the challenges imposed by opaque algorithmic systems, where algorithms are not study-ready. In addition, making sense of them and their impact on individuals and society proves difficult for experts, but even more so for non-expert citizens. Without needing to open the algorithmic black-box, we devised an approach for engaging with these systems by *thinking with*, – *through* and – *about* algorithmic artifacts.

We start from the premise that algorithms are always broken, since they always work for some, but not for everyone. They hold social power and they encompass societal power imbalances and in-equalities that are perpetuated via these systems. The understanding of algorithms as broken implies that they can be repaired. But each repair first requires a diagnosis, and each diagnosis requires knowledge and skills. Borrowing from the approaches developed by Shannon Mattern and Joseph Dumit, we take algorithmic data artifacts as entry points to start the interrelated and crucial processes of reflectivity and reflexivity, as crucial for knowledge and sense-making and ultimately, for repair.

The processes and insights we presented show that an approach like this can be useful for both researchers and educators. In our discussion, we pointed to the potential points of attention and the limitations of the study, especially if executed in a non-educational context. This type of approach can be useful outside of academia and future research could focus potentially on the suitability within non-expert and communities of practice, like activists and policy makers.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Fonds Wetenschappelijk Onderzoek; G054919N.

ORCID iD


Ana Pop Stefanija https://orcid.org/0000-0002-5700-602X

Note

1. Names of the participants have been changed, and a random names generator was used so they can be anonymized.

References

Ananny M and Crawford K (2018) Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society* 20(3): 973–989.

Angwin J, Larson J, Mattu S, et al. (2016) Machine bias. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (accessed 22 November 2023).

Aradau C and Blanke T (2022) *Algorithmic Reason: The New Government of Self and Other*. Oxford, New York: Oxford University Press.

Barabas C, Doyle C, Rubinovitz J, et al. (2020) Studying up: reorienting the study of algorithmic fairness around issues of power. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, New York, NY, USA, 27 January 2020.

Beer D (2017) The social power of algorithms. *Information, Communication & Society* 20: 1–13.

Benjamin G (2022) FuckTheAlgorithm: algorithmic imaginaries and political resistance. In: 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul Republic of Korea, 21 June 2022. Available at: https://dl.acm.org/doi/10.1145/3531146.3533072 (accessed 24 June 2022).

Bishop S and Kant T (2023) Algorithmic autobiographies and fictions: a digital method. *The Sociological Review*. Thousand Oaks, CA: Sage Publications Ltd.

Büchi M, Fosch-Villaronga E, Lutz C, et al. (2023) Making sense of algorithmic profiling: user perceptions on Facebook. *Information, Communication & Society* 26: 809–825.

Bucher T (2017) The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society* 20(1): 30–44.

Bucher T (2018) If...Then: algorithmic power and politics. In: *Oxford Studies in Digital Politics*. Oxford, New York: Oxford University Press.

Buolamwini J and Gebru T (2018) Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, New York University, NYC , February 23 and 24 2018, pp. 77–91. PMLR. Available at: https://proceedings.mlr.press/v81/buolamwini18a.html (accessed 7 March 2024).

Cheney-Lippold J (2011) A new algorithmic identity: soft biopolitics and the modulation of control*Theory, Culture & Society* 28: 164–181.

Colmellere C (2016) Repair in socio-technical systems. The repair of a machine breakdown that turned into the repair of a shop. *Italian Journal of Science & Technology Studies* 6(2).

Constantaras E, Geiger G, Braun J-K, et al. (2023) Inside the suspicion machine. Available at: https://www.wired.com/story/welfare-state-algorithms/(accessed 22 November 2023).

Cotter KM (2020) *Critical Algorithmic Literacy: Power, Epistemology, and Platforms*. East Lansing, MI: Michigan State University.

Couldry N and Mejias UA (2019) *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating it for Capitalism*. Stanford: Stanford University Press.

Crawford K and Paglen T (2019) Excavating AI. Available at: https://www.excavating.ai (accessed 26 November 2019).

Crawford K and Paglen T (2021) Excavating AI: the politics of images in machine learning training sets. *AI & Society* 36(4): 1105–1116.

Delacroix S and Veale M (2020) Smart technologies and our sense of self: going beyond epistemic counter-profiling *Life and the Law in the Era of Data-Driven Agency*. Cheltenham Glos, UK: Edward Elgar Publishing, 80–99. Available at: https://www.elgaronline.com/view/edcoll/9781788971997/9781788971997.00011.xml (accessed 13 January 2022).

Dumit J (2014) Writing the implosion: teaching the world one thing at a time. *Cultural Anthropology* 29(2): 344–362.

D'Ignazio C and Bhargava R (2016) DataBasic: design principles, tools and activities for data literacy learners. *Journal of Community Informatics* 12(3): 3.

Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: Martin's Publishing Group.

Fisher E (2020) The ledger and the diary: algorithmic knowledge and subjectivity. *Continuum* 34: 378–397.

Freire P (2000) *Pedagogy of the Oppressed*. 30th Anniversary Edition. London, Bloomsbury Academic.

Gitelman L (ed) (2013) *Raw Data Is an Oxymoron*. The MIT Press.

Graham S and Thrift N (2007) Out of order: understanding repair and maintenance. *Theory, Culture & Society* 24: 1–25.

Grosman J and Reigeluth T (2019) Perspectives on algorithmic normativities: engineers, objects, activities. *Big Data & Society*. SAGE Publications Ltd, 6(2), 1–12.

Hildebrandt M (2022) The issue of proxies and choice architectures. Why EU law matters for recommender systems. *Frontiers in Artificial Intelligence* 5: 1–17.

Iliadis A and Russo F (2016) Critical data studies: An introduction. *Big Data & Society*. SAGE Publications Ltd, 3(2), 1–7.

Jackson SJ (2014) Rethinking repair. In: T Gillespie, PJ Boczkowski, and KA Foot (eds) *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9780262525374.003.0011.

Kapsch PH (2022) Exploring user agency and small acts of algorithm engagement in everyday media use. *Media International Australia* 183: 16–29.

Kitchin R (2017) Thinking critically about and researching algorithms. *Information, Communication & Society* 20, 14–29.

Lehtiniemi T (2016) Reorienting datafication? New roles for users in online platform markets. *The Internet, Policy & Politics Conferences* 2016, 1–16.

Lindgren S (ed) (2023) *Handbook of Critical Studies of Artificial Intelligence*. Cheltenham Glos, UK: Edward Elgar Publishing. Available at: https://www.e-elgar.com/shop/gbp/handbook-of-critical-studies-of-artificial-intelligence-9781803928555.html (accessed 4 September 2023).

Lomborg S and Kapsch PH (2020) Decoding algorithms. *Media, Culture & Society* 42: 745–761.

Lupton D (2017) Feeling your data: touch and making sense of personal digital data. *New Media & Society* 19: 1599–1614.

Lupton D (2021) 'Not the real me': social imaginaries of personal data profiling. *Cultural Sociology* 15: 3–21.

Markham A (2019) Critical pedagogy as a response to datafication. *Qualitative Inquiry* 25: 754–760.

Markham A (2021) The limits of the imaginary: challenges to intervening in future speculations of memory, data, and algorithms. *New Media & Society* 23: 382–405.

Mattern S (2022) Requirements + assignments – redesigning the Academy 2022 Redesigning the Academy. Available at: https://redesigningacademy.wordsinspace.net/spring2022/requirements-assignments/ (accessed 14 November 2023).

McQuillan D (2022) *Resisting AI: An Anti-fascist Approach to Artificial Intelligence*. Bristol: Policy Press. Available at: https://academic.oup.com/policy-press-scholarship-online/book/45148 (accessed 21 November 2023).

Meuser M and Nagel U (2009) The expert interview and changes in knowledge production. In: A Bogner, B Littig, and W Menz (eds) *Interviewing Experts*. London: Palgrave MacMillan, 17–43. Available at: https://tertulia.com/book/interviewing-experts-a-bogner/9780230220195 (accessed 19 May 2023).

Miceli M, Posada J, and Yang T (2022) Studying up machine learning data: why talk about bias when we mean power? *Proceedings of the ACM on Human-Computer Interaction* 34: 1–14.

O'Neil C (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. 1st edition. New York: Crown.

Pink S (2022) Methods for researching automated futures. *Qualitative Inquiry* 28: 747–753.

Pop Stefanija A and Pierson J (2020) Practical AI transparency: revealing datafication and algorithmic identities. *Journal of Digital Social Research* 2(3): 84–125.

Pop Stefanija A and Pierson J (2023) Algorithmic governmentality, digital sovereignty, and agency affordances: extending the possible fields of action. *Weizenbaum Journal of the Digital Society* 3(2): 2.

Pronzato R and Markham AN (2023) Returning to critical pedagogy in a world of datafication. *Convergence* 29: 97–115.

Redesigning the Academy (2022) Our final projects – redesigning the Academy 2022. Available at: https://redesigningacademy.wordsinspace.net/spring2022/our-final-projects/ (accessed 23 November 2023).

Rieder B, Gordon G, and Sileno G (2022) Mapping value(s) in AI: methodological directions for examining normativity in complex technical systems. *Sociologica* 16(3: 51–83.

Ruckenstein M and Granroth J (2019) Algorithms, advertising and the intimacy of surveillance. *Journal of Cultural Economy* 13(1): 12–24.

Seberger JS and Bowker GC (2021) Humanistic infrastructure studies: hyper-functionality and the experience of the absurd. *Information, Communication & Society.* Routledge, 24(12), 1712–1727.

Suchman L (2006) Plans, scripts, and other ordering devices. In: *Learning in Doing: Social, Cognitive and Computational Perspectives*. 2nd edition. Cambridge: Cambridge University Press, pp. 187–205. Available at: https://www.cambridge.org/core/books/humanmachine-reconfigurations/plans-scripts-and-other-ordering-devices/F82500B3B78E0C3D6A04FB9906D11853 (accessed 25 September 2023).

Swart J (2021) Experiencing algorithms: how young people understand, feel about, and engage with algorithmic news selection on social media. *Social Media + Society* 7(2): 20563051211008828.

Sweeney L (2013) Discrimination in online ad delivery. *Communications of the ACM* 56(5): 44–54.

van Dijck J (2014) Datafication, dataism and dataveillance: big Data between scientific paradigm and ideology. *Surveillance and Society* 12(2): 197–208.

Velkova J and Kaun A (2021) Algorithmic resistance: media practices and the politics of repair. *Information, Communication & Society* 24: 523–540.

Wachter S and Mittelstadt B (2018) A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *SSRN Scholarly Paper, 5 October*. Rochester, NY: Social Science Research Network. Available at: https://papers.ssrn.com/abstract=3248829 (accessed 22 May 2019).

Weiskopf R and Hansen HK (2022) Algorithmic governmentality and the space of ethics: examples from 'people analytics. In: *Human Relations*. Thousand Oaks, CA: Sage Publications Ltd.

Yin L and Sankin A (2020) Google ad portal equated "black girls" with porn – the markup. Available at: https://themarkup.org/google-the-giant/2020/07/23/google-advertising-keywords-black-girls (accessed 22 November 2023).

Ytre-Arne B and Moe H (2021) Folk theories of algorithms: understanding digital irritation. *Media, Culture & Society* 43: 807–824.

Ziewitz M and Singh R (2021) Critical companionship: some sensibilities for studying the lived experience of data subjects. *Big Data & Society* 8: 20539517211061122.