Compelling truth

COMPELLING TRUTH: LEGAL PROTECTION OF THE INFOSPHERE AGAINST BIG DATA SPILLS .

Burkhard Schafer, School of Law, The University of Edinburgh, Old College, Edinburgh EH8 9YL

Abstract:

The paper explores if legal and ethical concepts that have been used to protect the natural environment can also be leveraged to protect the "Infosphere", a neologism used by Luciano Floridi to characterize the totality of the informational environment. We focus in particular on the interaction between allocation of (intellectual) property rights and "communication duties", in particular Data Breach Notification Duties.

Keywords Data Breach Notification**;** Copyright, totality operators; big data

I) BIG DATA, NEW OIL, BIG NEW PROBLEMS?

Big Data has been frequently characterized as "the new oil" (for examples see Hirsch 2013). This comparison is inevitably meant to be positive and encouraging, extolling the audience to fully embrace the new technology and the opportunities that it brings. Toonder's characterization can stand in for many others that made claims along similar lines:

> "Data in 21st Century is like Oil in the 18th Century: an immensely, untapped valuable asset. Like oil, for those who see Data's fundamental value and learn to extract and use it there will be huge rewards. [...]Without it, progress would halt."

The imagery that these statements try to evoke is of rich but underexploited potential for improved products, services and profits, potential for wealth generation and a better future for everybody. Just as oil was the energy source of the 19th and 20th century, the power behind the current industrial revolution is data; the combustion engine is replaced by the silicon chip.

However, even a cursory glance at the history of oil exploitation and the fossil fuel economy triggers a gestaltswitch. Big data may well be the new oil, but that means we should better be prepared for substantial environmental damage caused by massive spills; long lasting environmental degradation due to the premature use of untested and unsafe extraction technology (Ko and Day 2004); forced dislocation or destruction of small, indigenous communities (Gerlach 2003 p. 55); threats to the democratic system due to a combination of money and economic power held by a small number of large monopolies (Lloyd 1894; Hanne 2009); and political dependency on foreign states with questionable human right records (Le Billion & Khatib 2004). For all of these "old" dangers that came with the fossil fuel based

industry, we can easily find equivalent dangers in the data driven economy. The Data Barons of the 21th century are as ruthless in their ability to sideline or defeat political, legal and wider societal concerns in their quest to accumulate more and more information about us.

But is the equivocation between data spills and oil spills a mere metaphor of limited practical use, at best a pedagogical tool, at worst a polemical device to warn against unrestrained optimism when approaching the new paradigm? Or are there deeper structural similarities between informational and environmental hazards, so that we can draw relevant inferences from the failures and successes of environmental protection and environmental ethics to the question of the appropriate law and ethics of big data?

The paper will claim that there are indeed valuable lessons that can be learned from environmental regulation and environmental ethics for the regulation of the data society. Luciano Floridi's concept of the infosphere will be used as a broader theoretical framework to justify this analogy, (intellectual) property law and (data) spill notification duties will be the two specific examples that we transplant from environmental to data protection law. One will come from the field of civil law (ownership), the other from the field of public law, more specifically "regulatory criminal law" as the enforcement arm of administrative law. Taking this dual perspective serves two purposes. First, it ensures that the discussion is relevant for a number of legal traditions and approaches, accommodating both the US approach with its emphasis on property rights and the European approach with its greater trust in the state and its regulatory organs. Second, the analysis will show that despite their differences, civil and (regulatory) criminal law, EU and US approaches are experiencing changes that show some recurrent themes and structural properties. I hope to show that these similarities can be explained as a readjustment of our legal systems as a whole of mutually interdependent parts to the conceptual changes that the fourth industrial revolution and the "virtualization of reality" are bringing. In particular, I aim to show a duality between information rights and information duties as an ordering principle of the law of the infosphere.

The paper will discuss in its civil law part efficient property allocation for environmental protection very briefly, and with the main aim to make the case for a common good conception of privacy that requires communal or "stewardship" models rather than individual ownership regimes. These typically combine property rights with sets of "care duties". In the second section that takes an example from regulatory criminal law, the paper interprets Data Breach Notification duties as one such duty towards the shared informational environment. Ownership in data through an IP right regime gives the owner a bundle of *rights*, with the hope that this will *encourage* good data stewardship. Their role is largely preemptive. Data Breach Notification Duties by contrast impose a bundle of *duties* in the hope that this will *deter* bad data stewardship. Their role is largely reactive. Together, they from a specific type of *Totality*, the sum of all communication rights and duties –they are "the whole truth" that can be said about that aspect of the infosphere. We will see that talking about totalities of this type is a recurrent aspect of our analysis, making it desirable to explore the logic of this term in more detail.

2. BIG DATA BARONS

Regulating the Internet and the new data economy has proved to be a difficult, frustrating and often elusive task (Brown and Marsden 2013; Guadamuz 2013). While claims of the demise of the nation state and the emergence of a libertarian utopia were premature, protecting the safety of citizens online from criminal attacks or unintended damage through negligent service providers remains a significant challenge. Interlinked with is the protection of their privacy in from the preying eyes of private and state actors. It is widely accepted that traditional modes of regulation alone are insufficient to guarantee safety and respect for basic human rights in cyberspace. New forms of regulation are frequently proposed, in particular "regulation through code", by designing protective norms directly into the communication infrastructure (Brown 2001). But cyberspace is not the only context where traditional regulatory tools have reached their limits. Preceding the discussion on Internet regulation by more than two decades, concerns over the protection of the physical environment have identified very similar concerns to those posed by the Internet.

The natural environment is the paradigmatic example of a complex adaptive system, which makes it difficult to predict the effect of any top down legal intervention (Ruhl 1997). Similarly, the Internet as a complex adaptive network often defies "single point" regulatory attempts (Guadamuz 2013). Effective environmental protection is frequently a transnational challenge, with the polluter often located outside the borders of the country where the pollution causes the most damage. Effective environmental regulation and Internet regulation requires substantial technological and scientific expertise, which will often require judges to defer to expert determination or have regulators resort to technological solutions over legal remedies. The combination of these pressure led increasingly to the insight that despite a growth in the quantity of environmental legislation, the positive effect on the quality of our environment remained often weak, leading to an increased recognition of the limitations of traditional regulatory law (Wiethoelter 1986; Ackerman and Stewart 1986; Adam 2010).

These concerns about the limits of law's regulatory capacity led some to argue for de-regulation and minimal legal intervention. Law then is only facilitative of markets and scientific expertise (Rittich 2003 727ff; Kennedy 1973 351-383).

Reliance on markets requires efficient allocation of property rights. Undoubtedly, some success was achieved through environmental protection based on private property regimes, in particular NGOs acquiring land ownership for conservation purposes (Pasquini et al 2011; Kiesecker 2007) But in other cases the intrinsic limitations of a private property regime limited the success of otherwise beneficial privately owned and administered environmental protection schemes (LoBue and Udelhoven 2013) or even had detrimental effect (Cabral and Aliño 2011). A particular problem is the transient nature of these private property solutions, which can come to an end at any time the owner so choses (Langholz and Krug 2004). This insight had let to an increased recognition of the role of non-traditional property conceptions for the protection of the environment, in particular non-western conceptions of ownership by local and indigenous communities that are better understood as "stewardship" relation than the bundle of economy exploitation rights (Alcorn 1993, Brechin

et al 2002; Ross 2011 chap 6). While the empirical validity of claims of an intrinsically more environmentally conscious indigenous mindset have come under increased scrutiny (Alvard 1993; Fennel 2008), the general concept of non-traditional, communal property rights vested in local groups and linked to corresponding stewardship duties has increasingly found recognition by international bodies, and proved its usefulness as an additional mechanism for long-term, sustainable environmental protection (Berkes 1998; Arnold 1990; Agrawal 1999). Pooling of resources makes this approach inherently stronger than individual property rights in the face of adversity from large, well- financed corporations, while at the same time providing long term perspective and strategy that transcends the life span of individual community members. This has by now found recognition from international bodies and law makers such as the International Union for Conservation of Nature and Natural Resources (IUCN). IUCN recognizes four models of governance under its Protected Area System: governance by government, by shared governance, by private governance and governance by indigenous people and local communities. Private governance covers both individual, NGO (cooperative) and corporate ownership models, while indigenous "ownership" remains as a separate category to express the fundamental conceptual differences between the ownership concepts. (IUCN Guidelines, 2008 p.39ff) The latter often includes complex systems of joint responsibilities and benefit sharing (Larsen and Oviedo 2006), rooted in traditional systems of ethics, religion and cultural practices.

Three final points are needed to help us make the transition from a discussion of the physical environment and its protection to the protection of the informational environment. First, stewardship property regimes for environmental protection *are* intimately linked to the concept of knowledge and information. One reason for the failure of tradition government-led preservation efforts was lack of local knowledge and lack of understanding of the systems that the law tried to regulate. The stewardship system by contrast tries to develop strategies to maximize the use of traditional, local knowledge and information about the environment for its protection. Indigenous land property rights and intellectual property rights become essentially intertwined (Gadgil, Berkes and Folko 1993; Coombe 2005). Indeed, it could be argued that in some indigenous societies, certain type of knowledge is constitutive for land property rights – the group that holds the knowledge of certain ceremonies and rituals also controls access to the land (Rigsby 1999), destruction and oppression of these knowledge structures also destroy the economic and political control over the land (Pettipas 1994). Conversely, destruction of the land can in these belief systems also equate to a destruction of knowledge – flooding of burial grounds e.g. is deemed in some communities to destroy the knowledge of those buried there (Whitt, Roberts, Norman, and Grieves 2001).

Second, integrating local communities into the environmental protection effort, requires the free, prior informed consent of indigenous peoples (Durban Recommendation 5.24):

• "Free" means the consent has not been imposed or manufactured and is the result of voluntary consultations and negotiations

• "Prior" means the consultation and information process is sufficiently in advance to leave time for the indigenous communities to deliberate and decide

• “Informed” requires full disclosure of information in a form accessible and understandable to indigenous communities

• “Consent” requires full indigenous participation in all major phases through their own freely chosen representatives including the right to access relevant technical/ legal support (Motoc and Tebtebba 2004).

Consent plays of course also a significant role in data protection law, and the environmental conception of consent is if anything more elaborate and substantive than even the (by international standards) substantive conception of consent in the upcoming European General Data Protection Regulation. Consent in this model is a process that extends through time and is linked to a deliberative process of institutionalized information sharing.

Third, data is in (at least) one significant sense different from physical goods. Exploitation of data can be non-rivalrous. While only one person at a time can eat a specific piece of fruit (Locke’s example of property acquisition through labour+digestion) or ride a specific bike, indefinitely many people can use “the same” piece of data simultaneously. This does not harm the analogy that is presented here, on the contrary, it removes potential sources of conflict *within* the community, and should simplify the governance structure.

We can now make explicit the analogy on which this paper is build. State-led environmental regulation, typical for the early stage of environmental protection, focused on non-negotiable access restrictions to protected vulnerable biotopes and landscapes. Market based and technocratic solutions were promoted to address perceived failures and inefficiencies of this approach, by focusing on giving individuals control (and in this sense empowerment) through property rights. However, this put considerable burdens on under resourced (in terms of money, time, skills and knowledge) individuals vis-à-vis powerful commercial interests. Finally, indigenous communities became more closely involved in the conservation efforts, through a recognition of non-western, “stewardship” models of ownership. As property based solutions, they still emphasize the aspect of control, but do not locate property any longer in a libertarian understanding of rational, self-interested individuals maximizing their personal gain, but instead are based on “thick” notions of communal solidarity. Furthermore, the environment itself can become something that is owed a duty to, and not just merely a subject of exploitation and control.

In the field of data protection law, we can find a similar pattern. We find state-centric approaches that create (more or less) non-negotiable access restrictions and access control, mitigated by informed consent, an approach particularly strong in the European Union. In competition and contrast to this state-centric approach, we also find a US-led market-centric approach that invites us to think of personal data as a form of property of the data subject, similar to copyright. In this approach too, the aim is to “empower” users by giving them property rights and (market transfer) control over their data (so e.g. Zittrain 2000). More recently, Rubinstein proposed ‘Personal Data Services’ as a new business model to bring about a shift towards a “user- centric” concept of data protection, which seeks to “integrate diverse types of personal data while putting end users at the centre of data collection of use, subject to a set of global data principles that include transparency, trust, control and value creation” (Rubinstein 2012). However, just as with their environmental counterpart, these

market-based, private property solutions struggle to achieve adequate protection for two reasons. First, they ignore the very significant difference in bargaining power, resources and technological know how between users and the companies that control their data. The economic value of the data of any individual Facebook user is negligible, putting them into a very weak bargaining position, with little incentives to invest the time and resources to maximize the benefits from this new property and therefore "to look after it" carefully. Only collectively would they be able to engage with large data controllers as equals – pointing towards a collective conception of data property rather than an individualist one. This is an idea that was first proposed by Edwards (2004). The proposal here is in many ways similar to Edwards. The main difference is the stronger element of control that the community has over its stewards (who will typically be elected office holders and directly responsible to their constituency) than the control and degree of active involvement that a beneficiary of a common law trust has vis-à-vis the trustees. But with this control also come duties – in this case duties of active participation. A better analogy than trusts might be the pooling of resources in a union or a credit cooperative. While there will be members that chose to be more passive than others, they all elect collectively their leadership and office holders from amongst their ranks, and exercise control over them. This is important for the overall thrust of this paper that argues that we must not only rethink how we allocate communication rights and duties (and a "recall" or "vote of no confidence" is a communication for our purposes), but move away as far as possible from the focus on isolated individuals to democratic communities.

Following on from this, both the state-centric access control model and the market –centric property model see privacy primarily as an individual good. In both approaches, this incentivizes the right holder to maximize benefits in the short term, rather than looking for sustainable long term solutions. But this individualistic understanding of privacy misunderstands the importance of privacy as a political and societal good, which enables other democratic values and processes. Privacy empowers individuals to criticise and resist acts of government that are of an undemocratic nature. Bloustein e.g. (1964 p1003) argued that

"[t]he man who is compelled to live every minute of his life among others and whose every need, thought, desire, fancy or gratification is subject to public scrutiny, has been deprived of his individuality and human dignity. Such an individual merges with the mass. His opinions, being public, tend never to be different; his aspirations, being known, tend always to be conventionally accepted ones; his feelings, being openly exhibited, tend to lose their quality of unique personal warmth and to become the feelings of every man.

And indeed, the experience in many totalitarian regimes has shown that an absence of privacy has the potential for creating a "society of followers".

This interdependency between the protection of privacy and the protection of other essential features of a democratic society is also highlighted by Raab (2011) who argues that values like personal autonomy and self-determination

"are important not primarily because individuals may wish to live in isolation (for they do not, mostly), but so that they can participate in social and political relationships at various levels of scale, and so that they can undertake projects and pursue their own goals".

Just as a private property approach to environmental protection can lead to individuals cashing in on short-term advantages for themselves, while leaving the collective interest in a clean and diversified environment harmed, privacy theories that emphasise the notion of free alienation of privacy in market places, risk lasting harm to the common good. As Regan (1995 p. 233) argues, there is a risk that

"[i]f one individual or a group of individuals waives privacy rights, the level of privacy for all individuals decreases because the value of privacy [in the collective view of society] decreases".

In the privacy discussion, this has led many European commentators to be deeply sceptical of the very concept of a property/copyright approach to personal data (e.g. Purtova 2010) . They fear, not unreasonably, that once we think of private data as a property at all, this will lead to even more ruthless "trading" of data, where data subjects are paid a pittance. It is here though that we can learn most from the comparison to environmental protection.

Structurally identical problem in environmental protection led as we saw to the inclusion of collective, communal stewardship conceptions of "property in the environment" to the regulatory toolset. If we now take Rubinstein's notion of Personal Data Services as an "ecosystem" serious, then we can see the space for a communal notion of data –ownership that preserves the advantages of her property-centric approach, but without the inherent disadvantage of fragmentation and exploitation by economically stronger Data Barons. Facebook users, collectively, would then own the property in their data, administered through an independent, fiduciary structure (in the way indigenous land rights are typically protected through some form of elected political entity), crucially though not just for the benefit of the current group of members, but with a view on preserving a "clean" data environment also for future user groups and wider society. This will then have to be balanced against the expressed collective interest of the current owners who exercise a degree of control over the stewards. (In the same way one could say in which politicians balance the need to plan for future generations against the short-term goals of their electorate). While Zittrain asked privacy lawyers to learn from copyright lawyers, both may be able to profit from non-western conceptions of property that have been leveraged in environmental law.

It goes beyond the scope of this paper to develop such a "stewardship theory of data ownership" in any detail. As noted above though, the notion of "stewardship" means that in addition to use and exploitation rights, these new property titles also come with sets of new duties and obligations. As a first indicative example, we can see the emergence of such a stewardship model in the "replicable computation" movement. "Replicable" or "reproducible" computing concerns itself with the way research results are communicated to scientific peers and wider public. It is driven by the recognition that the traditional publication and peer review process is failing as a means to ensure quality, reliability and robustness of results. The more people are put in a position to test and replicate research,

the more false or premature results will be identified and weeded out. Staying with our analogy, we can think of any such misleading study as a pollutant of the informational environment, and the *collective* activity of testing and replicating as a clean-up. Copyright, if understood mainly as a traditional, individual right that prevents others from reproducing material, and hence an access barrier, seems prima facie inimical to this endeavour. However, Victoria Stodden amongst others has shown how a bundle of creative commons licenses for academic papers, software and data-sets together can facilitate replicability (Stodden 2009). This approach to copyright is intimately linked to a notion of community and its values – here the creative commons community and also the scientific community – and uses property rights to protect the openness of the data in perpetuity, secure from "land grabs" by say commercial publishers. Intellectual property thus changes form an access barrier to an enabler of sustainable information sharing. The communal values of the scientific community – in particular the specific type of rationality that comes with the concept of criticism and testing – then ensures that the shared research is subject to a much more rigorous post publication peer review than otherwise possible, thus leading to a collectively assured "clean" information environment. Transparency becomes intimately tied to the quest for a "clean" information environment. Laws such as copyright law can *enable*, so Stodden, this type of transparency, by creating a *right* to share (Lyon 2016). Other laws by contrast create a *duty* for information sharing, and these we will look at in the next section.

Many questions remain unanswered here, What is attempted is however to provide a conceptual framework that allows these questions to be asked. How would, in a model like this, the duty to only provide accurate data about themselves (as the opposite could be seen as "data littering" the info-environment), which would be a problem for obfuscation based methods of privacy protection.

3. SPILLING THE BEANS ON DATA SPILLS

Having discussed what we can learn from environmental law for private law based approaches to data protection, we now look at regulatory tools that use fines and similar punishments to prevent abusive behavior – regulatory crimes in environmental and data law. This section begins by an account of two significant data spills and the regulatory lessons that can be learned from them. Oil spills can trigger a duty to inform regulators about the spill, even if this means potential sanctions and fines. Similarly, many jurisdictions introduced a duty to inform regulators, customers or the general public of data spills: Data Breach Notification Duties. At a first approximation, DBND laws fall into a group of regulatory tools that Cass Sunstein called "regulation through disclosure"(Sunstein 1999). Regulation through disclosure has two effects. It gives incentives to the party who owes the duty to minimise the number of triggering events and avoid public embarrassment through safer procedures. Secondly, it enables affected third parties and the wider community to make prudent choices, take protective actions and bring market pressures to bear on repeat offenders. DBNLs perform both tasks, making them therefore also part of the "choice architecture" (Leonard, Thaler and Sunstein 2008). Schwarz and Janger (2007 p 915) thus

compare them to the duty of factories to disclose information about toxic releases of one's factory.

In early 2005, the data broker ChoicePoint became aware of a potential security breach. The pattern of activity indicated that the data had been used for the purpose of ID fraud. The personal data of 163,000 customers had potentially been compromised (Anon and David 2003). Ten years later, in August 2015, Ashley Madison, the extramarital-affair brokerage website, became the victim of a successful hacking attack by a group calling itself the "Impact Team". Potentially up to 32 million user profiles were affected, though this may exaggerate the number of victims, as one of the stated objectives of the hack was to expose "fake" profiles, which according to hackers were plentiful. The Ashley Madison (AM) data contained names, email addresses, physical addresses, login information, and partial credit card payment information. The consequences for many of them were devastating, driving some customers into suicide (Segall 2015).

While in both scenarios, customers of online service providers were put at risk when their personal identifiable data was leaked after an attack on their OSP, the legal protection offered to the affected groups of customers differed in one significant aspect. ChoicePoint, being an US company, informed initially the police of the suspected data breach, and soon after in compliance with the Californian Data Breach Notification laws also its 35,000 Californian customers. This limited release of information resulted in a public outcry, and the company soon afterwards approached all US customers whose records had been improperly accessed. While ChoicePoint accepted a duty to tell the truth about the danger it had created for its clients, it tried to avoid telling the whole truth as long as it could.

Ashley Madison customers were even less lucky. Being headquartered in Canada, AM was not subject to a similar legal obligation. When the "Impact Team" hackers initially approached AM, threatening the release of the data unless its parent company shut down Ashley Madison, AM kept the breach secret, hoping undoubtedly that the hackers would not act on their threat. They miscalculated, and it was through a statement released on the Internet by the hackers that AM customers learned about the risk of exposure that they were now facing. This prevented many AM customers from taking timely mitigating action – from coming clean to their spouse to changing their credit card.

The two cases raise interesting philosophical issues. In both of them, we find appeal to "truth telling duties", both moral and legal. Data breach notification laws had imposed on ChoicePoint a duty to tell the truth about data breaches to their customers, even though this was likely going to be a statement against the company's own interest. ChoicePoint complied, hesitantly, telling the truth, but never quite the whole truth. AM, at the danger of sounding puritan, was in the business of facilitating lies, enabling its customers to cheat on their partners. But it was not this that attracted the ire of hackers, but the suspicion that in doing so, they also lied to their own customers, or at least facilitated and benefited from a lie: using inflated numbers of (female) profiles to attract new (male) subscribers, even tough they new, or should have known, that many of the profiles were fake. Finally, AM lied by omission, keeping its customers unaware of the risk they were facing, and encouraging as a result continuous use of AM's services, disclosure of which could later harm them. If

translated back into the language of environmental damage, AM allowed degraded substances to accumulate (false profiles) and also allowed toxic information to leak into the environment On all three counts, AM were bad data shepherds who failed in their stewardship duties towards the data ecosystem that they had created.

Data Breach notification laws aim two things: mitigating the danger after a spill, but equally important, incentivizing data holders to prevent spills form happening in the first place. In 2002, California became the first country to enact a Data Breach Notification law, S.B. 1386, which took effect in July 2003. The law requires

> "a state agency, or a person or business that conducts business in California, that owns or licenses computerized data that includes personal information, as defined, to disclose in specified ways, any breach of the security of the data, as defined, to any resident of California whose unencrypted personal information was, or is reasonably believed to have been, acquired by an unauthorized person."

As of February 2016, forty-seven states, the District of Columbia, Guam, Puerto Rico and the Virgin Islands have enacted legislation Data breach notification laws (see Needless 2009; Lee 2006). However, the absence of a general tort of privacy intrusion and the difficulty for individuals to use the court system and prove they suffered individual harm means that this approach suffers enforcement issues (Winn 2009 1134):

> “Attempts to establish a right to damages following receipt of a security breach notice through class action lawsuits have generally only succeeded in clarifying the degree to which no such right exists, although many businesses suffering breaches have chosen on a voluntary basis to provide their customers with credit monitoring services to reduce the risk of harm from identity theft.”

As in our discussion above, collective, not individual conceptions of privacy are needed to address this deficit. In the European context, the situation is different. Here, in particular the upcoming General Data Protection Regulation will not only reaffirm the general right to information privacy, but also provide an enforcement mechanism. Europe’s attitude to DBNLs is a typical example of a “society centric” approach to mandatory breach disclosure that goes beyond remedies for any individuals customer who suffered harm. Directive 2009/136/EC created a two-tier framework for the reporting of data breaches. It imposes duties to inform both regulatory authorities *and* affected individuals.

The proposed Amendment to the draft General Data Protection regulation which will further strengthen this approach phrases the required content of the notification in an interesting way (our emphasis) in its Art 32:

> “The communication to the data subject referred to in paragraph 1 shall be *comprehensive* and use clear and plain language. It shall describe the nature of the personal data breach and contain at least the information and the recommendations provided for in points (b), (c) and (d) of Article 31(3) and information about the rights of the data subject, including redress.”

All breaches have to be notified to the authorities, but only those that involver personal data require notification of the person affected. Unlike the US laws, this requirement does not only affect subscribers or customers of the service provider, also affected third parties can benefit from this duty. Notifying affected non-customers can be particularly difficult. As the regulation notes:

> "While providers should have contact details of their subscribers, given their direct contractual relationship, such information may not exist for other individuals adversely affected by the personal data breach. In such a case, the provider should be permitted to notify those individuals initially through advertisements in major national or regional media, such as newspapers, to be followed as soon as possible by an individual notification as provided for in this Regulation" (Commission Regulation (EU) No 611/2013 at (14))

Importantly for our discussion, the duty of the service provider is not just an expression of a contractual bond between the provider and the affected data subject, duties between individual right owners. Rather, the duty is (also) owed to the wider public, protecting the collective interest in privacy as a common good.

Importantly, as we saw the Directive also entails an active duty by the provider to establish "the whole truth" of the breach. The hacked service providers are not just passive reporters of information that they have obtained. This establishes a rather unusual form of legal duty – a duty of a victim of a crime not only to report the event, but also to a degree to investigate it. We will return to this issue below.

Finally, the Directive adds a number of reporting obligations for providers, who will have to maintain an inventory of data breaches. This inventory has to note all the relevant facts that the provider was able to establish, the effects of the breach and any remedial action that was taken. Regulators will use the information gained from these registers to evaluate the efficiency of existing legislation and guidelines, and to develop best practice. We noted above that the European approach extends the scope of notification duties beyond customers to third parties whose data was held. Here we see this idea extended further – the beneficiaries of this aspect of the notification duty is not any individual who suffered harm, but through improved efficiency of the legal system, everybody.

We can now extend our environmental analogy. As we saw above, in stewardship models of land ownership a *collective* responsibility arises to "be concerned about" and "look after" the environment, something that can best be done if information about problems is spread widely. A duty to report oil spills e.g. will give an incentive to companies to invest in security equipment, since reporting a spill harms at the very least the reputation of the company, at worst exposes it to litigation. At the same time, it allows agencies and citizens to take appropriate action to mitigate the consequences of the spill, be it by direct involvement in the clean up, or by reinforcing the deterrent aspect of notification duties through a boycott call or similar use of market tools. Just as spill notification duties protect the natural environment, the biosphere, so do DBNLs protect the information environment – or as Floridi terms it, the infosphere (Floridi 2002). In Floridi's conception the infosphere has intrinsic value, which means in particular that we may have duties towards its constituent

parts, the “inforgs”. Inforgs, the “inhabitants” of the infosphere are entities made up of, processing and transmitting information. Individual humans are one example, but the collective of communication tools of a family also qualifies, as do indeed non-human and non-biological information creating entities – of particular importance if we consider ta quantitatively, already now most information is generated through machine-to-machine ienteraction.. Not only then do we have “information as environmental regulation” (Esty 2004), in Floridi’s framework we can now draw the inverse connection as well, and think of DBNL’s as “environmental protection regulation *of* information”, regulation that forces us to create a “clean”, i.e. truthful and transparent information environment.

4. PROTECTING THE DATA ENVIRONMENT – A DUTY FOR ALL INFORGS?

DBNLs carve an exemption into the right against self-incrimination, but as we saw, in this they are but part of a general trend of regulation through disclosure. In this sense, they are indeed very similar to duties to report oil spills or other failures by companies to ensure a safe and clean environment. However, they also go beyond this duty in a crucial sense. Ashley Madison did not only act carelessly with its customer data. They also were victim of a crime themselves, the hacking of their system, which ultimately will cost them customers, trust and business.

Imposing duties on the victim of a crime, potentially against his/her own interests and under the threat of criminal law sanctions, makes DBNLs conceptually unusual. Normally, we think of the victim as the passive party, wronged by the criminal, with a duty by the state to right this wrong. But when do victims have a moral duty to enable the prosecution of the crimes against them, and under which conditions is it appropriate, for a democratic republic under the rule of law, to turn this moral or civic aspiration into a legal duty, with criminal law sanctions in case of non-compliance? To answer these questions, we need a general theory of the relation between citizens and the criminal law, a theory that asks what actions are the appropriate subject of criminal sanctions, and what duties and obligations can the criminal law rightfully impose on the citizenry. With other words, the question needs to be answered within a general theory of (criminal) law, a theory that outlines the totality of rules and obligations that the criminal law creates.

Grand legal theory in the tradition of Austin, Kelsen or Hart struggles to provide us with an appropriate analytical framework for this task. Their gaze is primarily on the “officials”, the legislator as the source of law, the judiciary as its interpreter, and possibly other officials such as police for its enforcement. The citizen is in these systems a largely passive recipient of commands, his/her duties merely to obey the sovereign, or, in Hart’s version, obey those officials that they recognise as legitimate. Recast in the language of information, the information flow in these systems is (largely) uni-directional, from the top of the pyramid to its base. The criminal law is something external to the citizen, a body of rules laid down by others for them to obey (Duff 2010 p. 300).

The shortcomings of this conceptual approach have been recognised amongst others by Duff and Marshall in a series of papers on the role of the criminal law, the criminal trial and criminal punishment (see Duff 2010; Duff et al 2015). The traditional model, so they argue, cannot account for all those characteristics of the criminal trial that do not fall into the neat divide between officials and citizens. A broad range of phenomena, from the duty of citizens to serve as jurors, to the specific duties of witnesses as "truth speakers", or the role of criminals in their own rehabilitation, thus remains unaccounted for. It therefore gives at best a partial account of what it means to be a "system of criminal law". Duff and Marshall propose an alternative approach centred around distinct roles, connections that people can have relative to a crime, its investigation and ultimately, its punishment. In particular, they suggest that we can distinguish the following roles a citizen may play in the practical process that is a crime investigation and prosecution (Duff 2015 p3-4):

1) official and professional roles, where the citizen "in uniform" is employed by the polity and in this role given special powers, responsibilities and privileges. These include

- police officers
- judges
- prosecutors
- civil servants involved in legislative drafting

2) official roles filled by lay participants. Here too the polity endows the citizen with special rights and duties, but they serve not as employed professionals. These include amongst others

- jurors (the paradigmatic example)
- lay judges, justices of the peace etc
- special constables

3) the final type of roles are those that a citizen acquires by his or her connection to a particular crime. These so Duff and Marshal suggest include amongst others

- victim
- witness
- offender
- suspect/defendant
- 'ex-offender'

This framework then allows us to ask if our vision of the criminal law is compatible with recognising any of these roles in law – *should* we use e.g. lay people as police officers? It also allows us to ask precisely under what conditions we *should* or at least legitimately *can* turn a civic duty into a legal obligation. And finally, it allows us to ask how we can resolve conflicts between these roles while staying true to our vision criminal law in a democratic society.

As we saw above, being able to talk about specific roles also enabled us to place DBNLs into a first conceptual framework. DBNLs call upon online service providers in their role as professional businesses with special expertise, on their role as witnesses of a crime, and on their role as a victim of a crime. In attaching specific legal duties to these roles, we also saw how they are asked to take on functions that have previously been reserved to certain

officials, in particular the investigative duties they incur to be able to inform their customers about the "whole truth" of the breach. Here we see the problematic aspect of their involvement: As experts, they are tasked to investigate, as witnesses, to testify truthfully, while as victims and as enablers, they may be entitled to considerations that limits these duties.

Duff and Marshal integrate their role-based scheme within a wider communitarian theory of criminal law in a democratic society, which aligns their approach with communitarian conceptions of ownership that we saw in the first part for non-western property conceptions. They also link it to a specific conception of the trial as "public holding to account" (Duff 2001). As noted above, we do not normally in law require active participation of victims of crime in the prosecution of their aggressors. There are however a number of more or less isolated exceptions to this rules that can be explained in this combination of communitarian and communicative conceptions of the trial. One such example is prosecutions against victims of domestic violence who withdraw their allegations during the court proceedings. Marshall shows how we can make sense of this within a communication centric, communitarian model of the criminal law. Victims are not just witnesses; they are a *specific type of* witness, one that is irreplaceable even when there are other sources to ensure a conviction. In the witness, the harm done to the community is personified, his or her testimony therefore necessary to communicate the wrongness of the action to judge, public and indeed also the perpetrator (Marshall 2004). Marshall (2015 p. 299) writes:

> "A trial thus calls a defendant to answer not just to an alleged individual victim, though it does indeed do that, but also to the whole polity for the wrong that he allegedly committed; and it constitutes, in part, an expression, articulation, and application of what are purported to be the shared, 'public' values of the polity."

The communication acts that are performed as part of a trial are therefore identity-creating and constitutive. Again Marshall (2015 p. 298):

> "To understand the nature of law we have to understand its role as partly constitutive of a political community and therefore as an object for identification, as playing an important role in a people's sense of who they are. [...] Citizens are responsible both *to* the state (insofar as they are responsible to one another in their roles as citizens) and *for* the state."

This communicative conception of the trial enables us to connect their account to the information ethics of Floridi. Floridi's conception of the "Infosphere" is that it extends beyond cyberspace, as a general theory of information, and can thus accommodate pre-Internet institutions, processes and practices. Its sensitivity to the way ITCs "re-ontologise" the world in turn matches the constitutive aspects of information practices during the trial that Marshal emphasizes. If the Infosphere, then, as totality of all informational entities, their properties, interactions, processes and mutual relations is the whole of Being (Floridi 2010 p. 8-10), then we can say that the "whole of the law" becomes the totality of information duties that exist between informational agents. Victims are one type of informational agents, inforgs that hold information of a specific quality.

5. ALL THE FACTS, MA'AM

The paper will finish with a summary of what was achieved, and an outlook for future research.

In the past, legal systems have used, with mixed success, private law solutions (including data as property) and (regulatory) criminal law/public law to protect data subjects.

Protection of the environment too uses both approaches, but in each case with a significant difference: it replaces the focus on largely passive individuals who use rights as a protective shield to more complex communities where parties pat active roles and have certain communication duties towards others, in addition to rights.

In the property law case, this is the duty to participate in the communal decision making process that establishes collective consent. In the criminal law case, it is the duty of victims to participate in the communal communication process that is the trial

The paper so far has argued that the ethical treatment of Big Data can be informed by experience made with the regulation of environmental hazards. In both cases, institutionalizing the right type of communication duties is essential. We looked at two legal instruments in particular, the institution of private property, and breach notification duties in administrative law, as complementary strategies. Property laws, if recast as communal, stewardship models of property, can *enable* the right type of communication and data sharing, while giving incentives to "look after" the environment in perpetuity. Breach notification duties create duties to communicate. Together, they give us part of a regulatory framework where the granting of property rights comes with new, and admittedly burdensome, duties. In both cases, our approach relied on a substantial concept of community and community values. In the case of ownership, we argued for collective control over the assets; in the case of DBNLs, we argued for a duty to inform the whole community of a breach. Appropriate regulation of Big Data, just as appropriate regulation of the environment, works best within a framework that goes beyond individuals and their interest.

As an avenue for fruitful future research, capturing the precise nature of "truth telling duties" emerged as a key theme. In our analysis, we have also seen that on various levels, appeals to "the whole truth" were made. On the most basic level, we find appeals to the whole truth in the legal doctrine. Just as a witness in court, the victim of a data breach has the duty to report "the whole truth" of the breach. On a more abstract level, we saw how the trial itself can be seen as a communicative exercise, which involves state officials and citizens in a multi-directional exchange of information. Citizens owe to each other various communication duties, depending on their connection to the specific crime, as do legal officials. Victims owe subtly different communication duties to themselves, to the legal officials, to the jurors and indeed to the perpetrator to give a truthful and complete account of what happened. Similarly, judges have communication duties, in particular the duty to

give reasons, that hold towards victim, accused/convict and the wider public. But can vary in intensity, required communication mode or explicitness between these.

We contrasted this model with a more traditional jurisprudential approach to the criminal law, where communication between officials and citizens is uni-directional only, and based on simple commands by the sovereign to the citizenry. This latter view then captures only part of what constitutes a legal system, as opposed to "the whole" of law as the totality of communication rights and duties that citizens and officials owe each other. Finally, we indicated how this normative totality could be extended further still, by substitution "inforgs" for "citizen" and "officials". The resulting model of the infosphere, as suggested by Floridi, provides then the ontological underpinning of our attempt to combine concerns for the natural environment with concerns for the informational environment. The resulting totality encapsulates the whole of "Being", and with that both aspects simultaneously.

From a philosophical perspective, this raises a further question. How are we to make sense of the term "the whole truth" and, what, if anything, does it add to truth simpliciter? To give a precise formal account is interesting both for practical considerations, e.g. the development of software compliance tools that assist data holders to fulfill their duties under the DBNLs, and also for a more abstract philosophical analysis of what it means to have captured the totality of an issue. It is a natural thought that a description is complete, or is the whole truth, just in case it entails every truth. But this *entailment account* leads to a dilemma concerning the appropriateness of assertions of the form "it is the whole truth that p". Either p entails that it is the whole truth that p, or it does not. If it does, nothing is added to the bare assertion of "it is true that p" or even p, and the longer utterance violates Grice's maxim of manner. If p does not entail that it is the whole truth that p, however, then "it is the whole truth that p" is false. For if it were true, then there would be a truth – that it is the whole truth that p – that is not entailed by p. Failing to entail all truths, p would not be the whole truth after all, and the assertion would be false. On either horn of the dilemma, asserting "it is the whole truth that p" comes out as inappropriate. But that result is puzzling: clearly, as our analysis above has shown, there are scenarios where such an assertion seems neither false nor unduly prolix. The prefix "it is the whole truth that" does not seem to be redundant in the way that "it is true" has been taken to be.

In *Tractatus* 1.1.1, Wittgenstein writes: "The world is determined by the facts, and by their being *all* the facts. In recent years, philosophers have aimed to give a more precise meaning to this "quantifier over everything". This has involved efforts in metaphysics (Armstrong 1998; Chalmers 2001) and also crucially for us, ethics and the law (Holton 2010, 2011). Our discussion here indicates that these superficially very different appeals to totalities are indeed interrelated. Introducing a duty to "tell the whole truth" to a witness or victim calls upon both the metaphysical conception of "whole truth" (as a determinant of the precise content of the duties that the witness owes) and the ethical-legal concept (as a determinant of the form of that duty within an account of a complete legal system). Put differently, a more precise and ideally formally representable totality operator T* - "It is the whole truth that" would enable us to describe *what* the witness has to say, and at the same time allow us to better understand how this type of duty fits into the totality of communication obligations that actors (or inforgs) in a legal system have. The parallel reading of regulation

that pertain to the natural environment and regulation that pertains to the informational environment showed how the concept of a duty to tell "the whole truth" connected them both, and can thus also be seen as a first stepping stone for the development of a theory of a totality operator that spans law and metaphysics.

Competing interests

The author has no competing interests.

Acknowledgements

Stefan Leuenberger (University of Glasgow) and Martin Smith (University of Edinburgh) contributed substantially to the ideas that gave rise to this paper as part of a joint project on "the whole truth". Section 5 in particular benefited from the insights and ideas of Leuenberger. My thinking regarding "collective" administration of personal data was strongly influenced by discussions with Lilian Edwards, whose 2004 paper was the first to propose a similar solution.

Funding statement

This work was supported by the Arts and Humanities Research Council [grant number AH/M009610/1]

Bibliography

Ackerman, B. A., Stewart B. 1985 Reforming environmental law, *Stanford Law Review* **37** 1333-1365.

Agrawal, A. and Gibson C. C. 1999 Enchantment and disenchantment: the role of community in natural resource conservation, *World development* **27** 629-649.

Alcorn, J. B. 1993 Indigenous peoples and conservation, *Conservation biology* **7** 424-426.

Alvard, M. S. 1993 Testing the 'ecologically noble savage' hypothesis: Interspecific prey choice by Piro hunters of Amazonian Peru. *Human Ecology*, **21** 355–387.

Armstrong, D. 1989 *A Combinatorial Theory of Possibility,* Cambridge: Cambridge University Press

Arnold, J. E. M. 1990 *Social forestry and communal management in India*. Social Forestry Network Paper 11b, Overseas Development Institute, London

Berkes, F. ed. 1989 *Common Property Resources: Ecology and Community Based Sustainable Development,* Belhaven Press, London

Bille Larsen, P. and Oviedo, G. 2006 *Reconciling indigenous peoples and protected areas: rights, governance and equitable cost and benefit sharing*. IUCN Discussion Paper https://cmsdata.iucn.org/downloads/iucn_reconciling_ip_and_pa.pdf

Bloustein, E J. 1964 Privacy as an aspect of human dignity: An answer to Dean Prosser, *NYUL Rev.* 39 962-1007

Borgman, C. L. 2012 The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology* 63, no. 6 1059-1078.

Brechin, S. R., Wilshusen, P.R., Fortwangler, C. L. and West, P.C. 2002 Beyond the square wheel: toward a more comprehensive understanding of biodiversity conservation as social and political process. *Society &Natural Resources* 15, no. 1 41-64.

Brown, I., Clark, D.D. , Trossen, D. 2010 Should specific values be embedded in the Internet architecture? In *Proceedings of the Re-Architecting the Internet workshop*, ACM, doi: 10.1145/1921233.1921246.

Brown, I., Marsden, C. T., 2013 *Regulating code: Good governance and better regulation in the information age*. Boston: MIT Press,

Cabral, R. B., Aliño P. M. 2011 Transition from common to private coasts: Consequences of privatization of the coastal commons. *Ocean & Coastal Management* 54, no. 1 66-74.

Chalmers, D. 2010 The Two-Dimensional Argument against Materialism, in his *The Character of Consciousness,* Oxford: Oxford University Press.

Coombe, R. J. 2005 Protecting traditional environmental knowledge and new social movements in the Americas: intellectual property, human right or claims to an alternative form of sustainable development? *Florida Journal of Environmental Law* 115-135.

Duff, R.A. et al. 2015. *Criminalization: the political morality of the criminal law*. Oxford: Oxford University Press

Duff, R.A. 2010 A criminal law for citizens *Theoretical criminology* **14** 293-309

Duff, R.A. 2001 *Punishment, communication, and community*. Oxford: Oxford University Press.

Duff, R.A. 2010 *The boundaries of the criminal law*. Oxford: Oxford University Press.

Duff, R. A. 2015. Legal Reasoning, Good Citizens, and the Criminal Law. *Minnesota Legal Studies Research Paper* 15-18 Available at SSRN: http://ssrn.com/abstract=2618684 or http://dx.doi.org/10.2139/ssrn.2618684

Edwards, L., 2004. The Problem with Privacy. International Review of Law Computers & Technology, 18(3), pp.263-294.

Esty, D.C. 2004. Environmental protection in the information age. *NYUL Rev.* **79** 115-212.

Floridi, Luciano (ed). 2010. *The Cambridge handbook of information and computer ethics*. Cambridge: Cambridge University Press

Floridi, L. 2002 On the intrinsic value of information objects and the infosphere." *Ethics and information technology* 4: 287-304.

Gadgil, M., Berkes, F. and Folke, C., 1993 Indigenous knowledge for biodiversity conservation. *Ambio* **22** 151-156.

Gerlach, A. 2003 *Indians, oil, and politics: A recent history of Ecuador*. Lanham: Rowman & Littlefield.

Guadamuz, Andrés *2013Networks, complexity and internet regulation scale-free law*, London: Elgar

Fjelde, H. 2009 Buying Peace? Oil Wealth, Corruption and Civil War, 1985—99, *Journal of Peace Research* **46** 199-218.

Hirsch, D. D. 2013 The Glass House Effect: Big Data, the New Oil, and the Power of Analogy. *Me. L. Rev.* **66** 373- 396.

Holton, R. 2010 The Exception Proves the Rule, *The Journal of Political Philosophy* **18** 369-388.

Holton, R. 2011 Modelling Legal Rules in A. Marmor and S. Soames (eds.) *Philosophical Foundations of Language in the Law* Oxford:Oxford University Press 165-83.

Kiesecker, J.M et al. 2007 Conservation easements in context: a quantitative analysis of their use by The Nature Conservancy *Frontiers in Ecology and the Environment* **5** 125-130.

Ko, J. Y., & Day, J. W. 2004 A review of ecological impacts of oil and gas development on coastal ecosystems in the Mississippi Delta. *Ocean & Coastal Management*, ***47*** 597-623.

Langholz, J. A.,Krug W. 2004 New forms of biodiversity governance: non-state actors and the private protected area action plan." *Journal of International Wildlife Law and Policy* **7** 9-29.

Le Billon, P., & El Khatib, F. 2004 From free oil to 'freedom oil': Terrorism, war and US geopolitics in the Persian Gulf. *Geopolitics*, ***9*****(1),** 109-137.

Lee, S. 2006 Breach notification laws: Notification requirements and data safeguarding now apply to everyone, including entrepreneurs." *Entrepreneurial Bus. LJ* **1** 125-153.

Leonard, T. C., Thaler, R.H., Sunstein, C.R. 2008 Nudge: Improving decisions about health, wealth, and happiness." *Constitutional Political Economy* **19** 356-360.

LoBue, C., Udelhoven J. 2013 Private ownership of underwater lands in Great South Bay, New York: A case study in degradation, restoration and protection. *Marine Policy* **41** 103-109.

Lloyd, H. D. 1894 *Wealth Against Commonwealth*, New York: Harper and Brothers

Lyon, L., 2016 Transparency: the emerging third dimension of Open Science and Open Data. *LIBER Quarterly.* **25(4)** 153–171. DOI: http://doi.org/10.18352/lq.10113

Marshall, S. E. 2004 Victims of crime: Their station and its duties. *Critical Review of International Social and Political Philosophy* **7** 104-117

Marshall, S. 2015 It Isn't Just about You. Victims of Crime, their Associated Duties, and Public Wrongs. in Duff, Antony, et al, eds. *Criminalization: the political morality of the criminal law*. Oxford: Oxford University Press 291-306

Motoc, A-I., 2004 Preliminary working paper on the principle of free, prior and informed consent of indigenous peoples in relation to development affecting their lands and natural resources that would serve as a framework for the drafting of a legal commentary by the Working Group on this concept", E/CN.4/Sub.2/AC.4/2004/4 UN Doc E/C.19/2005/3

Needless, S. A. 2009 The Data Game: Learning to Love the State-Based Approach to Data Breach Notification Law, *NCL Rev.* **88** 267 – 310.

Pasquini, L., Fitzsimons J.A., Cowell, S. , Brandon S., Wescott G. 2011 The establishment of large private nature reserves by conservation NGOs: key factors for successful implementation. *Oryx* **45** 373-380.

Pettipas, K. 1994 *Severing the ties that bind: Government repression of indigenous religious ceremonies on the prairies*. Wiinipeg: Univ. of Manitoba Press.

Purtova, N. 2010 Private law solutions in European data protection: relationship to privacy, and waiver of data protection rights.? Netherlands Quarterly of Human Rights,. **28** 179–198

Raab, C. 2012 Privacy, Social Values and the Public Interest, *Politische Vierteljahresschrift* **46** 129-152

Regan, P. M., 1994 *Legislating Privacy: Technology, Social Values and Public Policy*, Chapel Hill: The University of North Carolina Press

Reichman, J.H., & Uhlir, P.F. 2003 A contractually reconstructed research commons for scientific data in a highly protectionist intellectual property environment. *Law and Contemporary Problems*, **66** 315–462.

Romanosky Sasha and Acquisti, Alessandro. 2009. Privacy Costs and Personal Data Protection: Economic and Legal Perspectives. *Berkeley Technology Law Journal* **24** 1061 - 1101.

Rubinstein, I. S. 2013 Big data: the end of privacy or a new beginning?. *International Data Privacy Law* **3** 74-87.

Ruhl, J. B. 1997 Thinking of environmental law as a complex adaptive system: how to clean up the environment by making a mess of environmental law." *Houston Law Review* **34** 933-1002

Schwartz P. M. , Janger E. J. 2007 Notification of Data Security Breaches. *Mich. L. Rev.* **105** 913-984.

Segall, L. 2015 Pastor outed on Ashley Madison commits suicide. *CNNMoney,* September 8 2015

Simitis, S., 1987 Reviewing Privacy in an Information Society, *University of Pennsylvania Law Review* **135** 707-746

Stevens, G. M. 2005 Data security breach notification laws. *CRS Report for Congress* R42475 https://www.hsdl.org/?view&did=706636 (accessed May 2016).

Stodden, V. 2009 The legal framework for reproducible scientific research: Licensing and copyright. *Computing in Science & Engineering* **11** 35-40.

Sunstein, C. R. 1999 Informational regulation and informational standing: Akins and beyond. *University of Pennsylvania Law Review* **147** 613-675.

Winn, J. K. 2009 Are "Better" Security Breach Notification Laws Possible? 2-3 *Berkley Technology Law Journal*, **24** 1133-1165.

Wiethölter, R. 1986 Materialization and proceduralization in modern law. In Teubner, G. (ed) *Dilemmas of law in the Welfare State.* Berlin: Walter de Gruyter, 221 – 249.

Zittrain, J. 2000 What the Publisher Can Teach the Patient: Intellectual Property and Privacy in an Era of Trusted Privication, *Stanford L. Rev*. ***52*** 1201- 1250