The trouble with European data protection law

Bert-Jaap Koops*

Abstract

- The trouble with European data protection law, as with Alfred Hitchcock's Harry, is that it is dead. The current legal reform will fail to revive it, since its three main objectives are based on fallacies.
- The first fallacy is the delusion that data protection law can give individuals control over their data, which it cannot. The second is the misconception that the reform simplifies the law, while in fact it makes compliance even more complex. The third is the assumption that data protection law should be comprehensive, which stretches data protection to the point of breaking and makes it meaningless law in the books.
- Unless data protection reform starts looking in other directions—going back to basics, playing other regulatory tunes on different instruments in other legal areas, and revitalising the spirit of data protection by stimulating best practices—data protection will remain dead. Or, worse perhaps, a zombie.

"The trouble with Harry is that he's dead, and everyone seems to have a different idea of what needs to be done with his body. . ."

(synopsis of Alfred Hitchcock, The Trouble with Harry, 1955[1])

Introduction: The trouble with Harry

As the technical possibilities for automated processing of huge amounts of data from multiple sources continue to grow, it is not surprising that data protection law has become a focal point of legal protection in both policy and academic literature. This is particularly visible in Europe, where the fault lines of the existing 1995 legislative framework have become ever clearer as the Internet age developed. The challenge of updating the legal framework, to enable it to stand the test of time in the coming decade or so, has been taken up by the European regulator, in the form of a proposal for a General Data Protection Regulation (GDPR) that will replace the Data Protection Directive (DPD).[2]

There is a problem, however, with European data protection law. The trouble with the law, as with Hitchcock's Harry, is that it is dead. What the statutes describe and how the courts interpret this has usually[3] only a marginal effect on data-processing practices. Data protection law is a dead letter; current ideas what to do with the body are not leading anywhere except that they offer entertainment to spectators. With the current reform, the letter of data protection law will remain stone-dead.

I am exaggerating here, of course. However, in essays such as this, hyperbole serves the function of clarifying an argument. My argument in this paper will be that the direction of the data protection reform is fundamentally flawed. It focuses too narrowly on solving too many ICT-related challenges to legal protection within a single general framework of data protection law, and by doing so diverges from the reality of 21st-century data-processing practices. In fact, all three of its new

* TILT, Tilburg University, PO Box 90153, NL-5000 LE Tilburg. Email: e.j.koops@uvt.nl.

1 Source: http://www.imdb.com/title/tt0048750/ (accessed 1 August 2014).

2 COM(2012) 11 final, 25 January 2012. An amended proposal has been adopted by the European Parliament in March 2014 (hereafter LIBE version), while the European Council is also defining its position with other amendments (I base myself here on the Addendum to the Note from the Presidency to the Council on Key Issues of Chapters I–IV of the Proposed General Data Protection Regulation (2012/0011(COD), hereafter Council version); I will refer to these where they substantially differ, but otherwise refer to the original proposal. This paper does not discuss the Directive proposal on data protection in the police and justice area, COM(2012) 10 final, 25 January 2012.

3 There are exceptions, where data protection law functions to a reasonable extent in practice, in particular in the simple forms of data processing (eg a database containing a limited number of records with personal data processed within a single organisation for a clear and uncontroversial purpose) that still occur in some contexts. In this essay, I am concerned with data-processing practices that are typical of the 21st century rather than of the 1980s, and for the sake of argument I will disregard the lingering remnants of ancient forms of data processing.

objectives, coming on top of the original objectives, put the reform on the wrong track:

> - increasing the effectiveness of the fundamental right to data protection and putting individuals in control of their data, particularly in the context of technological developments and increased globalisation;
>
> - enhancing the internal market dimension of data protection by reducing fragmentation, strengthening consistency and simplifying the regulatory environment, thus eliminating unnecessary costs and reducing the administrative burden (. . .)
>
> - to establish a comprehensive data protection framework covering all areas.[4]

I will argue that each objective is based on a fallacy: the delusion that data protection law can give individuals control over their data (fallacy 1); the misconception that the reform simplifies, while in fact it makes compliance even more complex (fallacy 2); and the assumption that data protection law should be comprehensive, stretching data protection to the point of breaking, and making it meaningless law in the books (fallacy 3). In the following sections, I will illustrate each fallacy by three key problems, to argue that the reform looks in the wrong direction. Although I do not have ready answers what to do with Harry's body, in the concluding section I will give some suggestions of better directions to look for reform.

Fallacy 1: too much focus on informational self-determination

Although data protection is not synonymous with informational self-determination, for many the two are closely related. Informational self-determination is the notion that people should be able to exercise control over what happens with their personal data; it is *their* data, after all. It implies, first, that individuals' free and informed consent is an important ground to legitimise data processing, and, second, that individuals have various rights to exercise control over the data, such as rights to correction or erasure. The perspective of informational self-determination has informed the development of data protection in important ways.[5] In data protection discussions, some advocate more radical forms of informational self-determination than others, but the notion that informational self-determination should underlie data protection law is dominant in data protection scholarship[6] and also often in policy rhetoric.[7] But in what world do people live who claim that individuals are able to exercise control over their personal data?

The first problem is the mythology of consent. The most obvious (although not the only) way in which individuals can exercise informational self-determination is by giving or withholding consent to certain forms of data processing. Particularly in private and commercial contexts, individuals' consent to data processing is usually considered the main legal ground for data processing. However, consent here is largely theoretical and has no practical meaning. It is generally recognised that with Internet-based services, most people just tick consent boxes without reading or understanding privacy statements, or that service providers sometimes assume that website visitors are somehow miraculously informed of the privacy statement and automatically give consent by merely visiting the website.[8] Surprisingly, however, the conclusion is too seldom[9] drawn that consent is simply not a suitable approach to legitimate data processing in online contexts; many scholars and policy-makers seem simply to suggest ways to make online consent more informed, more conscious, and more mandatory for providers.

The continued belief in consent as a major legitimating ground also in online contexts denies the reality of 21st-century data processing, which creates two fundamental challenges that make informed and voluntary consent a Sisyphean task. Often, there is little to choose: if you want to use a service, you have to comply with the conditions—if you do not tick the consent box, access will be denied. And there are no good alternatives: most other providers of the service you want apply the same practice and similar data-processing conditions, and with the most-used major services, such as Facebook, Google, or Twitter, there is no realistic alternative for

4 GDPR, p 102 (emphasis added).

5 A Westin, *Privacy and Freedom* (Atheneum Press: New York, 1967); BVerfG 15 December 1983, BVerfGE 65, 1.

6 See, eg J.C. Buitelaar, 'Privacy: Back to the Roots', *German Law Journal,* 13/3 (2012), 171–202, Antoinette Rouvroy and Yves Poullet, 'The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy', in Serge Gutwirth and others (eds), *Reinventing Data Protection?* (s.l.: Springer, 2009), 45–76.

7 Eg European Commission, *A Comprehensive Approach on Personal Data Protection in the European Union* (European Commission: Brussels, 2010b) at 7 (arguing that an important precondition 'for ensuring that individuals enjoy a high level of data protection' is 'the retention by data subjects of an effective control over their own data', emphasis in original).

8 Solon Barocas and Helen Nissenbaum, On Notice: The Trouble with Notice and Consent, *Proceedings of the Engaging Data Forum* (Cambridge, MA, 2009), Daniel J Solove, 'Introduction: Privacy Self-Management and the Consent Dilemma' (2013) 126 Harvard LR 1880–903.

9 There are refreshing exceptions, eg, Helen Nissenbaum, 'A Contextual Approach to Privacy Online' (2011) 140:4 *Daedalus* 32–48.

most people. Underlying this is the fact that there are practically no alternative business models that generate revenue from other sources than user-data-based profiling and advertising. Although paid and privacy-friendly services are theoretically possible, the move from free services to paid services is not something most Internet users want to make, conditioned as they are in thinking the Internet offers free lunches.

Another challenge of relying on consent is that convenience and people's limited capacity to make rational decisions prevent people from seriously spending time and intellectual effort on reading the privacy statements of every website, app, or service they use. Consent supporters focus on finding ways to make consent both more practical, eg using short texts and icons, and more meaningful, eg using plain language and technically enforcing that people actually see the text before ticking a box. However, they ignore the trade-off between practical consent and meaningful consent: the simpler you make the consent procedure, the less will users understand what they actually consent to; and the more meaningful you make the consent procedure (providing sufficient information about what will happen with the data), the less convenient the consent will become.[10] Steering between the Scylla of meaningless consent and the Charybdis of inconvenience will typically imply avoiding inconvenient Charybdis (which would swallow the whole boat) and verging alongside meaningless Scylla (losing six of the crew), resulting in—at best—some half-baked form of informed consent. There simply is no way in which ticking a consent box can 'ensur[e] that individuals are aware that they give their consent to the processing of personal data'[11] in any meaningful understanding of 'awareness' of data-processing practices and conditions.

This should lead us to conclude that data processing in most online contexts should be based on grounds other than consent. Sadly, one of the few sensible elements of the proposed GDPR to move away from consent-based data processing—Article 7(4): 'Consent shall not provide a legal basis for the processing, where there is a significant imbalance between the position of the data subject and the controller'—has been struck from both the LIBE and the Council-amended versions. Thus, the consent myth continues to hold sway over data protection law.

The second problem is that, whether or not data processing is based on consent, exercising control over personal data is extremely difficult, if not impossible, for individuals to realise in 21st-century data-processing practices. Informational self-determination is unenforceable. There are still simple situations in which a single data controller processes a relatively small set of personal data for a relatively clear purpose, but as the database, profiling, and Big Data era evolves, these situations will be few and far between. Most practices involve multiple data controllers and processors sharing sets of data, for multiple, not seldom fuzzy, purposes, and increasingly with automated operations on data—think of cloud computing and profiling—that data controllers themselves do not fully understand or know the details of. How can one effectively exercise the right to being informed, the right to access, to right to correction, and the right to erasure in such practices?

As is the case with consent, the exercise of data subject rights is highly theoretical. Yes, you can be informed, if you know where to look and how to read (but who knows, looks, and reads?). Yes, you can request controllers to let you know what data they process, if you know that you have such a right in the first place (but which controller really understands and seriously complies with all such requests, particularly if exercised on an above-incidental scale?). Yes, you can request correction or erasure, if you know whom to ask (but how are you ever going to reach everyone in the chain, or mosaic, or swamp, of interconnected data processing?). There are simply too many ifs and buts to make data subject rights meaningful in practice. The case against Facebook brought by Austrian law student Max Schrems is an eminent example of how useful the exercise of data subject rights can be to influence data controllers,[12] but the example tells us nothing about data subjects having effective control over how their personal data are being processed. Your average Internet user simply is not Max Schrems, and your average data controller does not neatly send you a 1,222-page cd with all the data they process about you (as Facebook sent to Mr. Schrems). Not even the firmest believers in informational self-determination can claim, at least not with dry eyes, that they actually know which of their data are being processed in what ways by data controllers, or that they have effective control on most data-processing operations they are subjected to.

The *Google Spain* case provides a glimmer of hope for effectively exercising control over data processing. In that case, the Court of Justice of the European Union found that search engines are data controllers over

10 Ibid, at 36.

11 GDPR, recital 25.

12 'Facebook could face €100,000 fine for holding data that users have deleted', *The Guardian* 20 October 2011, <http://www.theguardian.com/technology/2011/oct/20/facebook-fine-holding-data-deleted> (accessed 1 August 2014). See further <http://europe-v-facebook.org/> and <http://de.wikipedia.org/wiki/Maximilian_Schrems (accessed 1 August 2014).

search results containing personal data, and that they have to remove certain links from search results if the data subject considers that some information should, after a certain time has elapsed, no longer be linked to his name.[13] The fact that within two months after the Google Spain ruling, Google received 91,000 removal requests involving 328,000 URLs[14] shows that data subjects are actively exercising their right to request erasure. It remains to be seen, however, whether and how Google will process such large numbers of requests, which all require a case-by-case analysis, a process that can hardly be automated. Moreover, the Google Spain decision is nuanced and leaves considerable room for balancing competing interests and for interpretation, such as when a data subject is considered a public figure (in which case the interest in retaining the link will usually outweigh the person's interest in removal). Also, the situation is limited to search engines—leaving the information with the hosting providers intact—and in particular seems limited to searches for the person's name; searches using other terms (possibly including typos or spelling errors in names) will still link to the information. And going to a court to enforce your right to erasure can lead to increased publicity (the whole data protection community and half the reading public now know that Mario Costeja González was once a defaulter); a judgment paradoxically triggers new news items referring to the offending information, and erasure of these new pieces *cannot* be invoked since they constitute correct and relevant recent news. This demonstrates the complexities of erasure requests with online services, and the Sisyphean task of effectively exercising control over Internet information.

The third problem is that, although informational self-determination can—theoretically—function effectively in private relationships, it functions poorly, and in many cases is not supposed to function, in citizen–government relations. Citizens exercising control over what happens with their personal data, which is what informational self-determination involves, is at odds with the character of the public sector. First, consent cannot be an important ground for legitimising data processing by the government. Data processing in the public sector usually relies on legal obligations or a public interest; it would be difficult to maintain government records if they were compiled based on consent. And if consent is used as a basis for data processing, the idea of free and informed consent is an even larger myth than it is in the private sector. Consent implies a choice between realistic options; citizens, however, cannot choose another government or different government services with friendlier privacy policies. A case in point is the current decentralisation of important parts of social care in the Netherlands. Municipalities need to draft data protection policies for sharing personal data among many local authorities involved in social care, and many municipalities intend to base this on consent.[15] Thus, citizens in need of help from the government face a choice between consenting to extensive data sharing (without knowing or being able to influence the details of the protocol) in order to get social care, or not consenting and consequently not getting help from the government. Consent would only be meaningful if citizens could opt for a different form of social care with different forms of data processing—but that is not on offer.

Second, data subject rights apply, in theory, to most governmental forms of data processing (although not, or not to the same extent, in the law enforcement or national security context). However, this is limited to some basic standards of fair processing, such as having accurate and up-to-date data. Erasure might be requested, but depends on the government's determining whether it still needs the data. The rights do not involve any form of control over *how* the data are processed *for which purposes*. And that is the point: the government as data controller determines when, how, and why it processes data—citizens have nothing to choose here. Calling the exercise of rights to be informed, access, correction and, perhaps, erasure in citizen–government relations 'having control over one's information' is a travesty of the word 'self-determination'. Citizens do not determine which data the government can process in which ways, and there is no informational self-determination in the public sector.

And since there is also little informational self-determination in the private sector, as I argued above, trying to ground data protection law on a notion of informational self-determination, with a consequent focus on user empowerment, is a fundamental fallacy.

Fallacy 2: too much faith in controller actions

Data protection law not only relies on user empowerment, but it also relies on controllers' fulfilling their obligations, partly under the shadow of Data Protection Authorities' supervision but partly out of their own

13 CJEU 13 May 2014, Case C 131/12 (*Google Spain v AEPD and Mario Costeja González*).

14 Google, letter to Article 29 Working Party, 31 July 2014.

15 According to a survey among 50 municipalities, see 'Wie kunnen er straks allemaal je dossier inziens?', *NRC Handelsblad* 16 August 2014.

accord. After all, barring some exceptions, most organisations value legal compliance. Assuming that data controllers want to follow data protection law, can we realistically assume that they are in a good position to do so?

The first problem with controller compliance is that data protection law is complex. No data protection expert will deny that (indeed, data protection lawyers can be suspected of having an interest in complexity as it provides them with work). The 2007–2008 evaluation of the Dutch Data Protection Act is telling in this respect.[16] The first part, based on desk research, identified as a major deficiency the lack of clarity and vagueness of the statutory concepts and the open-ended terminology, especially in key terms and definitions; deficiencies also arose from the general, comprehensive character of the Act.[17] The second part, based on empirical research—significantly entitled *Where ignorance is bliss, 'tis folly to be wise*—concluded that the law's goals are 'not yet fully realised' in practice, and that the development of norms and guidelines through sectoral standards and case-law (which requires specific knowledge) had not yet materialised widely in practice.[18] In other words, the open norms were difficult for stakeholders to apply in the real world, and considerable work remained to be done to translate the open norms into workable, sector-specific, and context-specific rules and practices. This is significant, as it repeated the evaluation of the precursor legislation of the 1980s–1990s, which also found the general data protection legislation to be very complex, indeterminate, and little known at the practical level.[19] A similar conclusion, albeit formulated in a more politically correct manner, was drawn in a review of the European Data Protection Directive:

> it was also widely recognised that more value can still be extracted from current arrangements. A lot can be achieved by better implementation of the current rules, for instance by establishing consensus over the interpretation of several key concepts and a possible shift in emphasis in the interpretation of others.[20]

There was one advantage the European legislation had over the Dutch implementation. Whereas the Dutch law had 83 articles,[21] resulting in a law for specialists (and thus not for the average organisation that processes personal data)[22] with a complex set of rules that 'requires more from people than is humanly possible'[23], the Data Protection Directive only counts 34 articles and thus provides a relatively manageable whole. The proposed GDPR, however, tries to reinforce the thrust of data protection law by including 91 articles, almost tripling the size from 12,500 to 35,000 words. Although not all provisions are targeted at controllers, the law is a tightly-woven construction with many interdependencies, requiring controllers to grasp and implement a significantly more elaborate system of rules. Understanding the law is not made easier by a maze of interlocking constructions, such as the right to erasure of Article 17(1)(c), which applies when data subjects object on the basis of Article 19, which refers to the grounds of Article 6(1)(d-f), except when Article 80, 81, 83 or 17(4) apply, the latter being the case when for example Article 18(2) applies; and to make compliance easier, the Commission can adopt further rules to specify the criteria and requirements pertaining to all this for specific sectors and situations. If the Data Protection Directive was difficult to work with for controllers seeking legal compliance—and many well-meaning organisations have little knowledge of, and serious difficulty in understanding, the DPD[24]—the GDPR will certainly not make it easier for controllers to implement data protection law.

The second problem is that the *ex ante* focus of regulation—aiming at preventing unnecessary data processing—is reinforced by new obligations on controllers to conduct Data Protection Impact Assessments (DPIAs) (Article 33) and to implement 'data protection by design and by default' (Article 23). Obviously, prevention is better than cure, but current data protection law, with its many *ex ante* requirements, can hardly be said to have demonstrably led to preventing unnecessary data processing in practice (see Fallacy 3 below). In reality, controllers do not intend to restrict data processing to the bare minimum. Moreover, because of the many open and fuzzy norms, they can easily argue that what they do

16 For the context, see BJ Koops, 'The Evolution of Privacy Law and Policy in the Netherlands', (2011) 13:2 *Journal of Comparative Policy Analysis,* 165–79.

17 Gerrit-Jan Zwenne and others, *Eerste fase evaluatie Wet bescherming persoonsgegevens. Literatuuronderzoek en knelpuntanalyse* (eLaw@Leiden: Leiden, 2007).

18 HB Winter and others, *Wat niet weet wat niet deert. een evaluatieonderzoek naar de werking van de Wet bescherming persoonsgegevens in de praktijk* (WODC: Den Haag, 2008), 11.

19 JEJ Prins and others, *In het licht van de Wet persoonsregistraties: zon, maan of ster?* (Samsom: Alphen aan den Rijn, 1995).

20 Neil Robinson and others, *Review of the European Data Protection Directive* (RAND, Santa Monica, CA 2009), viii.

21 Wet bescherming persoonsgegevens [Data Protection Act], *Staatsblad* 2000, 302.

22 Zwenne and others, *Eerste fase evaluatie Wbp*, 74.

23 E Schreuders and H Gardeniers, 'Materiële normen: de kloof tussen de juridische normen en de praktijk', (2005) *Privacy & Informatie* 260–62, quoted in Zwenne and others, *Eerste Fase Evaluatie Wbp*, 73.

24 See also Lee A Bygrave, 'Data Privacy Law and the Internet: Policy Challenges', in Normann Witzlieb and others (eds), *Emerging Challenges in Privacy Law. Comparative Perspectives* (Cambridge UP: Cambridge, 2014), 259–89, 274, 77, Christopher Kuner, 'The "Internal Morality" of European Data Protection Law' (SSRN, 2008), 18.

is 'necessary' for the purposes they define themselves with usually less than razor-sharp precision, until, in rare cases, some supervisory authority stops them.

To be sure, DPIAs are useful instruments, and they will lead some controllers to actually think beforehand about which data they want to collect and what to do with them, instead of applying the prevalent catch-as-catch-can approach to data collection and processing. But whether a legal obligation to conduct a DPIA (in cases where processing presents 'specific risks') will cause an average controller to seriously think about data processing before setting up new systems or procedures, and actually adapt the plans to minimise the risks, remains to be seen. I fear that, as long as data protection is not in the hearts and minds of data controllers—and the law so far has done a poor job in reaching those hearts and minds (see also Fallacy 3)—mandatory data protection impact assessments will function as paper checklists that controllers duly fill in, tick off, and file away to duly show to auditors or supervisory authorities if they ever ask for it. Procedure followed, problem solved. But truly following an impact assessment, and particularly translating its findings into actual data protection-friendly designs and default settings as envisioned by Article 23, requires an attitude that looks beyond legal compliance, a vision that sees the logic behind the many legal rules of data protection, a mindset that is aware of the rationale of data protection. Articles 33 and 23 will not by themselves foster such an attitude, vision, or mindset, and we have yet to come up with good answers as to how data protection law can otherwise reach the hearts and minds of controllers. Including more sticks than carrots in data protection law will certainly not help to find a soft spot in most controllers' hearts for data protection.[25]

The third problem mirrors the second one: the GDPR also intensifies *ex post* regulation, with an increased focus on accountability and oversight, but as with *ex ante* regulation, provisions aiming at increased accountability run a risk of leading to more paper rather than more data protection. Current criticisms of European data protection law 'have often focused on the formalities imposed by the Directive (or by the transpositions thereof)'.[26] While an important objective of the data protection reform is 'simplifying the regulatory environment, thus eliminating unnecessary costs and reducing the administrative burden',[27] the reduction of some administrative burdens (such as notification) is amply compensated by the creation of new ones. Article 22 requires the controller to adopt policies and to 'be able to demonstrate that the processing of personal data is performed in compliance with this Regulation'. Controllers are required to keep documentation of all their data processing operations (Article 28). The Parliament and Council versions of these provisions diverge somewhat in the type of required red tape, but whatever compromise is reached, controllers will have to document what they do with personal data. Although the documentation obligation of Article 28 may apply only to organisations employing 250 persons or more (but the LIBE versions stipulates it for all controllers), it may also apply to small organisations, depending on how one interprets the rather vague exception;[28] and in any case, controllers need to be able to demonstrate compliance as per Article 22, which necessitates some form of documenting what they do. Will such documentation assist in increased compliance and better data protection?

That will depend not only on whether enforcement by supervisory authorities will be effective—a considerable challenge given a wide-spread scarcity of resources for DPAs to provide effective oversight over a myriad of data controllers—but also on whether the act of documentation will make controllers think about what they do, and adapt their practice accordingly if they realise, when documenting, that their activities are actually not compliant with the regulation. As with the *ex ante* instruments, this will only take place if controllers have a data protection *rationale* mindset, instead of a data protection *rule compliance* mindset, and such a mindset, as I perceive it, is all too frequently absent even on the part of well-meaning controllers. The result will be more paper (or disk space) and more work for data protection practitioners, but not, I fear, more protection of personal data. On the contrary, filling in forms about compliance with rules runs the risk that rules are blindly followed in their letters, but that their spirit is overlooked: the spirit of data protection can hardly be captured in documentation.

Hence, the complexity of data protection law, which is increasing instead of decreasing, together with more *ex ante* and *ex post* paperwork and checklist obligations, creates a situation in which controllers will, at best, blindly follow a set of rules to be rule-compliant, while (still) not understanding much about data *protection*. The fallacy of creating more rules to ensure controller compliance will not lead to data protection law being a

25 Bygrave, 'Data Privacy Law and the Internet', 288–89.

26 Robinson and others, *Review of the European Data Protection Directive*, viii.

27 GDPR, p 102.

28 Article 28 would also apply to small or medium organisations that process personal data as an activity that is more than 'ancillary to its main activities' (original version), or where data processing 'involves specific risks' (Council version).

dead letter, as the letters are (at best) obediently followed, but it will result in the law being a zombie: it seems to live, but lacks a vital spirit.

Fallacy 3: regulating everything in one statutory law

The roles allocated by data protection law to data subjects and data controllers, based on the fallacies of expecting too much from either, are embedded in a script, namely the Data Protection Directive and the proposed General Data Protection Regulation. Focusing fully on defining rights and duties in statutory law, particularly in a single comprehensive framework, is a third fallacy. Law in the books does not always become, nor does it always resemble, law in action. In data protection law, the gap between law in the books and law in action is particularly glaring, and this will be compounded by expanding data protection law in an attempt to address new challenges, such as profiling. The on-going focus on command-and-control regulation to the neglect of other regulatory tools does not help either to achieve better data protection in practice.

The first problem is an enormous disconnect between the law and reality. The basic, and for many data protection experts major, notion underlying data protection law is, and continues to be, data minimisation, which is a combination of the traditional principles of collection limitation, data quality (requiring data to be relevant), purpose specification, and use limitation.[29] These principles are maintained in the GDPR. Recital 30 reads:

> The data should be adequate, relevant and limited to the minimum necessary for the purposes for which the data are processed; this requires in particular ensuring that the data collected are not excessive and that the period for which the data are stored is limited to a strict minimum. Personal data should only be processed if the purpose of the processing could not be fulfilled by other means.

And these are not only legal requirements (Article 5(b),(c), (e) GDPR); the principles should also be translated into technical or organisational measures, in Article 23(2)'s vision of data protection by design and default:

> The controller shall implement mechanisms for ensuring that, by default, only those personal data are processed which are necessary for each specific purpose of the processing and are especially not collected or retained beyond the minimum necessary for those purposes, both in terms of the amount of the data and the time of their storage.

Think for a moment of all the databases in which your personal data are stored, including the databases you do not really know but can reasonably assume to exist (almost all online providers you have interacted with, scores of governmental databases, credit-reporting agencies, etc.), and then reread recital 30 and Article 23(2). Type in 'big data' in a search engine and browse what is happening in the field of massive data collections—'Data, data everywhere' is an instructive *Economist* report[30]—and then reread recital 30 and Article 23(2).

In 2010, the total amount of information processed globally was estimated to be 1.2 zettabytes (that is 1,200,000,000,000,000,000,000 bytes) and growing at a rate of 60% per year;[31] of course, much of that is not personal data, but with increasing technological capacities to combine and interpret data, personal data will show up ever more frequently in the zettabytes of 21st-century information flows. The Data Protection Directive has done little to prevent the development of massive databases or the advent of the Big Data era, and it is folly to think that the GDPR will fare better in preventing 'unnecessary' data processing.[32] Who in his right mind can look at the world out there and claim that a principle of data minimisation exists?

The second problem is that the response to the regulatory disconnection between data protection law and data-processing practice is more law, and specifically more-of-the-same law. Not only does the law triple in size and create more obligations (see Fallacy 2), it also expands its scope. One aspect of expansion is the definition of personal data—already a much-debated issue under the DPD[33]—that remains largely the same (data relating to 'an identified natural person or a natural person who can be identified, directly or indirectly, by means reasonably likely to be used by the controller or by any other natural or legal person'), but now also refers to identifying individuals on the basis of 'an identification number, location data, online identifier'[34] (Article 4(1) GDPR). While data protection law has traditionally focused on 'lookup identifiers'

29 See OECD, *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (OECD: s.l., 1980); OECD, *Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data (2013)* (OECD: s.l., 2013); Article 6(1)(b)-(c) Data Protection Directive.

30 The Economist, 'Data, Data Everywhere', (27 February 2010) *The Economist* 1–13.

31 Ibid.

32 Cf. Bygrave, 'Data Privacy Law and the Internet', 272–74.

33 See, eg Article 29 Working Party, *Opinion 4/2007 on the Concept of Personal Data* (Article 29 Data Protection Working Party: Brussels, 2007), 26; Paul M Schwartz and Daniel J Solove, 'Reconciling Personal Information in the United States and European Union', (2014) 102 California LR 877–916.

34 The LIBE version uses the term 'unique identifier' instead of 'online identifier'.

(L-identifiers), ie an identifier associated with a register, directory, or table in which the connection between the identifier and a named individual can be looked up,[35] it now also looks at 'recognition identifiers' (R-identifiers), ie data that allow an individual to be recognised as a previously known individual, without (necessarily) being able to associate the identifier with a *named* individual.[36] Such R-identifiers can be combined with other data to identify the individual by name (or otherwise uniquely pinpoint the individual), in which case the R-identifier also counts as personal data. This will trigger much debate when online identifiers, such as cookies, should be considered personal data: companies will argue they should not be (as they will not use them for name-identification purposes), while data protection activists will argue they can be easily combined to link the identifier with a named individual—essentially repeating a debate that has been going on for some time.[37] Although using R-identifiers to take decisions about (not necessarily identified or identifiable) persons raises privacy and non-discrimination concerns that are somewhat similar to decision-making based on personal data,[38] not all identifiers function in the same way, and it makes sense to differentiate in the legal regimes for different types of identifiers.[39] Instead of allowing for differentiation, however, EU data protection law applies an all-or-nothing approach: data is either personal data (triggering the whole regime), or it is not (triggering nothing), but it cannot be something in between or something else.[40]

A similar issue occurs with profiles, which the GDPR also aims at regulating through Article 20, which says (in simplified terms) that persons have the right not to be subjected to decisions based solely on profiling. This seems to suggest that data protection law will be extended to regulate profiling at large—an important development given the normative challenges of profiling.[41] The key feature of profiles is that they do not necessarily relate to individuals, but often to groups ('someone with characteristics x, y, and z'), which makes them non-personal data (until they are applied to identified individuals) and hence the creation and much of the processing of profiles traditionally falls outside of data protection law.[42]

But while the impression is created that the GDPR will regulate profiling, this is actually not the case. Confusingly, Article 20(1) talks about 'natural persons'[43] not being subjected to purely profiling-based decisions, suggesting that this applies not only to identifiable persons but also to unidentifiable persons. However, the exception in Article 20(2) talks about 'data subjects' and Article 20(4) about controllers (ie those who process *personal* data); this calls into question whether Article 20(1) indeed covers situations in which group profiles are applied to unidentifiable individuals. While the formulation itself carries the impression it does, the rest of Article 20 contradicts that impression. More importantly, the material scope of the GDPR is defined as 'the processing of personal data' (Article 2), and the GDPR's subject matter is 'rules relating to (. . .) the processing of

35 Ronald E Leenes, 'Do They Know Me? Deconstructing Identifiability' (2008) 4:1–2 University of Ottawa Law Technol J 135–61, 148.

36 Ibid, 149–50.

37 Cf. Article 29 Working Party, *Opinion 4/2007 on the Concept of Personal Data*, 14.; Bygrave, 'Data Privacy Law and the Internet', 265–67.

38 Arnold Roosendaal, *Digital Personae and Profiles in Law. Protecting Individuals' Rights in Online Contexts* (Wolf Legal Publishers: Oisterwijk, 2013).

39 Leenes, 'Do They Know Me? Deconstructing Identifiability'; Schwartz and Solove, 'Reconciling Personal Information in the United States and European Union', 15–16.

40 The amended versions do introduce a new category, namely of pseudonymous data (Article 4(2a) GDPR), which could become a useful in-between category. However, the currently used definition is problematic ('personal data that cannot be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organizational measures to ensure non-attribution'), as it constitutes a contradiction in terms: it refers to personal data that (because non-attribution is ensured) cannot be linked to an individual, and hence are not personal data.
Also in a teleological reading, the definition is problematic. The formulation seems to refer to data relating to an individual who could, in theory, be identified but where identification is made difficult through technical–organisational measures. Now, there are two possibilities. If identification is made really difficult, the data do not fall under the definition of personal data (since there are no means that can reasonably be expected to be used to make the link); this constitutes a meaningful sub-category of non-personal data, but it is self-contradictory to regulate such data in a legal instrument that restricts its material scope to personal data (cf. my discussion of profiles below). If, on the other hand, identification is not made really difficult, in the sense that the technical–organisational measures are in place but the connection can still, with reasonable means, be made by the controller, the pseudonymous data are personal data. This constitutes a meaningful sub-category of personal which could be attributed, in some respects, a lighter regime. However, since personal data have to be secured with appropriate technical–organisational measures anyway (Article 30 GDPR), it is unclear why these personal data would have a privileged position, with a lighter regime, over personal data that are secured in other (and possibly equally or more effective) ways, such as in stand-alone computers or with state-of-the-art protection against leaking. Hence, although creating a new category of data seems a promising way of allowing differentiation, introducing differentiated regimes for certain types of personal data or certain types of processing requires more reflection and careful crafting than the current proposal, with its non-technology-neutral preference for just one type of data security, achieves.

41 Mireille Hildebrandt, 'Profiling and the Identity of the European Citizen', in M. Hildebrandt and S. Gutwirth (eds), *Profiling the European Citizen* (Springer: s.l., 2008), 303–26; BHM Custers and others (eds), *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases* (Springer: New York, 2013); M Hildebrandt and Katja De Vries (eds), *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology* (Routledge: Abingdon, 2013); Roosendaal, *Digital Personae and Profiles in Law.*

42 Hildebrandt, 'Profiling and the Identity of the European Citizen'; Roosendaal, *Digital Personae and Profiles in Law.*

43 The Council version more consistently uses the term 'data subject'.

personal data and rules relating to the free movement of personal data' (Article 1). Given its self-declared scope, the GDPR cannot cover the creation and application of group profiles in general, but only the creation and use of individual profiles. This will create similar demarcation problems (when is a profile related to an identifiable individual?) as online identifiers.

The underlying problem is that all data processing is seen through the lens of personal data, in the frame of data protection law. As socio-technological developments raise new regulatory challenges—behavioural advertising, profiling—the regulatory response is to include these in the data protection framework. This, however, requires stretching the concept of personal data (sometimes to the point of breaking, or perhaps rather of becoming void of meaning), or stretching the regulatory problem so that it becomes a problem of processing personal data. The result is a framework that theoretically separates the world of data processing in two parts, namely the processing of personal data and the processing of non-personal data, which practically leads to large and increasing border conflicts.

Another primary example of framing Internet-related problems as data protection problems is data portability, a regulatory challenge that is now appropriated by data protection law (Article 18 GDPR (Article 15(2a) LIBE version)). By its nature, data portability would be more at home in the regulation of unfair business practices or electronic commerce, or perhaps competition law—all domains that regulate abuse of power by commercial providers to lock-in consumers. Framing such power abuse as a data protection problem leads to introducing new types of protection into an already complex system, leading controllers to lose sight of the forest of data protection's rationale for the trees of rules, and requiring supervisory authorities to expand their staffing and scope with expertise that already exists with competition and consumer supervisory authorities.

In short, the second problem of relying on law in the books is the expansion of data protection legislation to include almost all types of data processing, leading to an artificial framing of data-processing problems in terms of personal data. This compounds the regulatory disconnection by moving data protection law further away from how people working with data processing perceive the world.

This leads to the third issue: data protection law has a communication problem. It does not have a particularly positive connotation with controllers: instead of perceiving data protection as a basic and reasonable set of rules for decent treatment of people, controllers often, in my experience, perceive data protection as an obstacle and a nuisance. Many controllers see data protection law as a traffic light that is usually red or else a warning yellow (or orange, depending on the country), but seldom green. The misconception that data protection law only *restricts*, and not also *enables*, is wide-spread, and of the two functions of data protection—protecting fundamental freedoms and stimulating the free flow of personal data—the latter is often overlooked.

Apparently, regulators have not managed to speak to the regulatees in the right register, to make them realise not only the limitations but also the opportunities that data protection law provides for data processing practices. The miscommunication persists, I think, partly because many data protection practitioners (corporate lawyers and possibly also data protection officers) likewise tend to see data protection as a yellow/red traffic light rather than a traffic light that is generally meant to allow you through (green), although you may have to wait while you take appropriate measures (yellow), and only for some situations provides a no-go (red). The yellow/red framing of data protection also persists because it is the main frame in the media and in political debates. A reframing of the dual nature of data protection, as both affording and restricting—in other words canalising rather than prohibiting—is urgently needed, but the current reform does little in that area, creating an overall impression of sticks (increased obligations, new obligations, and stepped-up enforcement and fines) rather than carrots (the main carrot being that if you comply with certain obligations you more easily avoid the stick).

Another aspect of the communication problem is poor expectation management. Despite the huge gap between law in the books and law in action, regulators keep using language suggesting that data subjects will gain control over personal data through data protection law.[44] Few data subjects actually have a feeling of data control, and the difficulty of enforcing their rights in practice does not help to get a better feeling of control. Also telling is the error of judgment in using the label 'right to be forgotten' (Article 17 GDPR, original version) for what is nothing else than a slightly expanded form of the existing right to erasure. The term 'right to be forgotten' has been floated as an appealing ideal for doing something about the persistence of embarrassing data on the Internet, but it is pretty obviously a misnomer: not only is it very difficult to have all copies of data removed from the Internet, but removing content

44 *Supra*, note 6.

from the Internet also cannot be equated to people actually forgetting what they have already read.[45] The error is being corrected in the amended GDPR (the LIBE version now calls Article 17 only the 'right to erasure'), but the damage has been done: the label sticks, and the expanded (but by no means absolutely effective) right to erasure will be discussed for a long time in the frame of 'being forgotten', raising expectations that the right could not deliver in the first place.

Underlying the communication problem is a too narrow focus on law, and particularly command-and-control law, leaving the other tools of the regulatory toolbox[46] largely lying idle. The regulators now also apply 'code' as a regulatory tool, through the notion of data protection by design and by default, but still in the form of command-based law (Article 23 GDPR provides a legal obligation to employ 'code'-based protection). Although relying on 'code' seems an attractive proposal, there are significant challenges to make it work, and in some important respects data protection law cannot be hardcoded,[47] making it largely a useful tool for assuring compliance with the easy, routine parts of data protection law.[48]

Self-regulation, as a form of consensus-building, is stimulated through the encouragement of codes of conduct (Article 27 DPD, Article 38 GDPR), but within a strong legislative framework this is co-regulation with relatively little space for regulatees to develop their own rules; moreover, the added value and efficacy of codes of conduct are contested.[49] Communication as a regulatory tool is used to some extent, with useful guidelines and self-assessment tools being published by Data Protection Authorities, but regulators have spent relatively little effort to communicate best practices (perhaps because best practices are altogether too scarce?) and, as observed above, they have a communication problem. Competition as a regulatory tool is also not yet being widely employed; this may be due to market structures of the data economy, but given the dominance of certain multinational Internet companies, a focus on providing market incentives for alternative providers with more privacy-friendly policies and default settings might be more helpful than command-based rules for data processing.

Altogether, the regulatory disconnection of data protection law, which risks being enlarged rather than diminished through stretching the scope of data protection law to embrace new regulatory issues, together with a narrow focus on command-and-control law, demonstrate a fallacy of regulators to believe that every problem related to Internet data flows can be regulated by data protection law in the books. This does not work, as any realist looking at 21st-century data processing practices will acknowledge.

Conclusion: what to do with Harry?

Outsiders might enjoy the data protection reform as 'a comedy about a corpse',[50] but for insiders—European data subjects—it feels more like a zombie horror movie. We see data protection bodies moving all around, but they do not provide us with real protection. The fundamental fallacies featuring in data protection law lead to the conclusion that, as it stands, data protection law is dead. So, the question, as with Harry in Hitchcock's film, is what needs to be done with its body.

A first option is to resurrect it. I am not sure whether this can be achieved, but if it is to be done, it requires a different approach than the current reform. Instead of making data protection law broader and more detailed in how it is to be implemented and enforced, which makes it more complex and more rigid and therewith unrealistic for 21st-century data processing, data protection law should be simplified and focus more on the main underlying principles. In other words, it should go back to its roots, the basic data protection principles such as those stipulated in the OECD data protection guidelines.[51] These provide a general framework in which the spirit of data protection is clearly visible, in contrast to the EU law's tree-obscuring forest of rules. The principles could come alive on the work-floor of

45 Cf. Paulan Korenhof, 'Forgetting Bits and Pieces. An Exploration of the "Right to Be Forgotten" as Implementation of "Forgetting" in Online Memory Processes', in M. Hansen and others (eds), *Privacy and Identity Management for Emerging Services and Technologies* (Springer, 2014), 114–27.

46 Bronwen Morgan and Karen Yeung, *An Introduction to Law and Regulation: Text and Materials* (Cambridge UP: Cambridge, UK/New York, 2007) identify command, consensus, communication, competition, and code (i.e., technology or architecture) as regulatory tools.

47 BJ Koops and Ronald Leenes, 'Privacy Regulation Cannot Be Hardcoded. A Critical Comment on the "Privacy by Design" Provision in Data-Protection Law', (2014) 28:2 Int RL Comput 159–71.

48 BJ Koops, 'The (in)Flexibility of Techno-Regulation and the Case of Purpose-Binding' (2011b) Legisprudence 171–94.

49 Koops, 'The Evolution of Privacy Law and Policy in the Netherlands', 169–70 (arguing that the promise of self-regulation has not been fulfilled in Dutch practice); Dennis D Hirsch, 'Going Dutch? Collaborative Dutch Privacy Regulation and the Lessons It Holds for U.S. Privacy Law', (2103) 2013:1 Michigan State LR 83–166 (arguing, at 161, that Dutch codes of conduct 'have worked in some important respects' and, at 151, that they have 'important virtues' but also 'significant weaknesses'); Yves Poullet, 'The Directive 95/46/EC: Ten Years After', (2006) 22 *The Computer Law and Security Report* 206–17, 210–11.

50 One of the taglines used for Hitchcock's The Trouble with Harry, IMDb, <http://www.imdb.com/title/tt0048750/taglines?ref_=tt_stry_tg> (accessed 1 August 2014).

51 *Supra* note 28.

each data controller, if data protection 'is integrated into its governance structure and [the controller] establishes internal oversight mechanisms'.[52] This will be more easily (but still not easily) achievable if the basic principles are communicated as useful corporate governance or good governance guidelines that assist organisations to respect the people they are working with. Enforcement by supervisory authorities will be needed, but should be fostered by investing in Data Protection Authorities' resources and expertise, at least as much as by creating strong enforcement powers on paper.

One thing that going back to the roots of basic principles would not achieve is harmonisation. This is the price to be paid for relying on framework legislation, and it might seem a high price to pay given the current emphasis that both regulators and companies put on creating a Regulation that unifies the law in all member states. But if we can choose between a harmonised law that is dead and a fragmented law that is living, I would rather choose the latter. Whether a framework approach that only stipulates the basic principles will be enough to resurrect data protection, is hard to say; it depends on many accompanying regulatory measures—focusing on consensus, communication, and competition—that stimulate organisational awareness, compliance, and self-monitoring. For those who want to restore data protection to life, it might be worth a try.

For those who, like me, believe that the first option is not going to work in the foreseeable future, a second option is to let data protection law rest in peace, and to regulate the protection of individuals against the risks of data processing elsewhere. Part of the solution might lie in creating *sui generis* regimes for types of data, such as online identifiers, or types of problems that fall somewhere between data protection and consumer (or other weak-party) protection, such as profiling. Current data protection law does not allow for an uncertainty principle, but it might be considerably more productive if, instead of trying fitfully to establish where the border lies between personal and non-personal data, we would allow for categories of data that have certain effects on people when they are processed, regardless of whether or not they relate to identifiable individuals. Just as light sometimes acts as a particle and sometimes as a wave, data sometimes act as personal data and at other times as non-personal data, and we simply cannot always predict which of the two occurs. For certain types of data processing, such as the use of tracking cookies, profiling, and Big Data Analytics, it does not matter that much whether data are particles or waves, if they can be treated jointly as light that is canalised in certain ways.

It would help considerably if regulators would not try to solve every problem associated with data processing through data protection law. Issues emerging through developments in technology-facilitated data processing can be regulated in *sui generis* legislation, such as a legal instrument regulating profiling, or a legal instrument regulating mass surveillance. Some problems also resemble problems dealt with in other fields, and data protection can also be achieved by making fair data processing part and parcel of those fields. Data portability, for example, could well be regulated in consumer protection law, while the use of online identifiers can be regulated (and to some extent is regulated, although not particularly well[53]) in electronic commerce, telecommunications, and media law. Abuses of power in governmental data processing need not necessarily be regulated in data protection law itself, but can also be protected against in public procedure law, for example, by strong limitations to collecting (blanket) personal data and rigorous scrutiny of the proportionality and subsidiarity of governmental interferences with private life. This is not easy in the current preventative, risk-minimising paradigm that pervades public policy, but the infrastructure of Article 8 ECHR oversight might, although requiring lengthy and cumbersome litigation, ultimately prove more effective to curb disproportional governmental data processing than data protection law oversight by DPAs. Profiling and protection against automated decision-making can also be embedded in consumer protection (regulation of unfair trade practices) or non-discrimination law, or, for the public sector, in administrative procedure law. Indeed, given the increasing prevalence of these practices, which cannot be realistically stopped by a preventative focus on data minimisation, data protection should be sought much more in regulating the decision-making stage than in regulating the data collection and data processing stages.[54]

A third option, which can be combined with the second or even the first, is to commemorate data protection, so that, even while its body remains dead, its spirit is kept alive. Governments could, and should, demonstrate themselves to be much more of a role model than they have done in past decades, to show that data protection is to be taken seriously. Public-sector databases have expanded as much as those in the private sector (for

52 Article 15(a)(iv) revised OECD guidelines, supra note 28.

53 RE Leenes and E Kosta, 'Taming the Cookie Monster with Dutch Law – a Tale of Regulatory Failure' (2015) 31:1 Computer Law & Security Review (forthcoming).

54 BJ Koops, 'On Decision Transparency, or How to Enhance Data Protection after the Computational Turn', in M Hildebrandt and K De Vries (eds), *Privacy, Due Process and the Computational Turn* (Routledge: Abingdon, 2013), 196–220.

example, there are no less than 18 large-scale EU initiatives to feed databases for the purposes of 'law enforcement or migration management'),[55] which makes it facetious of regulators to proclaim a principle of data minimisation. If the public sector cannot get its own information security in order, as witnessed by on-going frequent reports of data leaks from public-sector databases, how credible are mandatory data protection rules to provide 'adequate security'? If regulators advocate the use of Privacy-Enhancing Technologies since the mid-1990s, but few government agencies actually employ PETs, why should controllers believe that 'data protection by design and by default' is important? As long as governments pay lip service to data protection but in practice fail to comply with its tenets, and while regulators proposing strict data protection laws at the same time pass legislation allowing governments to create massive databases, the spirit of data protection will remain dead. Setting good examples, for example by widely employing PETs, adapting information systems to actually minimise risks to data subjects (instead of merely checking off a Data Protection Impact Assessment list), and curbing government's own data hunger, would help in making data protection law more credible. This would have positive effects on data protection not only in the public sector but also, through generating best practices, in the private sector.

To conclude, I believe that the current data protection reform is on the wrong track, since it disregards the problems underlying the current lack of actual data protection in practice. As I have argued in this essay, each of the reform's new objectives is based on a fallacy. Too much is expected from informational self-determination, which is impossible in the 21st century. Too much is expected from controllers, for whom compliance is too complex even if they want to follow the law. And too much is expected from regulating everything within a single framework of law in the books. With such fallacious objectives, data protection law is moribund if not already dead, and future regulators will have to deal with its body. Unless data protection reform starts looking in other directions—going back to basics, playing other regulatory tunes on different instruments in other legal areas, and inducing administrations and regulators to revitalise the spirit of data protection by setting good examples—data protection will remain dead. Or, worse perhaps, a zombie.

Declarations

The research for this paper was made possible by a VICI grant from NWO, the Netherlands Organisation for Scientific Research. I thank Lee Bygrave, Paulan Korenhof, Christopher Kuner, Ronald Leenes, and Nadezhda Purtova for their helpful comments on an earlier version of this paper. No conflict of interests.

doi:10.1093/idpl/ipu023
Advance Access Publication 8 October 2014

55 European Commission, *EU Information Management Systems* (European Commission: Brussels, 2010).