Hermes. Joke; Natalia Berger

Friend and Foe. An Analysis of Expert Advice on Educating for Data Literacy in Journalism and the Creative Industries

2024

https://doi.org/10.25969/mediarep/23181

Veröffentlichungsversion / published version

Zeitschriftenartikel / journal article

volume 13 issue 25/2024

VIEW

JOURNAL OF EUROPEAN TELEVISION HISTORY & CULTURE

FRIEND AND FOE

AN ANALYSIS OF EXPERT ADVICE ON EDUCATING FOR DATA LITERACY IN JOURNALISM AND THE CREATIVE INDUSTRIES

Joke Hermes
Inholland University of Applied Sciences
joke.hermes@inholland.nl

Natalia Berger
Inholland University of Applied Sciences
natalia.berger@inholland.nl

Abstract: Analysis of practical advice in regard of developing a data literacy course for students in higher education and for early-career professionals in interviews with 56 experts in data analysis, offered an unexpected shared story among the experts around the perception, affect and power of data. By attending to what data are felt to be and do to you unintentionally, the analysis delivered a new story of how to teach for data literacy. Rather than (only) focus on formal skills, it points to how building a relationship with data may help more students and young professionals become curious, critical and engaged while recognizing the strengths and the fickleness of data sets.

Keywords: MediaNumeric, Data literacy, Perception of data, Affect of data, Power of data

1 Introducing Our Educating for Data Literacy Project

Data and data-based multimedia materials have come to play an increasingly crucial role in journalism and other forms of public storytelling. The fast development and introduction of generative AI is a case in point. The credible texts across media genres, the video segments and audio fragments in the voices of existing persons who utter sentences they never themselves have spoken, boggle the mind. How to educate students in journalism and the creative industries to use data well if this is what data can do? Just before the introduction of easy-to-use AI applications, such as ChatGPT-3 (and the resulting wide-spread public discussion), we conducted a needs analysis for data literacy education among a broad group of experts in news and media production-related fields. The needs analysis was part of a European-funded project to build an open-access course to train future public storytellers (whether in journalism or in other types of content production) in data literacy called MediaNumeric[1]. This article will focus on how teaching data literacy needs attending to (1) how data are perceived, (2) how data (singular) makes people feel (their affect) and (3) what the power of data is.

Using data in the singular has become customary. 'Data' is then understood as referring to digitally stored information, a mass noun, the individual elements in which are beyond counting. When used in the sciences, data refers to groups of single observations that lead to e.g., a data set. It is used in the plural. As will become clear from the interview excerpts below, those who work with data shift between the everyday and the academic use of the term. Understanding that even mass data start with individual bits of observed reality helps in gaining a sense of where data come from and how they may be more -or for that matter less- correct depending upon one's point of view. As a well-known example in journalism education goes: one person's freedom fight is another's terrorism. In this article we are interested in how to educate students and early-career professionals in media professions to understand and work with data as mass; where even experts wrestle with backtracking to where the individual elements that make up the mass originated. Where, as media users, we have to take it on trust that content producers know where their data comes from and have used it in good faith rather than bad and where data can turn into any of the forms of misinformation.[2]

The gap between what data (as observations) are and what data (as a mass of datapoints that only computers can handle) is, suggests how educating for data literacy might be a complicated project. Not only because computing requires specialized competencies but because we need to understand how data (both in the singular and the plural) relate to how humans describe other humans and how we, in turn, relate to being able to grasp (or be bewildered) by what is going on there. Data are not 'mere' description, they are assembled with specific goals for their usage in mind from specific perspectives. Data, as used by computers, needs to be treated with care. This article will present the slice of the MediaNumeric project in which we sought to find out what might make newcomers take to data and data analysis (or feel distanced from it).

2 The Need for Data Literacy

Data literacy in general is predicted to be the most in-demand employee skill by 2030.[3] According to business magazine *Forbes*:

> The value of data literacy is continuing to increase in companies all over the world. (-) 70% of employees are expected to work heavily with data by 2025 — up from just 40% in 2018. It's clear that organizations that build data literacy into their workplace culture will gain a competitive edge in our increasingly complex global economy.[4]

Journalists, given this prediction, will need a skill set to allow them to reverse engineer how data is used, whether in business, policy development or elsewhere. Other content makers too will want to be able to engage in these relatively new ways of understanding the world. Beyond having to be intelligent and responsible storytellers, like journalists in 'hard news' media, they now need to know when and how data matters, what kind of 'facts' come with data sets and how to work with the specialists that can unlock data at advanced levels. We are, in addition, living in an age of synthetic media and 'deep fakes', a broad intervention not just in journalism education but in the professional training for the creative industries education is therefore of the utmost urgency.[5] We need new models of thinking what can count as 'truth'. Before we can find those, such an intervention has the unappealing task to counter what Dutch philosopher and 'slow journalism' inventor Wijnberg typifies as 'cynical sensationalism'.[6] Out to attract viewers, readers or users, a majority of media forms will lean towards stories and storytelling that make an impression. To counter the depletion of value and trust that are the result of such strategies, we need to invest in sincere curiosity and in understanding truth finding as a never-finalized collective effort, he argues.

Separate from attending to how data is increasingly used for decision making, data itself is in need of critical understanding. Feminist data science warns us that data is more likely to be sexist, racist and so on than attentive to long ongoing historical exclusions given older regimes of data collection.[7] Even though creatives might hesitate to

expand, pivot or develop their data-driven practice, it is important that those not attracted to data in and of itself, take on the collective responsibility to use and improve data-based practices.[8]

3 Research Question and Method

The interviews this article is based on were collected to answer the very broad question of what needed to be part of a higher-education course in data literacy.[9] In 2021-2022, members of the Erasmus+ MediaNumeric consortium, interviewed 56 experts on data literacy.[10] They were interviewed on the practical challenges, obstacles, and needs in making others more familiar with using data well. Our expert informants are based in the Netherlands (16), France (11), Poland (9), the USA (6), as well as in the UK, Spain, Denmark, Sweden, Belgium, Austria. They are an even mix of journalists and editors, archivists, fact checkers and data specialists, and academic researchers and teachers. They represent media enterprises, media archives, higher education, NGOs and fact-checking organizations.

The interviewers asked them about the must-have skills, knowledge, and tools for data work in their field (1); whether they could share their best and worst experiences with data work (2) and how data skills should be taught in higher education (3)? The interviews, ranging from 45 to 90 minutes, were recorded and transcribed in full, resulting in an average of 25 A4 pages per transcript and a total data set of approximately 1,500 pages. Inspired by grounded theory, MaxQDA qualitative data analysis software was used to reduce this expansive textual data set.[11] Three coders employed open coding techniques to form a draft codebook by blind-coding each other's transcripts and agreeing on codes to be shared after.

Regardless of whether experts represented creative or technical know-how, they agreed that effective and responsible data-driven storytelling needs foundational skills. Journalistic and storytelling skills are crucial for composing compelling narratives. Secondly, ethical considerations are equally crucial in the handling and representation of data. Thirdly, basic statistical knowledge and data manipulation skills, including understanding data collection, cleaning, analysis, and distinguishing between correlation and causality, are imperative for accurate and insightful data interpretation. Lastly, critical thinking was emphasized as a fundamental skill. Practitioners should have a rigorous and questioning approach to data and its implications and be able to question, analyze, and debunk misinformation.

Open coding in grounded theory research allows going beyond a descriptive summary such as this, however informative and useful. After open coding, in what are called axial and selective coding, themes that cut across the material can be identified. We found that emotional responses to data were surprisingly often discussed. Our informants observed such responses in their places of work, experienced them themselves and took them for granted. After taking a closer look at our axial codes, we isolated three selective codes. Those are codes that in addition to cutting across the material represent a particular perspective. We decided to follow up on whether the selective codes of perception, affect and power together would give us a better understanding of what makes working with data scary and what makes it exciting. Both were referenced in the interview material. In grounded theory research, selective codes acquire persuasive impact by extensive quoting from the interview material which is offered as proof for the emerging theory the researchers present. In addition, the research literature will be referenced where appropriate for further validation. Below, therefore, space has been allowed generously for what our experts had to say.

Our analytic intervention as researchers has been to reconstruct what we see as the underlying *discursive logic* in the material. Rather than analyze individual quotes, we sought to identify the shared cultural knowledge of a diverse group of insiders, bound by the professional reputation they are accorded for their work with and on behalf of data analysis and data literacy. The material is, therefore, presented as what anthropologists Marcus and Fischer have termed a polylogue.[12] We have included our experts' job titles in the ongoing narrative to give a better sense of the material and to ground the validity of our argument. Aiming for a light form of theorization, we hope to make the outcomes of our research more easily transferable than, e.g., a more descriptive account would be.

As is usual in qualitative research, our research question gradually focused. From the broad question: how to educate students in journalism and the creative industries to use data well, we moved to how feelings might matter in working with data and might therefore be relevant to data literacy training. Emotion was often in adjectives and asides: feeling 'weird about data', pessimistic or cautiously optimistic, finding data overwhelming, daunting, threatening but also powerful and delicious. After checking back with the original codes, three entry points presented themselves to discuss the material from this new vantage point. As mentioned, they are perception, affect and power. We will discuss them below.

In a sense, asking this new question of our material is a form of secondary qualitative analysis.[13] This involves re-examining existing interview data originally collected for a different purpose. It allowed us to derive valuable, unbiased insights, as no direct questions about the emotional difficulties in acquiring data literacy skills were asked initially when interviewees were approached for advice on building an open-access data literacy course for students in higher education.

4 Key Theoretical Concepts

Above we briefly discussed why data is/are both referred to in the singular and in the plural and how that is a meaningful distinction. In this section we will briefly introduce key theoretical notions as they will be used below where we discuss our analysis and conclusions.

Data in the singular is also referred to as 'big data,' which boyd and Crawford define as a phenomenon that rests on the interplay of technology, analysis and mythology. It verges around the ability to search, aggregate and cross-reference massive quantities of information produced by and about people, things, and their interactions. They warn that we need to wonder whether

> large scale search data help us create better tools, services, and public goods? Or will it usher in a new way of privacy incursions and invasive marketing? Will data analytics help us understand online communities and political movements? Or will analytics be used to track protesters and suppress speech?[14]

It is data in the singular, or 'big data' that concerns us here.

Data literacy, in relation, is the ability to effectively understand, use, and argue with these large data sets that is essential for both decision-making and storytelling.[15] The ability to critically reflect upon collections of data we use is crucial to data literacy. It empowers creative industries workers and citizens alike.[16] Data literacy then encompasses the ability to comprehend, manipulate and assess data in order to make decisions. It involves understanding and critically reflecting upon the practices of collecting and utilizing data and recognizing associated risks. This form of literacy empowers individuals to effectively engage in discussions supported by data and apply their knowledge for well-informed internet (or platform-based application) usage. For us, data literacy is a form of critical thinking. It is both a crucial antidote to 'fake news' (or disinformation) and a necessary skill for those doing creative and artistic work using data and AI tools in order not to reassert dominant representational codes that associate gender, race, sexuality, ability and so on with negative qualities.

Perception refers to how we form images of independently existing reality.[17] Perception as a selective code was built on such open codes as 'utility,' 'complexity', 'assessability' and 'privacy concerns,'. They point to how the way an image is formed of data is a crucial component in how data is (or is not) felt to be understandable and manageable by individuals.

Affect denotes intersubjective energy.[18] It focuses on how the world, or of course data, impacts us as human beings. How, literally, we are moved. Wetherell has suggested that in social sciences and humanities research, 'affective practice' is the more useful term to use. It shows

> the emotional as it appears in social life and tries to follow what participants do. It finds shifting, flexible and often over-determined figurations rather than simple lines of causation, character types and neat emotion categories.[19]

Hill adds that for Wetherell:

> affective practice includes the meaning of practice as forms of order, habits, expression of emotions and actions, and the meaning of practice as what could be otherwise, thus how we might imagine ways of being, or alternative relations.[20]

Affect, then, allows as much for emotion as for the imagination. Leys supports such an interpretation in her review of affect across a range of scientific fields. For her the emotional dimension that exists subconsciously stands out, grounded as it is in bodily responses, which Leys calls 'anti-intentional'. That makes affect a type of reaction that is not styled or occurs from manipulative motives.[21]

In our material, the intersection of **power** and knowledge is also important and notably how data might be used as a means to wield power against others.[22] Data has become a currency of power, Catherine D'Ignazio argues. She is an advocate of data literacy for people in non-technical fields to conquer inequity "in terms of who has access to computer power and know-how to be able to make sense of data".[23] Knowledge is never created in isolation nor is power solely a top-down force.[24] How knowledge is shaped and felt to be accessible or ownable is deeply influenced by existing power dynamics and social power relations. Given their entanglement, the empowering potential of knowledge in the context of data and the ways in which access to knowledge about and insight into data excludes, becomes an important facet to explore.

Below, we will turn to the analysis of our interview material to show how efforts to make data literacy more accessible and more encouraging, may well depend on recognizing how specific perceptions, forms of affect and power relations shape data literacy as an object of knowledge and an asset for professional media making.

5 Analyzing Data Literacy

The analysis section is organized around the three key concepts discussed above. As is usual in a grounded theory approach, we offer a discussion of our interpretation of the data that consists of presenting relevant quotes paired with insights from the research literature. The goal is to develop a theory-based, well-theorized answer to our question, which is: how to understand what data feel like, in order to build as strong as possible a higher education course on data literacy and working with data for students and early-career professionals in the creative industries.

5.1 … In Terms of the Perception of Data

Counter-intuitively, data should be seen as '*human,*' we heard – as being about actual living beings, which comes with a responsibility for data analysts:

> [...] because, at the end of the day, a lot of data sets are about humans and human lives and how human relationships are. And I think because of that, I do feel kind of weird about automating that; the idea that there is no human touch in the data collection process when what we're trying to tell stories are about humans. Maybe in five years I will have changed my mind and now I'm going to just have all of the machines do it for

> me, but as of right now with the state of the technology as it is, I think I'm cautiously optimistic about where it could be in the future. (Independent creator of data visualizations, freelance, USA)

This quote addresses how we should perceive data as human-made. It also contains a first lesson about how to teach digital literacy beyond the formal requirements we summarized above (good storytelling, statistical competency and so on). Digital data is neither simply 'true' nor untrue. Although it consists of ones and zeroes that machines can work with, there is always a 'human-in-the-loop'.[25] The research literature points out that this makes data doubly prone to bias: the original material collected may well be biased and then machine learning protocols may not be geared to correcting or even signaling this. As 'the tools of big data research are increasingly woven into our daily lives', those who use them should always be aware of the potential impact on people's lives and the possible harm caused by data.[26]

Teaching data literacy, this ethical component of how data are produced, used, and presented, and the impact they may have, is crucial. Huff's 1954 *How to lie with statistics* should be required reading for all concerned with how digital data and digital presence easily compromise the agency of vulnerable groups, among them women but also LGBTQI+ teenagers, for example.[27] Huff's main premise continues to hold true.

Mejias and Couldry also address the transformative impact of data and data-driven practices on human life as a focal point of concern for academics, practitioners, and policymakers.[28] This is not simply a question of personal data usage and the dangers that presents to individuals. There is the more general argument that perceiving data in terms of 'objectivity' or 'accuracy' is deeply misleading.[29] Data is about human beings and made by them: data is always collected following human-made rules and an object of interpretation that is not meaningful without human intervention (including all the biases that come with such interventions).[30]

Not only, then, does working with data come with great responsibility, it quickly becomes complex and effort-intensive once you are out of the pre-set tasks of the classroom. Across the expert interviews, there is, therefore, the advice to make collaboration an integral element of a data literacy course by, for example, having newsrooms and students cooperate. Teamwork can be a short route to changing one's perception of data: it is far easier to enjoy working with data as an intriguing collective puzzle. It also shortens a process of discovery that can be long:

> That's how I discovered data, because in sports, economics and ecology, I was confronted with data and, I must admit, I didn't know what to do with it. I had been taught to interview people, [observe] situations, to do reports, but I didn't know how to *interview numbers*, so I sort of started out on my own, asking myself the following questions: What does it mean? Can I trust this figure? What's the source? And to not just copy and paste press releases. I had more of a scientific background in high school, so I had an appetite for it, but during my years of political science, I had completely skipped those parts. (Co-founder dataviz platform, FR)

Here is another informant for whom data became a 'delicious' tool, a promise of intellectual discovery. As in the previous quote, there is a mix of expectations (which pertain to how data is perceived) and feelings (related to affect, to which we will return below):

> If you look at it like that, the really kind of delicious data moments though are where you get literally spreadsheets with numbers on it. And you *interrogate* that in a way that creates something that can illuminate an issue. That's when it gets really powerful. (Editor of national broadcast company, UK)

The feeling of empowerment mentioned here will reappear later in our discussion of both affect and power, it starts though with a perception of data as an invitation to engage.

Such engagement may require hard work:

> I think that there is also a great fantasy around data, in the world of big data, which tends to lead people to believe that all you have to do is retrieve a file, press three keys and lines of code appear, etc. But it doesn't

> work like that at all. It's a job where the human element is very important. It's not the machine that does everything, it's the human who configures it, who checks, analyses and interprets the data, etc. (Co-founder dataviz platform, FR)

Working with data, says the below quote, can also be exhausting, especially when there is no organizational support. It presents a perception of data as costly specialist work (whether or not as part of a professional team) that again links through to affect (on which more below).

> I think there is a fear of engaging in this field because there is a lack of data culture among editors-in-chief. They don't really understand what it is — they only see the time-consuming aspect. It requires work from a lot of people, it's very expensive, and they ignore the editorial aspect that could never have been produced without all that work. (Co-founder dataviz platform, FR)

5.2 … In Terms of the Affect of Data

Perceiving data as delicious bespeaks of a form of affect. The stronger reference to how data affects people, though, speaks of fear. Some of our experts from the field of journalism pointed to how fear is very much part of their own and their colleagues' relationships with data. Data can be difficult, time-consuming, and unpredictable, they said. It is capricious and requires care and time; one cannot always depend on it to deliver. Data's undependability upsets the fantasy of the powerful machine that otherwise is also part of how data (and working with data) is perceived.

> I think because big data sets have so much information, there is a big chance that I might get really overwhelmed or side-tracked, so it's really helpful for me to have a goal and an intended audience so that I can focus my data exploration. (Independent creator of data visualizations, freelance, USA)

Fear of data, it was suggested, could be a generational issue, not among those working with data but their managers:

> I think that there is also perhaps a generational issue and that until we have a generation at the head of newspapers that 'got their hands dirty' with data in order to really understand what it is and to master it from an intellectual and practical point of view, data will remain this sort of 'black box' that scares people. (Co-founder of dataViz agency, FR)

We are not sure that a new generation of managers will necessarily deal more easily with data, especially if data remains the realm of what is called a 'small class of specialists' who know how to work with data effectively.[31] A creative producer from Poland elaborates:

> It is pretty abstract for me (...). I would like to have a technologist in my institution. I would like to remain the person who does not get her head around it.

'Fear of data' was mentioned in almost all interviews. Experts, however, recognized but were not greatly interested in general fear of numbers or statistics. Taking a closer look at where they talk about 'fear of data' makes clear that they felt there are good reasons to treat data and data analysis with the greatest respect. We will return to this in the next section on 'power'.

In relation to fear of data there was a good tip on how to incorporate data in training for creatives that we would like to share. (And then the quote) good tip though on how to incorporate data in training for creatives:

> People who come to our schools, like 80%, have zero technical knowledge. But we create a sort of low threshold. So, it's like, 'Okay, you should watch these video tutorials on how to collect data from YouTube, and then here is how to analyze it. And here's how to visualize it. And then here's how to tell stories with what you've found.' And that's it. (Academic digital culture researcher, NL)

It was used to build the MediaNumeric course where it worked well. Students in general education might need more reassurance and guidance for a genuine cumulative learning process. Using the lens of affect, however, we see how it would be too easy to follow straightforward rational arguments and discard the mixed feelings that are also alluded to. Fear, excitement, awe and reticence interweave in the material after all. Different forms of engagement, excitement and keeping one's distance, whether fearful, joyous or enthusiastic, will determine in large part who will use data and how. Not only does this take us beyond listing formal training requirements, it underlines how a significant part of our reactions are not styled or occur from primarily manipulative motives. Understanding that affect operates autonomously from cognitive processes allows us to understand that dealing with resistance against embracing data and data analysis will need care and time as much as enthusiasm needs careful fostering. More generally, understanding that affect is an important part of meaning-making that links perception, representation and identity formation allows us to see what is at stake in relation to data for individuals from their personal as well as from broadly social and cultural perspectives.

A feeling of overview and ownership of one's data and of the process of analysis are therefore important when training for data literacy. Data literacy teaching, another informant argued, is best built on real-life examples of practices of data-based storytelling. Data has the potential to inspire, ignite a desire to emulate (e.g., by reverse engineering examples) and dispel misconceptions about data work. This expert is the head of data journalism at a national newspaper in the Netherlands. She taught herself data journalism techniques in the 2010s by deconstructing data-based stories to understand their structure and then meticulously retracing the steps involved:

> Based on the story presented, my first step would be to figure out what his [the data journalist's] data set looked like, what variables were included, and what calculations were performed to support the claims made in the story. My approach would be to reverse-engineer the article, deconstruct it to identify the 'ingredients'—the data points that were incorporated, the analytical methods applied, and the questions posed during the investigative process. Through this analysis, I would seek to grasp the construction of the story, recognizing that if certain data tables were used and specific questions asked, then the answers found—assuming that the journalism was sound—would validate the results reported. (Head of data journalism at a national newspaper).

The head of data science at a major TV company offered a similar argument. The aspiration for data literacy can be fueled by curiosity, he argued. 'You think you want to figure something out, so you sort it out.' In terms of affect, the approach that data analysis is 'doable' or exciting has its own affective load. Bringing all these possibilities into play will provide for a different conversation about the power of data than the (negative) one that is often had. It suggests that in as far as fear leads to caution, building self-confidence is an evident challenge to realizing data literacy which might be easier to achieve when we take the affect of data into account.

5.3 … In Terms of the Power of Data

Data literacy matters because data confer power most easily on to the powerful. That makes 'power' an important part of how data is understood beyond its affective force. An editor of the online archive of one of a big broadcasting corporation in the UK said:

> You want to be able to hold powerful people to account for the decisions they make that affect everybody in society, and believe me the people that control the data are going to be in power. Your ability to understand that and to interrogate that is absolutely fundamental to your purpose as a journalist.

The editor suggests that we are entering a pivotal era for the free press, comparable to the invention of the printing press:

> I think data is the key ingredient in that. I don't know whether I'm optimistic or pessimistic. Most days I'm quite pessimistic because I think the powerful people who control the data have such an advantage that it's going to be very difficult for journalism to perform the role that it performs in society as effectively as it has done up until now.

Data is perceived here as knowledge and hence as power. It is as likely to be a tool of oppression as a tool for political struggle and critique. For journalism to remain the watchdog of democracy, one of its cherished 20th-century professional identities, it must step up its game. This is a well-rehearsed theme that structures much of the conversation about data analysis and data literacy in the interviews. It sidesteps that journalism has also been a domain of the re-establishment of a particular type of hegemonic power, personified in a dominant Western white male outlook, with 'objectivity,' or 'a view from nowhere', as its main tenet.[32]

Informants in journalism and data visualization stress data's role in accountability and the need for critical data interrogation reminiscent of what Hallin identified as journalism's high modern period.[33] Data for them is sexy. Others marshal a threatening vista to inspire engagement with what data is and what data does. Data-driven technology, after all, easily fools us by offering highly credible disinformation:

> Data can be manipulated for malicious intent. For example, deep fakes, face swaps and other applications using known algorithms can ultimately threaten the safety of society, whether these actions are carried out by conspiracy, or simply just for simple goals like money, and the pursuit of power. (PhD in computer science, PL)

Even without malicious intent, data analysis is fraught with danger for the less than fully initiated, though these can be different from what the non-initiated might think. A journalist with a technical background told us that it is important to "be very careful with the outliers of the data" and be aware of the temptation to embellish a story at the expense of accuracy, which they find easiest to express in statistical terms:

> If you want to define what data means, you always need to be very careful about insecurity, confidence intervals, and those kinds of things. So, I think that data journalism requires a little bit more of a scientific approach. And in a way as journalists, we of course want to write the story, but that story needs to sell, and therefore we like to, if we have a fact, we like to over-exaggerate it a little bit so the story is nicer. With data, it just requires a little bit more carefulness. (Data journalist, Investigative journalism agency, NL)

Data is nothing without interpretation and therefore requires more than a technical approach. The technology in turn needs to be treated with respect. Interpretation requires mastery of analytical procedures and their pitfalls, presenting us with the very down-to-earth ways in which power is threaded through the discussion of data literacy. The banal reality of data analysis is that it needs a skill set as much as it needs open data that is qualitatively okay, because data, data analysis and data literacy are produced by and productive of deep social rifts.

A head of a fact-checking division in a journalistic platform (UK) talks about exactly this. Access to data is limited: it is behind a paywall or there are restrictions on data formats in an organization, sometimes simply the quality of available data is bad. "If we want inputs to the fact-checking process, we need high-quality, well-published data that we can trust." Open data published, for instance, by governmental bodies do not always meet these standards. Those who do fact-checking lament this. They insist we need technological and cultural shifts to make such data interoperable and easily usable, describing this necessary change as "fixing the plumbing of the web":

> It's a very unglamorous kind of world to work in, but fixing the plumbing of the web, making sure that this data is published in a way that it can be used and is interoperable with more of the fact-checking process, is a technology thing and is also a cultural thing that will take time to fully see through (Head of fact-checking division, UK).

The power of data, in contrast to this image of plumbing, is thus twofold. It can hold the powerful to account but it also, much more mundanely, lies in its potential to impress and dazzle and to allow for visuals that will end arguments. The power of data in storytelling is also Machiavellian; it can be manipulated for the good and the bad. It can bestow personal power: having data at your command allows access to debates, production processes, and jobs for individuals. It offers forms of control for institutions and regimes. In journalistic and other forms of media production, data is powerful even when it is not explicitly visible:

> It may be that you use zero numbers, zero graphics in your final article, even though there may have been important data work behind it. (Programme director of school of data, FR).

The skilled use of data sets can broach new alliances:

> Generally speaking, journalists are not held in the highest regard by scientists and researchers, and I think that working on data has made it possible to re-establish a much more constructive level of exchange. I think that data journalism brings a lot to the table through collaborations outside the office, and in particular, the world of research; we saw this with the Panama Papers leaks. (Co-founder dataviz platform, FR)

If talking about how data is perceived shows us data as weird but useful, delicious to some and scary to others, the theme of affect tilts that conversation 90 degrees and affords a sense of how all of us are not just touched by data but invested in what data is and does. We all have feelings about data and working with data, whether we are scared or fascinated. In the end, though, and turning further, data is seen as a tool. Whether data in media content helps hold powerful institutions to account, makes for better journalism, for lovely visuals or is used by bad people to misinform or manipulate us. Data literacy, this last quote says while quoting the Panama Papers leak, offers storytelling perspectives that other forms of journalism or documentary media making do not have.

6 Conclusion

Gaining a deeper understanding of how experts in media-related professions talk about data serves the goal of knowing better what to take into account when educating students and early-career professionals in the creative industries for working with data and for data literacy. Theorizing the interview material with a light touch, we hope, will help make our interpretation of the interview material more easily transferable to other professional sectors such as policy development or research.

As the creative industries continue to mature in response to technological advancements, understanding and employing data becomes increasingly crucial.[34] Powerful AI-tools will further reshape creative work, storytelling and how society provides in its need for news, entertainment and the overlapping zones between them. Our research demonstrates that successfully integrating data literacy in these fields requires more than technical training, it needs a nuanced understanding of the emotional landscape surrounding data and a commitment to addressing the multi-faceted relationship we have with data in educational and professional development programs. Pivotal elements in understanding that relationship are perception, affect and power.

These are not neatly separate concepts but present as chained aspects of understanding. Different from thinking in terms of formal requirements, they suggest that working with data requires a relationship, that once you engage with data and offer time and attention, data will have something to give. It is not a lack of imagination or of a skill set to feel that data sets are a foe. Everybody who works with data needs to befriend it, acknowledge good and bad qualities and come to terms with what rules and limitations apply. As in all relationships this will register affectively as positive or negative energy. Data, last but not least, are the kind of friend parents feel weary about: will the friendship lead to trouble or bolster your ambitions and upward mobility? It is difficult, at the outside, to know what it will be.

Interestingly, then, we now think that presenting data as fickle, as needing attention and care, may well work better than insisting on technical skills. As a long-term educator among our informants told us, these students can easily pick up such skills from YouTube tutorials. The added value of a data literacy course needs to be in forms of relationship counseling. All of us engaged in teaching and working in the media and creative industries need to come to terms with how each of us understands what they can do to (re)humanize datafication. The increasing power and significance of data require that we do so. That will require some teaching for technical understanding of what digital data sets are but mostly, interestingly, it needs a combination of emotional and critical intelligence of those who teach.

Acknowledgement

The MediaNumeric project was co-funded by the Erasmus+ Programme of the European Union. To learn more about the project visit **www.medianumeric.eu**.

Grant acknowledgment: The research leading to these results has received funding from the European Commission under grant agreement No. 621610-EPP-1-2020-1-NL-EPPKA2-KA.

Notes

1. See: **https://www.medianumeric.eu/** Medianumeric ran from 2021-2023. Between 2022 and 2023 MediaNumeric hosted three in-person courses on data-driven journalism, creative storytelling and fact-checking. The week-long courses were open to twenty students at a time from universities in Poland, The Netherlands and France. The project also delivered an open online course: **https://www.medianumericacademy.eu/**
2. For definitions see Claire Wardle and Hossein Derakhshan, "Information disorder: Toward an interdisciplinary framework for research and policy making," 2017 (**https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html**).
3. QlickTech and Data Literacy Project, "Data literacy: The upskilling evolution." **https://thedataliteracyproject.org/data-literacy-the-upskilling-evolution/** (December 19, 2022)
4. Bernard Marr, "The Importance of Data Literacy and Data Storytelling." *Forbes*. **https://www.forbes.com/sites/bernardmarr/2022/09/28/the-importance-of-data-literacy-and-data-storytelling/** (September 28, 2022).
5. Graham Meikle, *Deepfakes* (Oxford: Polity Press, 2022).
6. Rob Wijnberg, *Voor ieder wat waars (A little truth for everybody)*. (Amsterdam: de Correspondent, 2023)
7. Catherine D'Ignazio and Lauren Klein, "Seven intersectional feminist principles for equitable and actionable COVID-19 data." *Big data & society*, 7, no 2 (2020)
8. In this edited collection, all chapters are accompanied by case studies from across the creative industries: from the music industry to advertising, to the theatre. Melissa Terras, Vikki Jones, Nicola Osborne and Chris Speed (editors) *Data-driven Innovation in the Creative Industries*. (London: Routledge, 2024). Ramy Rahimi offers a useful survey in their paper: "A survey of technology acceptance models in the creative industry: exploring key limitations." In 2020 13th International Conference on Developments in eSystems Engineering (DeSE), (2020), 9-14.
9. Our analysis of data literacy needs within the creative industry was published as an open-source report: Natalia Berger, Joke Hermes, Faye Mercier and Emma, "Updated Needs Analysis Report," (2021) **https://www.medianumeric.eu/resources/**. It served as the foundation for an intensive one-week training course on storytelling with data. The course's online version is open-access: **https://www.medianumericacademy.eu/**
10. The interviewers included Natalia Berger, Maria Drabczyk, Mirosław Filiciak, Barbara Giza, Ewa Korzeniowska, Clément Malherbe, Claude Mussou, Johan Oomen, Jacqueline Pietsch, Kuba Piwowar, Sten-Kristian Saluveer and Daniel Sorabji. The list of 56 informants was curated by this group to reflect a broad group of data-related professions in news production and media content production more generally, in data research and in higher education.
11. Juliet Corbin and Anselm Straus, "Grounded theory research: Procedures, canons, and evaluative criteria." *Qualitative sociology* 13, no 1 (1990), 3-21

12. George Marcus and Michael Fischer, *Anthropology as cultural critique: An experimental moment in the human sciences*. (Chicago: University of Chicago Press, 1999)
13. See Janet Heaton. *Reworking Qualitative Data*. (London: Sage, 2004).
14. danah boyd & Kate Crawford, "Critical questions for big data." *Information, Communication & Society* 15, no 5 (2012), 662-679.
15. Ellen Mandinach and Edith Gummer, "A Systemic View of Implementing Data Literacy in *Educator Preparation. Educational Researcher* 42, no 1 (2013), 30–37; Edith Gummer and Ellen Mandinach, "Building a conceptual framework for data literacy" *Teachers College Record* 117, no 4 (2015), 1-22,; Catherine D'Ignazio and Rahul Bhargava, "Approaches to building big data literacy [Conference session]," Proceedings of the Bloomberg Data for Good Exchange Conference, New York, NY (2015); Chantel Ridsdale et al, "Strategies and best practices for data literacy education: Knowledge synthesis report." Dalhouse University (2015). **https://dalspace.library.dal.ca//handle/10222/64578**.
16. Ina Sander, "Critical big data literacy tools—Engaging citizens and promoting empowered internet usage." *Data & Policy* 2, no 5 (2020).
17. William Alston, "Perception and representation." *Philosophy and Phenomenological Research* 70, no 2 (2005), 253-289; Klaus Solberg Søilen, "Users' perceptions of Data as a Service (DaaS)," *Journal of Intelligence Studies in Business* 6, no 2 (2016); V. Allahverdov & V. Bogdanova, *Psychology: Textbook (2nd ed.)*. (Moscow: Prospekt Publishing House, 2001), 67.
18. Sara Ahmed, "Affective economies" *Social text* 22, no 2 (2004), 117-139; Teresa Brennan, *The transmission of affect*. (Ithaca and London: Cornell University Press, 2004); Brian Massumi, *Politics of affect*. (Cambridge and Malden: Polity, 2015)
19. Margaret Wetherell, *Affect and emotion: a new social science understanding*. (London: Sage, 2012), 4. See also: Margaret Wetherell, "Trends in the turn to affect: A social psychological critique". *Body & Society* 21, no 2 (2015), 139-166.
20. Annette Hill, *Media experiences: Engaging with drama and reality television*. (London and New York: Routledge, 2019) 12.
21. Ruth Leys, "The turn to affect: A critique." *Critical Inquiry* 37, no 3 (2011), 434-472
22. Safiya Noble, *Algorithms of oppression*. (New York: New York University Press, 2018)
23. Catherine D'Ignazio, "Creative data literacy: Bridging the gap between the data-haves and data-have nots." *Information Design Journal* 23, no 1 (2017), 7
24. Chris Barker on 'power' in *The Sage Dictionary of Cultural Studies*. (London: Sage, 2004)
25. Jiangtao Wang, Bin Guo, and Liming Chen, "Human-in-the-loop machine learning: A macro-micro perspective." arXiv preprint arXiv:2202.10564 (2022).
26. Matthew Zook, S. Barocas, danah boyd, et al, "Ten simple rules for responsible big data research. *PLoS Computational Biology* 13, no 3 (2017), e1005399,
27. Darnell Huff, *How to Lie with Statistics*. (New York: W. W. Norton & Company, 1953/1993); Heather Brook Adams, Risa Applegarth, and Amber Hester Simpson, "Acting with algorithms: Feminist propositions for rhetorical agency." *Computers and Composition* 57 (2020), 102581; Koen Leurs, "Communication rights from the margins: Politicising young refugees' smartphone pocket archives". *International Communication Gazette* 79, no 6-7 (2017), 674-698; Sander de Ridder, "Sexting as sexual stigma: The paradox of sexual self-representation in digital youth cultures." *European Journal of Cultural Studies* 22, no 5-6 (2019), 563-578.
28. Ulises Mejias and Nick Couldry, "Datafication," *Internet Policy Review* 8, no 4 (2019), 1-10.
29. Helen Kennedy, Susan Oman, Mark Taylor, Jo Bates and R. Steedman. "Public understanding and perceptions of data practices: A review of existing research." CBS Research Portal (2021) **https://research.cbs.dk/en/publications/public-understanding-and-perceptions-of-data-practices-a-review-o**; boyd and Crawford, "Critical questions for big data".
30. D'Ignazio and Klein, "Seven intersectional feminist principles"; Noble, *Algorithms*
31. D'Ignazio, "Creative data literacy," 6.
32. Candis Callison and Mary Lynn Young, *Reckoning: Journalism's Limits and Possibilities*. (Oxford: Oxford Academic, 2020).
33. Daniel Hallin, "The passing of the "high modernism" of American journalism." *Journal of Communication* 42, no. 3 (1992), 14-25.
34. See Melissa Terras, Vikki Jones, Nicola Osborne and Chris Speed (ed), *Data-driven Innovation in the Creative Industries*. (London: Routledge, 2024).

Biographical Note

Dr. Natalia Berger is senior researcher and lecturer Creative Business at Inholland University of Applied Sciences, the Netherlands. Her doctoral dissertation at St.-Petersburg University (Russia) was dedicated to the methodology of investigative journalism. Her areas of interest include critical discourse analysis, framing and other forms of textual research applied to analysis of social change. Her recent projects focus on fashion in public discourses, data literacy and data-driven technology in creative industries.

Dr. Joke Hermes is a professor of practice-based research in Inclusion and the Creative Industries at Inholland University (Netherlands) and a lecturer in Television and Cross-Media at the University of Amsterdam. She studied political science in the University of Amsterdam and completed her doctorate with a dissertation on the reading of women's magazines. Her expertise lies in the area of qualitative research, citizenship and popular culture. Joke Hermes was a founder and co-editor of the international journal *European Journal of Cultural Studies*. Her work on popular culture and gender, television audiences, internet communities and cultural citizenship, among other subject areas, has been widely published.

VIEW Journal of European Television History and Culture Vol. 13, 25, 2024
URL: **https://doi.org/10.18146/view.327**
Publisher: Netherlands Institute for Sound and Vision in collaboration with Utrecht University, University of Luxembourg and Royal Holloway University of London.