MAKING SENSE OF TRIFLES: DATA NARRATIVES AND CUMULATIVE DATA DISCLOSURE

Callum Nash / Dan Carey / Emma Nicol / Amal Htait / Burkhard Schafer / Jo Briggs / Wendy Moncur / Leif Azzopardi

Senior Research Assistant, Northumbria University, School of Design, Sutherland Building, Newcastle-upon-Tyne, NE1 8ST, UK
callum.nash@northumbria.ac.uk; https://northumbria.design/people/nash/

Independent designer, danpaulcarey@gmail.com

Postdoctoral Researcher, University of Strathclyde, Computer and Information Sciences, 16 Richmond Street, Glasgow, G1 1XQ, UK, emma.nicol@strath.ac.uk; https://www.strath.ac.uk/staff/nicolemmadr/

Research Associate, University of Strathclyde, Computer and Information Sciences, 16 Richmond Street, Glasgow, G1 1XQ, UK
amal.htait@strath.ac.uk; https://www.strath.ac.uk/staff/htaitamalms/

Professor of Computational Legal Theory, University of Edinburgh, School of Law, Old College, Edinburgh EH8 9YL, UK
b.schafer@ed.ac.uk; https://www.law.ed.ac.uk/people/professor-burkhard-schafer

Associate Professor, Northumbria University, School of Design, Sutherland Building, Newcastle-upon-Tyne, NE1 8ST, UK
jo.briggs@northumbria.ac.uk; https://www.northumbria.ac.uk/about-us/our-staff/b/jo-briggs/

Professor, University of Strathclyde Computer and Information Sciences, 16 Richmond Street, Glasgow, G1 1XQ
wendy.moncur@strath.ac.uk; https://www.strath.ac.uk/staff/moncurwendyprofessor/

Reader, University of Strathclyde Computer and Information Sciences, 16 Richmond Street, Glasgow, G1 1XQ
leif.azzopardi@strath.ac.uk; https://www.strath.ac.uk/staff/azzopardileifdr/

Schlagworte: ***Privacy by design, Data Protection, visualisation***

Abstract: *The law does not concern itself with trifles. If a risk is deemed minimal, or an infraction negligible, invoking the authority of the law often seems unnecessary. However, there are increasingly fields of human activity where this principle leads to gaps in the protection necessary for a flourishing society. This paper reports findings and ideas from a research project in cumulative data disclosure, where an aggregation of in themselves harmless data points can expose the users of social media to significant personal risk.*

1. A Trifle of an Introduction

This paper contributes to the discussion around the "privacy paradox" – the observation that many users of online services profess that they value their privacy, but yet routinely take decisions that expose them to privacy risks. We argue that one of the reasons for this discrepancy is that in order to reduce their cognitive load, users have to use incomplete risk models that at best assess the risk associated with individual pieces of data and their disclosure, but cannot take into account the cumulative risk that the sum of these decisions over time and platforms creates for them. A risk that appears thus minimal and trifling, if repeated over time and across platforms, can nonetheless become significant.

In law, we find this idea expressed in the legal maxim that the law does not concern itself with trifles: *De minimis non curat lex.* While there are as we will see sound justifications for this rule, the accumulation of "trifling" rule violations or harms can in the long term cause significant damages, which poses a difficult regulatory challenge. The paper presents some of the findings and ideas of an interdisciplinary research project[1] that aims to assist citizens and businesses to adopt safer online data practices.

[1] https://cumulative-revelations.github.io/revelations/publications.html.

In the first section, we give a description of the problem of cumulative harm and put it also in a wider historical context. In the second section we discuss some of the findings from a series of qualitative interviews with users of online services that we conducted in 2020. We then introduce findings from a workshop where we helped people to visualise data risks by asking them to take on the role of adversaries who conduct an adversarial media search of a target person. We conclude with a brief indication of future work.

2. De Minimis non Curat Lex

It is commonplace that the law does not concern itself with trifles. De Minimis Non Curat Lex is a legal maxim that can be found across times and across jurisdictions. Despite the Latin expression, it does not seem to have been recognised explicitly in Roman law under that name, though we find the idea endorsed in Ulpian's Digests.[2] By the 15th century, in the common law world, we find a similar pattern, with the idea discussed by Braxton in the 13th century though not under this name, while Blackstone' Commentaries on the Laws of England lists it explicitly as a valid principle.[3] In one of the earliest reported cases, *York v York* from 1431, the court held that "*No action lies of a waste but to the value of a penny; for de minimis non curat lex*."[4]

In the common law world, it found application across a broad range of legal fields, and, confusingly, with varying rationales and criteria. Some countries incorporated it explicitly in their statutes, for instance California, where it became in 1871 a general principle of legal interpretation enshrined in Art 3533 of the Civil Code that simply states: *The law disregards trifles.*

The increasing importance of the principle was emphasised in the US by the US Supreme Court, which ruled in 1992 that:

> "the venerable maxim de minimis non curat lex ... is part of the established background of legal principles against which all enactments are adopted, and which all enactments (absent contrary indication) are deemed to accept."[5]

A key justification of the principle from the beginning is that enforcement against merely trifling violations ultimately harms respect of the law more generally, and prevents socially beneficial activities through overly cautious behaviour by citizens. In the US, the seventh Circuit put this rationale well:

> "to place outside the scope of legal relief the sorts of intangible injuries ...that must be accepted as the price of living in society,"[6]

Tracing the historical roots of this principle is important for one of the arguments of this paper, which links the principle in complex and sometimes paradoxical ways with the history of technology and technology regulation. The de-minimis maxim was conceived before the industrial revolution, and embodies a simple, causal model of reality: in an action for damages for instance, it is always possible in principle to determine the cause of a harm. However, if the harm is minimal, the social costs of enforcement would be disproportionate and ultimately harm society and social cohesion.

That the principle became ever more important from the 19th century onwards is, we argue, due to a number of factors that will later allow us to understand better some of the difficulties data protection law is facing today. First, advances in science and technology also allowed us an increasingly comprehensive understanding of how our actions, even on an apparently small scale, impact our world. Improved methods to measure and detect causal effects of our actions in turn enabled us to prove, in principle, smaller and smaller harms. In law, we see this in particular in the field of toxicology and its increasing importance for forensic investigations and

[2] D. 4. III. 9–11.

[3] Bracton, De Legibus (Twiss ed. (I878)) at I. I, fol. 9, 607; Blackstone Commentaries 3 15 p. 262.

[4] Cited from Veech/Moon, De Minimis Non Curat Lex, 45 Michigan Law Review 1947 p. 537–570 at 5.

[5] Wis. Department of Revenue v. William Wrigley, Jr., Co., 505 U.S. 214, 231 (1992).

[6] Swick v. City of Chicago, 11 F.3d 85, 87 (7th Cir. 1993).

prosecutions.[7] The more we detect and causally explain even small instances of harm, or minimal rule violations, the more important it becomes to guard against over-zealous enforcement of rules that were conceived at a time where such a nuanced analysis was impossible.

A second element is the increased reliance of laws as a way to structure all aspects of our lives, especially in the modern administrative welfare state after the Second World War. Sometimes called "juridification",[8] this development also meant that with more and more rules, their judicious application has become more and more important. This in turn requires appropriate discretion by the enforcement agencies.[9]

Finally, we encounter from the early 19th century onwards a new *mode* of regulation in response to the industrial revolution and the technologies that enabled it. The danger of exploding steam engines, and an ever-increasing death toll especially when deployed aboard ships, forced governments in the industrialising nations to take on a radically new role. While law traditionally had been evoked only after a harm had materialized, appropriating blame, restitution and punishment as appropriate, it now had to adopt also a forward-looking perspective. For the first time, legally mandatory safety certification schemes were introduced, often against considerable objections from industry, that tried to pre-empt harm from occurring rather than merely dealing with its aftermath.[10] As part of a historical compromise however, these new regulations did not try to prevent *all* harms from occurring, something that arguably would have shut down the new industries in their infancy. Rather, it now became a role of governments, and the experts advising it, to define acceptable levels of minimal harm. Risks below these thresholds became "de minimis" by law, a risk we have to accept for the greater social benefits that these technologies bring. The newly emerging disciplines of statistics and probability theory provided the theoretical underpinning of this "actuarial" approach to governance, culminating in the "risk society" of the 20th century.[11]

The GDPR and its "risk based" approach[12] is but a recent heir of this tradition that began in the 19th century. With this ancestry, it also shares some of the conceptual limitations and dangers of this approach. Risk, from the beginning, was defined and understood as an average that a society was willing to tolerate, or, more honestly, the new question became: how many additional deaths are we as a society willing to accept? What this approach ignored however, especially in the early days, was the unequal distribution of risk, especially if this was caused by cumulative risk exposure. True, the newly certified boilers were less likely to explode, making it potentially a rational decision to agree (consent) to be transported by them. However, what this approach overlooked was the cumulative effect of risk exposure: for someone travelling only occasionally, the risk was now indeed manageable. But what about frequent travellers? Or members of the crew who for years would spend most of their time in the vicinity of these machines? For them, the chances that *eventually*, they would be involved in a serious accident were considerable.

The GDPR intersects with the issues we have discussed so far in several ways. We can for instance think of the "household exemption" as a statutory expression of the de-minimis principle. Data processed by individuals within their normal household activity are unlikely to create the type of harm that requires legal intervention. The risk assessment and risk compliance required of data controllers, but also the corresponding enforcement discretion of the various Data Protection Agencies directly follows the trajectory of quantitative regulation and allows, again like the de-minimis principle, the focussing of scarce resources on those dangers and harms that are, or seem to be, the most significant.

[7] Watson, Poisoning crimes and forensic toxicology since the 18th century. Academic forensic pathology, 2020, 10 p. 35–46.

[8] Teubner, Juridification concepts, aspects, limits, solutions. In ibid (ed) Juridification of social spheres, 2012, de Gruyter, pp. 3–48.

[9] See e.g. Bressman, Beyond accountability: Arbitrariness and legitimacy in the administrative state. NYUL Rev. 2003, Vol 78, p. 461.

[10] See Burke, Bursting boilers and the federal power. Technology and Culture 7, no. 1, 1966, pp. 1–23; for a direct connection to computer technology see Leveson, High-pressure steam engines and computer software. Computer 27, 1994, 65–73.

[11] Fressoz, Beck back in the 19th century: towards a genealogy of risk society. History and Technology 23, no. 4, 2007, pp. 333–350.

[12] See e.g. Quelle, Enhancing compliance under the general data protection regulation: The risky upshot of the accountability-and risk-based approach. European Journal of Risk Regulation 9, no. 3, 2018, p. 502–526; Demetzou, GDPR and the Concept of Risk. In Kosta et eal (eds) IFIP International Summer School on Privacy and Identity Management (). 2018 Springer, Cham, pp. 137–154.

Intuitively, it seems to make sense to have fewer obligations on controllers who only have small amounts of prima facie less dangerous data. Equally, one should prosecute an accidental use of a staff photograph without permission on the company website for advertising purposes less vigorously than a data breach that exposed 200,000 credit card details to fraudsters. The argument that a de-minimis principle benefits the respect for the law that we encountered above for the de-minimis principle is visibly in operation here. In the UK for instance, a barrage of press reports about over-zealous approaches to the GDPR have become a staple diet, used not only to discredit data protection law, but to feed scepticism of the EU (in pre-Brexit times) in general.[13]

However, while thus rationally justifiable, the de-minimis principle and focus on permissible risk in data protection law suffers from the same limitation that we saw for the risk-based approach to technology regulation. Individual events may indeed be harmless and trifling, but their cumulative effect may be less so, and different data subjects are differently exposed to these cumulative dangers. The unauthorised photograph with name on a staff website for instance will not normally expose the person to risks, and neither will a post on Twitter that this person is celebrating their birthday, but taken together, an adversary may use them for a social engineering attack against the employer that exploits knowledge of the significant date with knowledge of the physical appearance, to pretend closer familiarity with an employer than is the case to gain trust and access.

An illustrative real-life example took place on Twitter during the election campaign for mayor of London in 2020, at the height of the Covid-19 pandemic. One of the candidates campaigned on a platform critical of the government's vaccination policy. In a much-circulated Tweet, he announced that he was using his "deep needle anxiety" to claim exemption from any vaccination mandate. What he had apparently forgotten though was that some time earlier, he had proudly shared on social media an image of his latest tattoo. Readers quickly picked up on the inconsistency of a medical phobia of needles and the enthusiasm for tattoos, to the significant embarrassment of the candidate.[14] While here the political process arguably benefited from the ability of citizens to form a rounded impression of a candidate, we can easily see how ordinary citizens can fail to properly assess the risk of information sharing if they do not keep track, in theory, of all the other pieces of data they have disclosed about themselves, not only on platform, but across them all.

We find this problem particularly in the third intersection between GDPR and cumulative risk assessments, namely, the problem of consent. Consent is of course a cornerstone of the GDPR and similar data protection regimes, even though its problems and limitations are now widely understood.[15] Critical discussions surrounding consent focus on the power and information imbalance between data controllers and data subjects, and the difficulty of providing "informed consent" in the absence of knowledge about processing that is, on the one hand, detailed and comprehensive and on the other hand, clear and easy to process without disproportionate investment of time. In other words, without an understanding of the risks involved, consent cannot be informed.[16]

The way our research frames this problem takes it a step further: no amount of information provided by the controller alone can lead to a comprehensive assessment of the risks that data disclosure entails. This risk is cumulative and depends on past behaviour of the data subjects including their behaviour on other platforms. When deciding whether sharing a piece of information on a social platform is safe, we make decisions under "bounded rationality",[17] with limited information and time at our disposal. Centring data protection law

[13] For one example of many, Daily Telegraph: GDPR chaos as churches stop prayer requests and charities prepare to halt meals on wheels https://www.telegraph.co.uk/news/2018/05/25/gdpr-chaos-churches-stop-prayer-requests-sick-charities-prepare/.

[14] https://www.thepoke.co.uk/2021/04/15/laurence-fox-mask-and-vax-avoidance-pointed-responses/.

[15] Bergmann, The Consent Paradox: Accounting for the Prominent Role of Consent in Data Protection. In Kosats et al (eds) IFIP International Summer School on Privacy and ID Management 2017, Springer, Cham. pp. 111–131.

[16] E.g. Padden//Öjehag-Pettersson, Protected how? Problem representations of risk in the General Data Protection Regulation (GDPR). Critical Policy Studies, 2021 pp. 1–18.

[17] Simon, A Behavioral Model of Rational Choice, Quarterly Journal of Economics, 69(1) 1955 pp. 99–118.

around a "magic moment" of consent,[18] even under the best conditions, then becomes an assessment of whether *this* platform can be trusted with *this* piece of information about us.

But as our examples above show, this is a necessarily incomplete picture. To fully understand the cumulative risk exposure, it is not sufficient to look at an individual piece of information, but diachronically, the data subjects' past sharing behaviours, and synchronically, their data sharing behaviour across platforms. This need to assess cumulative risks marks also a critical difference between the technologies of the first and second industrial revolution, and the third. While even for steam engines, certification of individual machines did not give the full picture, it worked reasonably well, as it was clear which acts exposed someone to the potential risk (i.e. travelling on a steam boat), and how many of these individual journeys could expect a critical event over time. The decentralised Internet changed this equation not just quantitatively, but qualitatively. Now it became possible to aggregate individual actions (or data) at an unprecedented scale, leading to situations where the whole truly is more than the parts.

This can be used for good – Wikipedia as a resource where professionals contribute a "trifle" of information each, at low costs to themselves, but cumulatively creating an invaluable resource. Similarly, that ability can be abused. Denial of Service attacks are a particularly good illustration of the problem that the law faces: each individual action of an attack falls below the de-minimis threshold, and indeed raises questions of causality, as the harm would also have occurred had *this one ping* been omitted. Nonetheless, their cumulative effect is serious damage to a service platform, such that we must find a way to conceptualize how an aggregate of individually permissible actions can collective become an illegal act. We are therefore facing a paradox: on the one hand, the de-minimis principle is also a response to scientific and technological innovation and therefore needed in any complex legal system. On the other, it can hinder adequate legal responses to actions that individually fall below the de-minimis threshold, but collectively are harmful, and which are typical for the digital world, where data is sticky, difficult to expunge, and can be combined across contexts. We, by contrast, explore how users of social platforms can be helped to better manage their cumulative data risk. In the next section, we introduce some of our work that will eventually lead to ways in which cumulative exposure risks can be intuitively visualised.

3. O wad some Power the giftie gie us.

As part of our project into the cumulative effects of data disclosure, we carried out a series of seminars with users of social media that allowed them to explore, and to re-think, their data practices. These workshops build on and were informed by a series of qualitative interviews that we carried out in 2020. Those interviews were conducted using a data narrative approach[19] in order to understand risks, issues and consequences of the digital traces that people leave online. The data narrative approach is intended to capture participants' descriptions of their data, device use, channels and networks of communication; data and information practices.

The study was conducted in May-July 2020. From an original target of 12–20 participants, we expanded to recruit around 50 percent more with a view to better explore and represent demographic effects on changing practices such as those arising from differences in age, gender, educational background and level of technical knowledge.

Interviews followed the following themes: questions to capture information about communications channels, apps, data storage/management systems, and devices used, including whether/how any of these were shared; everyday practices and behaviour patterns around e.g. conducting searches, posting and other digital information sharing; questions concerning participants' awareness of the unanticipated potential for self-disclosure

[18] Van Ooijen/Vrabec, Does the GDPR enhance consumers' control over personal data? An analysis from a behavioural perspective. Journal of consumer policy 42, no. 1, 2019, pp. 91–107.

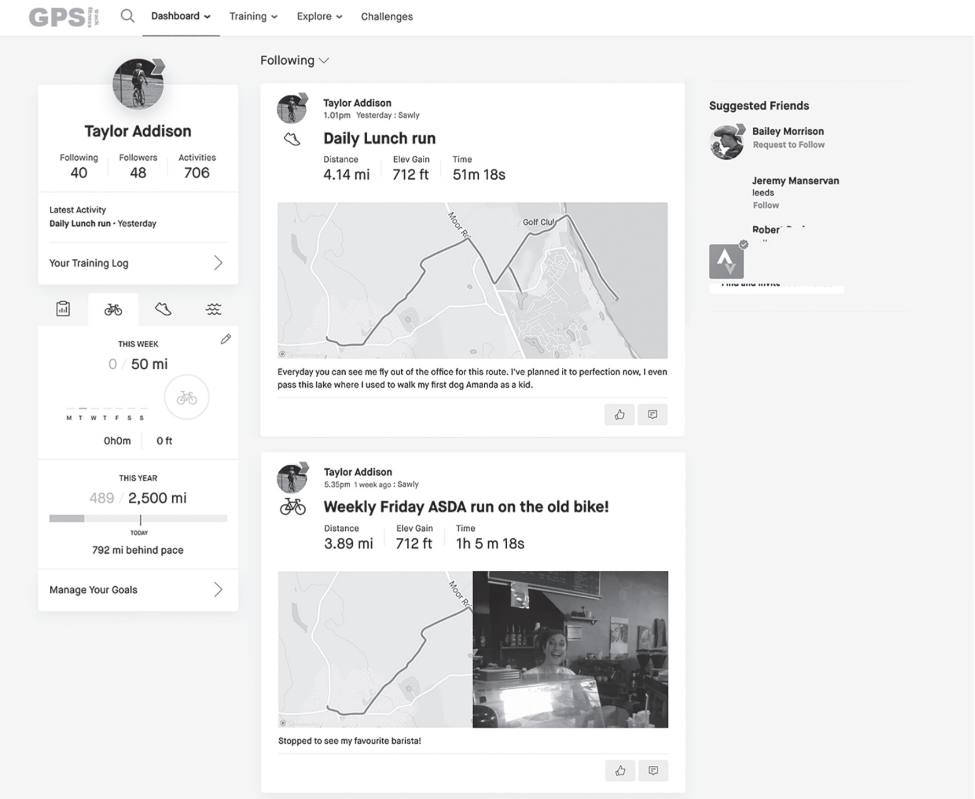
[19] Vertesi et all, Data {Narratives}: {Uncovering} tensions in personal data management. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing ({CSCW} '16). Association for Computing Machinery, San Francisco, California, pp. 478–490. https://doi.org/10.1145/2818048.2820017.

through digital traces and their associated level of concern; a series of questions relating to information management, security setting behaviours, and Lockdown-related changes – especially regarding working from home. Finally, participants were asked to envision a scenario where someone else had to write a book about them based only on their digital traces, and to think about what the resulting book would comprise.

While a detailed discussion of the findings of these interviews will be the subject of another paper, we note here that the overall pattern confirmed our hypothesis that the way people make decisions about data risks is heavily bounded. While many participants voiced concerns about their privacy, the type of stories they reported to illustrate their understanding of the risks overwhelmingly looked at isolated leakage of "high risk items": the mother who disclosed the regiment and location of the advanced training of her son in the military, or the job interview that did not result in an offer because of the embarrassing photo shared on Facebook. However, once participants were asked to shift their perspective and think how a third party would possibly write a book about them, the temporal and cumulative effect of data disclosure became much more pronounced, though individual data pieces remained significant in themselves.

Building on these results, we conducted a follow up workshop that took the idea of "writing a book about someone" as a starting point. Scotland's national poet, Robert Burns, expressed the idea in his Poem to a Louse: O wad some Power the giftie gie us To see oursels as ithers see us – if only we had the power to see ourselves as others see us.

Echoing this idea, we created fictional profiles for a number of social media users. We created for each of them a number of digital traces, like a Twitter post, an image uploaded to Facebook etc. Most, though not all of these traces were intentionally "low risk", that is information one could easily disclose unthinkingly about oneself. A typical example are these screenshots from the profile of the fictional "Alex Smith":



Taylor Addison

Following	Followers	Activities
40	48	706

Latest Activity
Daily Lunch run · Yesterday

Your Training Log

THIS WEEK
0 / 50 mi
0h0m | 0 ft

THIS YEAR
489 / 2,500 mi
792 mi behind pace

Manage Your Goals

Taylor Addison
1.01pm Yesterday : Sawly
Daily Lunch run

Distance	Elev Gain	Time
4.14 mi	712 ft	51m 18s

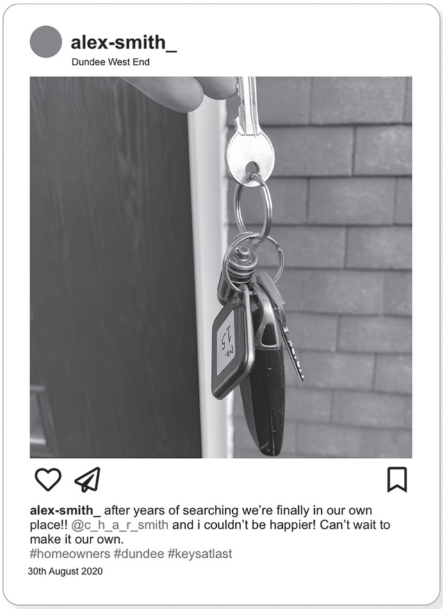
Everyday you can see me fly out of the office for this route. I've planned it to perfection now, I even pass this lake where I used to walk my first dog Amanda as a kid.

Taylor Addison
5.35pm 1 week ago : Sawly
Weekly Friday ASDA run on the old bike!

Distance	Elev Gain	Time
3.89 mi	712 ft	1h 5 m 18s

Stopped to see my favourite barista!

Suggested Friends
Bailey Morrison
Request to Follow
Jeremy Manservan
leeds
Follow
Rober



alex-smith_
Dundee West End

alex-smith_ after years of searching we're finally in our own place!! @c_h_a_r_smith and i couldn't be happier! Can't wait to make it our own.
#homeowners #dundee #keysatlast
30th August 2020

Images 1 and 2: Alex's data traces

The picture shows, not easily seen, the house number. The visibility of the shape of the key would also be sufficient to allow a copy to be made. The map then enables location of the street. For each character, we created a number of similar items, including Facebook profiles that show their friends, data shared from their fitness apps etc. Our workshop participants were then asked to take on one of several "adversarial" roles – the potential employer, the journalist looking for a story, and the concerned friend who noticed changes in online behaviour. A Miro board then allowed the participants to explore how much information they could glean about their target from combining these sources, as shown in Image 3 below.

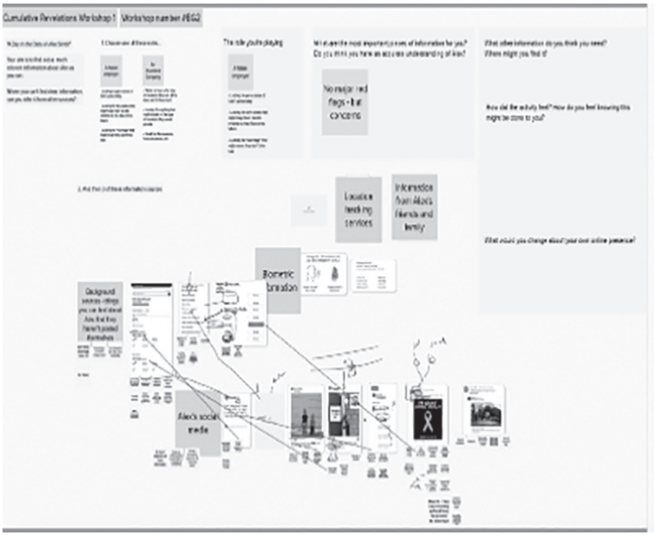
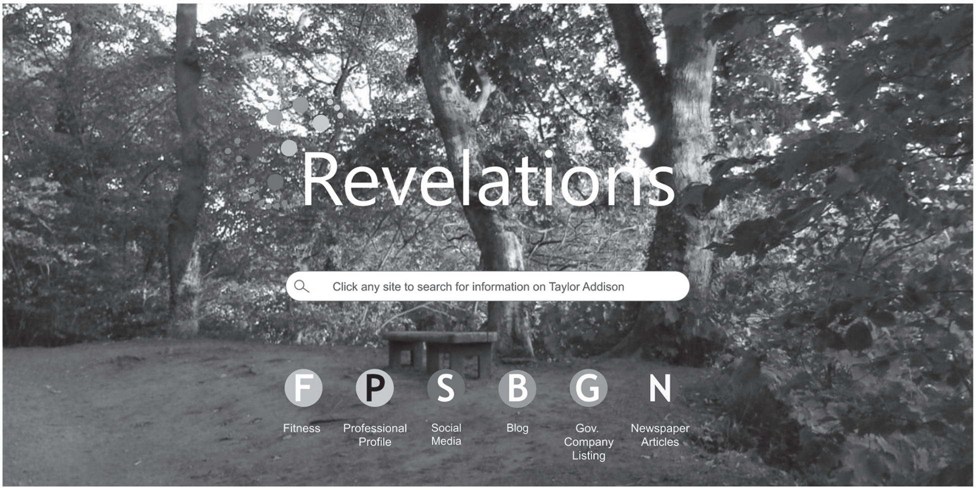

Cumulative Revelations Workshop 1
The role you're playing
No major red flags - but concerns
What other information do you think you need? Where might you find it?
How did the activity feel? How do you feel knowing this might be done to you?
Location tracking services
Information from Alex's friends and family
What would you change about your own online presence?
Biometric information
Background sources - things you can find about Alex that they haven't posted themselves
Alex's social media

Image 3: Miro board of adversarial media search

The participants embraced the scenarios with enthusiasm, and discovered hidden and exploitable connections that we had not anticipated. In their reflective analysis after the activity, they also showed a heightened sensitivity towards the cumulative effect of data disclosure, indicating the pedagogical value of this type of activity for better data practices.

We then turned the activity into a data game, playable without our supervision. The next picture, Image 4, shows the entry screen for the data game.[20] The green, environmental background conveys a message: just as the ecosphere is at danger from the cumulative release of toxins that remain for long periods of time in the environment, so is the health of the "infosphere" at risk from data traces that stay in the environment for much longer than their authors realise, which can combine with other traces to create synergetic hazards that are substantially more "toxic" than their constituent parts.

We cannot in this paper fully develop the analogy between pollution of the 'Infosphere" and what we have learned about environmental pollution and its regulation. We have previously developed a similar analogy for "Big Data". Sometimes called the "new oil", it creates the danger of "oil spills" that can have devastating effects.[21] For "small data", the analogy is even more promising, but also complicated. Environmental science was one of the first disciplines to recognise the danger of cumulative environmental degradation (bioaccumulation) through the release of toxins that sometimes interact with others in unexpected ways, sometimes remaining unobserved in the environment for very long periods of time (persistence), slowly accumulating. Scenarios like these could not easily be handled through risk-based approaches that prescribe simple maximum thresholds of toxins,[22] but required a new model of risk assessment[23] to adequately address issues of environmental justice.[24] Our long-term aim is to see if these models of risk assessment that work well for toxins also work for "toxic data".



Click any site to search for information on Taylor Addison

Image 4: Start page of the data game 1

[20] We are grateful to Melissa Duheric for her work on the design.

[21] Schafer, Compelling truth: legal protection of the infosphere against big data spills. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374, no. 2083, 2016, 742016011420160114 http://doi.org/10.1098/rsta.2016.0114.

[22] See e.g. Wennig, Threshold values in toxicology–useful or not?. Forensic science international 113, 2000, pp. 323–330. Mackay/McCarty/MacLeod. On the validity of classifying chemicals for persistence, bioaccumulation, toxicity, and potential for long-range transport. Environmental Toxicology and Chemistry: An International Journal 20, no. 7 (2001): 1491–1498.

[23] Moretto et al , A framework for cumulative risk assessment in the 21st century, Critical reviews in toxicology 47, no. 2, 2017, pp. 85–97.

[24] Sexton/Linder. The role of cumulative risk assessment in decisions about environmental justice. International journal of environmental research and public health vol. 7, 2010, 4037–4049. doi:10.3390/ijerph7114037; Callahan/Sexton. If cumulative risk assessment is the answer, what is the question?. Environmental health perspectives 115, 2007, pp. 799–806.

Once a character (here, Taylor), a scenario (comprised of the data types listed in the circles) and a task (e.g. journalist looking for a story) are chosen, a visualisation of the digital traces, ordered by categories, is automatically generated, as seen in Image 5 below:

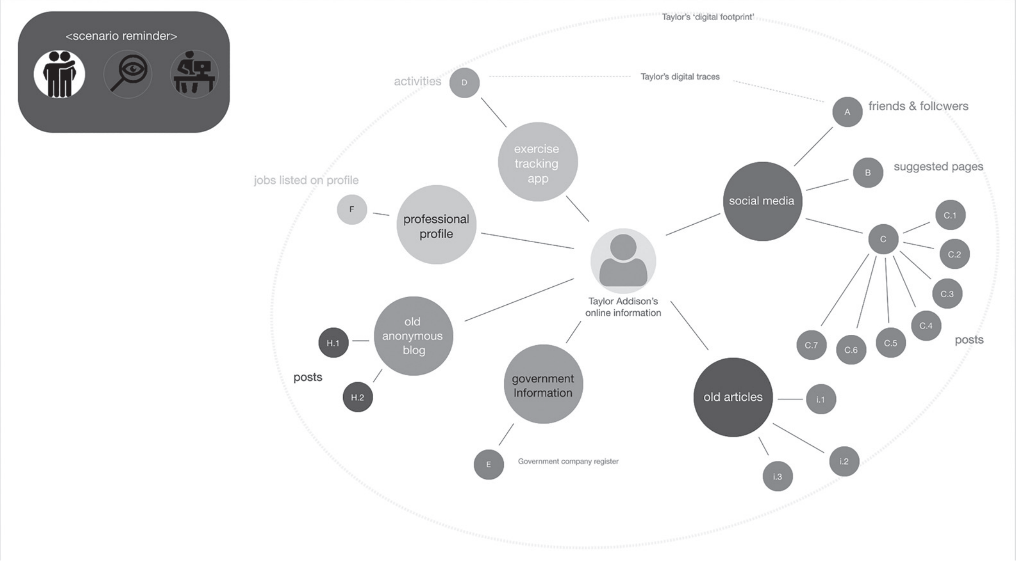


Image 5: A digital traces visualisation

Seeing the risk associated with cumulative data from "the other side" allows, we hope, safer data practices. But ultimately, we aim to use the same interface to allow users to generate their own profiles, compare them to those they studied "adversarially" and thus turn what was learned directly into action. Persistence of data, just as persistence of toxins in the environment, poses unique regulatory challenges. The "right to be forgotten" in this analogy becomes a tool to "clean" the data ecosystem, but exercising this right requires the data subject to keep track of their interactions in much more systematic ways than is currently feasible. It requires a new way to think about risk. As this process is less intuitive and more cognitively demanding than "atomistic" risk assessment, it also requires intelligent tools as support that lower that burden.