Oxford Journal of Legal Studies, (2014), pp. 1–30
doi:10.1093/ojls/gqu011

Information Concepts in Law: Generic Dreams and Definitional Daylight

Lee A Bygrave*

Abstract—This article concerns the legal significance of basic information concepts, such as 'data' and 'information'. It describes how the application of laws increasingly turns on the meaning of these concepts. This development partly reflects lawmakers' generic dreams—that is, a desire to keep abreast of technological change by drafting laws using relatively generic terminology rather than technology- or media-specific rules. The overriding argument is that while the regulatory importance of basic information concepts grows, legal discussion about their meaning is far from satisfactory. In particular, technological developments that are challenging the distinction between information and biological material require jurists to address the meaning of these concepts more rigorously. This involves recalibrating the regulatory principle of technology neutrality so that laws intended for the informational world do not haphazardly enter the biological world.

Keywords: data, information, biological material, technology neutrality

1. *Introduction*

A striking feature of the legal system is that while information is in many respects the stuff of law, the law often reflects an underdeveloped, if not poor, understanding of information. The chief issue tackled herein is whether this feature gives cause for concern. In the past, it seems to have been little troublesome. Law could usually be sensibly applied without lawmakers having first extensively analysed what is meant by 'information' and closely related concepts, such as 'data' and 'communication'.

That approach, however, sits uneasily with key features of present society, most notably the processing of unprecedented amounts of data and the placement of ever greater premiums on the information that the data conveys.

* Professor, Norwegian Research Centre for Computers and Law, Department of Private Law, University of Oslo. Email: lee.bygrave@jus.uio.no. Special thanks go to Bert-Jaap Koops for thorough criticism of an early draft of the article. Thanks go also to Jon Bing, Dag Wiese Schartum, Emily Weitzenboeck, Tobias Mahler, Ole-Andreas Rognstad, and Graham Greenleaf for constructive feedback along the way, and to Andrew Murray and two anonymous reviewers for helpful suggestions in finalising the article. The usual disclaimer nonetheless applies. References to legal instruments are to their amended state as of 1 March 2014.

Data and information, especially when organized as knowledge, are now commonly regarded as linchpins for innovation, productivity, and competitiveness.[1] Thus, we might reasonably expect the legal community to have developed, through systematic reflection, a stable analytical apparatus for defining 'information' and related concepts. Mostly, it has not.

A rich body of academic literature (mainly in languages other than English) has analysed the general relationship between law and information.[2] Some of this literature discusses the meaning of information concepts, but much of it has been concerned with other matters, such as how to arrive at legally relevant explications of various kinds of information and communications technology (ICT). The propriety of according property rights over information has also attracted extensive debate.[3] The debate touches on the nature of information but the treatment is superficial and, again, tends to be dwarfed by other issues, such as the vexed question of what amounts to property in law.

What are the causes of this state of affairs? One possible cause lies with those scholars who were amongst the first in the legal community to familiarize themselves with the relatively elaborate and rigorous ways in which information concepts are explicated in the field of informatics.[4] Many took the view that strictly applying an information scientist's understanding of these concepts is legally unnecessary. In the opinion of one such scholar, information is a concept that everyone basically understands, so attempts to reach a more detailed determination of its content are futile, strained, and redundant, at least

[1] Manuel Castells, *The Rise of the Network Society* (Blackwell 1996); Daniel Bell, *The Coming of Post-Industrial Society: A Venture in Social Forecasting* (Basic Books 1973).

[2] In roughly chronological order, see eg Jon Bing, *Rettslige kommunikasjonsprosesser* (Universitetsforlaget 1982); Peter Blume, *Fra tale til data* (Akademisk Forlag 1989); Egbert J Dommering, 'An Introduction to Information Law: Works of Fact at the Crossroads of Freedom and Protection' in Egbert J Dommering and P Bernt Hugenholtz (eds), *Protecting Works of Fact: Copyright, Freedom of Expression and Information Law* (Kluwer 1991) 1; JN Druey, *Information als Gegenstand des Rechts* (Schulthess Polygraphischer Verlag / Nomos 1995); Viktor Mayer-Schönberger, *Information und Recht: Vom Datenschutz bis zum Urheberrecht* (Springer 2001); John Cahir, 'Understanding Information Laws: A Sociological Approach' (2002) 3 J Information, Law and Technology (now EJLT) <www2.warwick.ac.uk/fac/soc/law/elj/jilt/2002_3/cahir> accessed 17 April 2014; Michael Kloepfer, *Informationsrecht* (Beck 2002); Mads Bryde Andersen, *IT-retten* (2nd edn, Gjellerup 2005); Henrik Udsen, *De informationsretlige grundsætninger. Studier i informationsretten* (Jurist- og Økonomforbundets Forlag 2009); William Kingston, *Beyond Intellectual Property: Matching Information Protection to Innovation* (Edward Elgar 2010); Inger Marie Sunde, *Automatisert inndragning* (Unipub / Senter for rettsinformatikk 2011) part II.

[3] Notable contributions include DF Libling, 'The Concept of Property: Property in Intangibles' (1978) 94 LQR 103; R Grant Hammond, 'Quantum Physics, Econometric Models and Property Rights to Information' (1981) 27 McGill L Rev 47; R Grant Hammond, 'Theft of Information' (1984) 100 LQR 252; Arnold S Weinrib, 'Information and Property' (1988) 38 U Toronto LJ 117; James Rule and Lawrence Hunter, 'Towards Property Rights in Personal Data' in Colin J Bennett and Rebecca Grant (eds), *Visions of Privacy: Policy Choices for the Digital Age* (University of Toronto Press 1999) 168; Paul Kohler and Norman Palmer, 'Information as Property' in Norman Palmer and Ewan McKendrick (eds), *Interests in Goods* (2nd edn, LLP 1998) 3; Paul M Schwartz, 'Property, Privacy, and Personal Data' (2004) 117 Harv L Rev 2055; Nadezhda Purtova, *Property Rights in Personal Data: A European Perspective* (Kluwer Law International 2012).

[4] The term 'informatics' is used here to encompass a range of overlapping disciplines, most notably information science, computer science, information philosophy, and knowledge management. The central remit of informatics is scientific analysis of the ways in which information is constituted, represented, processed, and communicated. See further Michael P Fourman, 'Informatics' in John Feather and Paul Sturges (eds), *International Encyclopedia of Information and Library Science* (2nd edn, Routledge 2003) 237–44. This is a simplified description of the field, but sufficient for present purposes.

outside of 'technical contexts'.[5] Still others have despaired over the definitional instability of 'information', with at least one proclaiming that the concept 'cannot be defined'.[6] Partly because of this mix of pessimism and pragmatism, few serious attempts have been made to push the rest of the legal world into more conscious reflection on the meaning of basic information concepts.

The central question is whether such reflection is now needed. This article argues that it is. The basis of the argument lies not so much in the above-described characteristics of our 'information age' but in a combination of other factors.

One such factor has scientific and technological roots. Advances in the biological sciences, biotechnology, and ICT are enabling an ever-greater generation of information from biological material. Increasingly probative forms of genetic testing highlight this trend well.[7] So too does the upsurge in the deployment of biometrics[8] and in the systematic accumulation of biological samples in biobanks.[9] Biological material is not only increasingly a source of information; it is also starting to provide the core constitutive elements of new information systems.[10] We see this in the nascent development of biocomputers made up of organic microprocessor units that are programmed by 'wetware' based on DNA (deoxyribonucleic acid).[11] Additionally, contemporary discourse in and about the biological sciences makes a great deal of metaphorical or analogical use of information concepts to describe biological processes.[12] In some cases, the processes are portrayed as genuinely informational.[13] One may question the conceptual veracity of such discourse,[14] arguing that at least some of the processes could perhaps be analysed more accurately in terms of cause

[5] Blume (n 2) 29–30.

[6] Thomas Hoeren, 'Tractatus germanico-informaticus – some fragmentary ideas on DRM and information law' in Arno R Lodder, Alfred Meijboom, and Dinant TL Oosterbaan (eds), *IT Law – The Global Future: Achievements, Plans and Ambitions* (Elsevier 2006) 149. See too Kloepfer (n 2) 24 [52] (indicating that the concept of information is well nigh impossible to define).

[7] See generally Francis S Collins, *The Language of Life: DNA and the Revolution in Personalized Medicine* (HarperCollins 2010).

[8] Biometrics denotes systems for determining or verifying the identity of persons based on their bodily characteristics. Further on their increasing popularity along with the regulatory challenges they present, see generally Nancy Liu, *Bio-Privacy: Privacy Regulations and the Challenge of Biometrics* (Routledge 2011).

[9] Biobanks are basically structured collections of organic material, which may be human or animal in origin (eg blood samples) or non-human (eg plant seeds). The material is usually linked to associated data sets consisting of the information derived from the organic material, together with other data enabling linkage between the material and its source. For more on their characteristics, growth, and regulatory implications, see generally Herbert Gottweis and Alan Petersen (eds), *Biobanks: Governance in Comparative Perspective* (Routledge 2008).

[10] At the risk of spelling out the obvious, the term 'information system' is used here to denote a system that is concerned with 'analysis, design, delivery and use of information for organizations and society': David Avison and Guy Fitzgerald, 'Information Systems' in Feather and Sturges (n 4) 306.

[11] See generally Martyn Amos, *Theoretical and Experimental DNA Computation* (Springer 2005).

[12] See eg Collins (n 7) and Matt Ridley, *Genome: The Autobiography of a Species in 23 Chapters* (Fourth Estate 1999). For older examples, see Fritz Machlup, 'Semantic Quirks in the Study of Information' in Fritz Machlup and Una Mansfield (eds), *The Study of Information: Interdisciplinary Messages* (Wiley 1983) 641, 652–55 and references cited therein.

[13] See eg Ridley ibid 7 ('The idea of the genome as a book is not . . . a metaphor. It is literally true').

[14] As does Machlup (n 12).

and effect or stimulus and response. Nonetheless, such discourse seems strongly entrenched. With all of these trends, maintaining a distinction between biological material and information becomes practically and normatively more difficult.[15]

A second factor, with a legal dimension, springs from these difficulties. As scientific and technological developments challenge the boundary between information and biological material—and, in effect, the traditional distinction between the message and the medium—they can also trigger the application of laws that employ information concepts to biological material or, in some cases, to physical material more generally. This is particularly so when the concepts are embodied in generic terminology—that is, terminology that is not exclusively tied to a particular technology, medium, or form—and when the lawmakers provide little or no guidance about the terminology's intended scope.

A third factor is that laws are increasingly formulated using such terminology, the meaning of which is frequently unclear due to a paucity of authoritative guidance. This development and underdeterminacy[16] reflect, in large part, a strategy of lawmakers for coping with technological change. The strategy seeks to inject flexibility into laws by (re)formulating them using generic terminology that is left undefined or only partially defined. 'Data', 'information', and 'communication' are three terms in point.

A fourth factor is that such terms have nebulous boundaries and are, concomitantly, open to various interpretations. In particular, both 'data' and 'information' can connote something corporeal as well as items or processes that are more intangible. Accordingly, a rule that turns on the meaning of such terms risks being over-inclusive. By 'over-inclusive' is meant that the rule may 'include within its scope conduct to which it is not appropriate, in that there is a poorer correlation between the rules and its justification when applied to that conduct than otherwise'.[17] There is, in other words, a risk of what Brownsword terms 'normative disconnection'.[18] A basic worry then is that if lawmakers fail to take steps to clarify the terms' meanings, they fail not only to provide sufficient legal certainty for those affected by the rules in which the terms feature, but to limit the rules' potential for over-inclusiveness and to ensure proper regulatory connection. Of course, *ex post* clarification, primarily in the form of case law, can offset this worry. Yet, it will not necessarily offset another, perhaps more serious worry—and might even exacerbate it: the paucity of

[15] See eg Irma van der Ploeg, 'Biometrics and the Body as Information: Normative Issues of the Socio-Technical Coding of the Body' in David Lyon (ed), *Surveillance as Social Sorting: Privacy, Risk, and Automated Discrimination* (Routledge 2002) ch 3.

[16] In Solum's sense of the word: see further Lawrence B Solum, 'Indeterminacy' in Dennis M Patterson (ed), *A Companion to Philosophy of Law and Legal Theory* (2nd edn, Blackwell Publishers 2010) 479, 480–81.

[17] Lyria Bennett Moses, 'Adapting the Law to Technological Change: A Comparison of Common Law and Legislation' (2003) 26 U New South Wales LJ 394, 400.

[18] Roger Brownsword, *Rights, Regulation, and the Technological Revolution* (OUP 2008) 166.

legislative guidance aggravates the risk of courts or other actors expanding the laws' scope of application to, say, biological material in a haphazard way, perhaps well beyond what legislators actually intended, and without adequately reflecting over the consequences.

The structure of the article is as follows. After this introduction, the article presents two case studies highlighting the problems outlined above. It then explicates, firstly, the increase in express reference to basic information concepts in legal rules and, secondly, lawmakers' reluctance to provide detailed definitional guidance on the concepts and terms in question. The propriety of such reluctance is then assessed. Attention is directed at information concepts' polysemantic character, points of disagreement over the appropriate properties of data, information, and communication, and the legal relevance of such disagreement. The article concludes with a discussion of the principal regulatory implications of its findings.

Given the subject matter of the article, it would be remiss not to provide working definitions of 'data', 'information', and 'communication'. Unless otherwise stated, references to 'data' denote signs, patterns, characters, or symbols that can convey information about facts, concepts, processes, or objects. References to 'information' denote the meaning (semantic content) assigned to data, while 'communication' refers to the process whereby information is imparted from one entity to another. These definitions conform roughly to those often applied in informatics.

2. *Case Studies*

The following case studies derive from different areas of law: data protection and intellectual property. Each study concerns a legal issue or set of issues that turns on the meaning of a term embodying a basic information concept. The significance of the issue(s) is augmented by the fact that the terms in question are employed in legal instruments of a large number of jurisdictions including international codes that are high in the legal hierarchy of their respective fields and, thus, exercise considerable influence over 'lower-order' codes in those fields. Although involving legal-dogmatic analysis, the ultimate point of the studies is not to ascertain the state of law as it is but to elucidate some of the consequences of ambiguity in the relevant terms and concepts.

A. *'Bodily' Data and Information*

The first case study concerns the question of whether the terms 'personal data' and 'personal information' in data protection law may extend to human biological material *per se* (eg blood cells as opposed to any data or information that can otherwise be derived from the cells). This is a key question for how data protection law relates to the human body and for the extent to which such

law may regulate the use of human biological material, particularly in biobanks. It is also an important point of departure for assessing the propriety of using data protection law as an instrument for biobank regulation. I have dealt with this question at length in another article,[19] so treatment of it here is brief.

Most data protection legislation fails on its face to provide a definite answer to the question. It simply defines 'personal data' or 'personal information' as data or information that can be linked, directly or indirectly, to identifiable individual persons.[20] The precise meaning of the terms 'data' and 'information' in those definitions has usually been taken for granted, with most effort directed at elucidating the link between data or information and persons.[21] Certainly, the architects of data protection law have generally intended to give the terms a broad ambit and to draft the legal provisions in a technology-neutral way. The preparatory works for the EU Data Protection Directive are a case in point.[22] Yet, they fail to indicate conclusively that 'personal data' in the Directive is intended to embrace human biological material as such.[23] The same can be said for many other data protection laws and their preparatory works.[24] This underdeterminacy compounds the laws' already considerable potential for over-inclusiveness—a potential brought on not only by the expansive breadth of their definitions of 'personal data' but also by the frequent obscurity of their aims.[25]

Scholars, lawmakers, and other policy entrepreneurs diverge in their respective stances on how this underdeterminacy should be resolved.[26] Many of those who take the view that human biological material does not constitute

[19] Lee A Bygrave, 'The Body as Data? Biobank Regulation via the "Back Door" of Data Protection Law' (2010) 2 Law, Innovation and Technology 1.

[20] See eg Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31 (hereinafter 'Data Protection Directive'), art 2(a); Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (hereinafter 'Data Protection Convention'), art 2(a).

[21] See eg Lee A Bygrave, *Data Protection Law: Approaching its Rationale, Logic and Limits* (Kluwer Law International 2002) 41–50, 210–15, 315–19.

[22] See the European Commission's commentary with respect to its amended proposal for a data protection Directive of 15 October 1992: COM (92) 422 final – SYN 287, 9. See also its commentary with respect to the original Directive proposal of 24 September 1990: COM (90) 314 final – SYN 287, 15.

[23] See further Bygrave, The Body as Data?' (n 19) 14–15 and references cited therein.

[24] On the other hand, the European Commission's recently proposed General Data Protection Regulation to replace Directive 95/46/EC seems to distinguish between biological material and personal data: see Proposal for a Regulation on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) (COM(2012) 11 final). The proposal elaborates what constitutes 'personal data relating to health' by stating that such data 'should include . . . information derived from the testing or examination of a body part or bodily substance, including biological samples' (recital 26). The syntax in the English version (and, indeed, Danish, Swedish, and French versions) is not entirely free of ambiguity: while 'biological samples' most likely refers to 'body part or bodily substance', it could also be read as referring back to 'information derived . . .'. The German version, however, rules out this possibility ('Informationen, die von der Prüfung oder Untersuchung eines Körperteils oder einer körpereigenen Substanz, darunter biologischer Proben, abgeleitet wurden').

[25] Further on these aspects of the laws' aims and scope, see eg Lee A Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 117–18, 135.

[26] See generally Deryk Beyleveld, David Townend, Ségolène Rouillé-Mirza, and Jessica Wright (eds), *Implementation of the Data Protection Directive in Relation to Medical Research in Europe* (Ashgate 2004); Bygrave (n 19) 12, 16–20.

personal data or information ground their stance in conceptual logic, claiming that data is a formalized representation of objects or processes, while information comprises a cognitive element involving comprehension of the representation. On that basis, biological material as such cannot be treated as information and its treatment as data is doubtful.[27] I discuss this claim in more detail in the penultimate section of the article.

Those pushing the view that biological material may be personal data or information tend to pay more regard to pragmatic considerations than conceptual logic. Chief amongst such considerations are the need to fill lacunae in biobank regulation, the growing ease with which persons can be identified from biological material, and the fact that such material is often only stored for generating information.[28]

Sometimes, the decision is made to treat biological material as personal data without a full explanation being given. A significant instance is the judgment of the European Court of Human Rights (Grand Chamber) in *S and Marper v UK*.[29] The case dealt with the legality of police agencies retaining DNA profiles, fingerprints, and cellular samples taken from persons who have been suspected of having committed legal offences. On its way to unanimously concluding that the retention practice violated article 8 of the European Convention on Human Rights, the Strasbourg court held, in a single sentence, that the cellular samples (along with the fingerprints and DNA profiles) 'constitute personal data within the meaning of the Data Protection Convention as they relate to identified or identifiable individuals'.[30] It elaborated no further. The finding is remarkable not just for its terse formulation and reasoning but for the fact that the legal sources upon which it is based offer scant, if any, firm support for it.[31]

Although the *Marper* line does not create a legally binding precedent for interpreting the Data Protection Convention, it is highly persuasive authority and undoubtedly creates pressure to apply the Convention to a range of activities that would fall outside the Convention's ambit were samples or other biological material not to be regarded as personal data. It also creates pressure to treat other 'lower-order' data protection codes more directly as instruments of biobank regulation (or transform them into such) in the interest of maintaining regulatory consistency down the normative chain.

It cannot be assumed that such an extension will sit happily with the respective history, rationale, and customary application of the lower-order instruments, or enable their sensible application to biobanks. Some of these instruments distinguish between samples and data or information.[32] Moreover,

[27] Bygrave (n 19) 15–17, 19–20.
[28] For examples, see ibid 16, 19 and references cited therein.
[29] *S and Marper v UK* (2009) 48 EHRR 50.
[30] ibid para 68.
[31] See further Bygrave (n 19) 8–9.
[32] For examples, see ibid 9, 12.

applying data protection laws to regulate biobanks will usually involve an extension of the laws—and thereby an extension in the scope of data protection authorities' competence—beyond what the legislators have consciously considered. While good grounds exist for closely calibrating biobank regulation with the principles of data protection law,[33] and while it is tempting to apply the latter to body samples given the parlous state of biobank regulation under other codes,[34] such application is unlikely to be 'plain sailing'. At the very least, the rules in data protection laws will need fine-tuning if extended to biological material. Some of the current rules (eg the requirement that personal data be 'accurate', or erased, or rectified if inaccurate)[35] are logically difficult if not impossible to apply to biological samples. Moreover, usage of some biological material (eg DNA) involves weighing up group interests that are not readily accounted for by current data protection rules.[36]

Thus, the decision to extend data protection laws to biological material ought not to be taken lightly but be preceded by careful discussion. When assimilating the human body into the body of data protection law, we should avoid quick fixes, especially when they are not the product of considered thought.

B. *Data in Databases*

Database protection under intellectual property law is another context where lawmakers' failure to address thoroughly the meaning of basic information concepts causes difficulty. The difficulty concerns legislative definitions of the terms 'database' and 'compilation'. The open-ended, ambiguous way in which the terms are defined has generated questions about their meaning, in particular, over whether or not they embrace biobanks along with other collections of physical objects. Again, I have dealt with this question at length in another article,[37] so treatment of it here is brief.

The point of departure for the discussion is the EU Database Directive.[38] With few qualifications, this protects databases 'in any form' (article 1(1)). A database is defined as 'a collection of independent works, data or other materials arranged in a systematic or methodical way and individually

[33] See further ibid 20–22; Mark Taylor, *Genetic Data and the Law: A Critical Perspective on Privacy Protection* (CUP 2012) ch 7.

[34] As Gibbons and Kaye note, 'to date we still have no clear, co-ordinated or coherent framework for governing genetic databases and related activities, even at the domestic level': Susan MC Gibbons and Jane Kaye, 'Governing Genetic Databases: Collection, Storage and Use' (2007) 18 King's LJ 201, 203.

[35] See eg Data Protection Directive, art 6(1)(d).

[36] Consider eg the issue as to when a person ought to be accorded a right of access to bodily samples of their first-degree genetic relatives (ie siblings, parents, or children), along with the issue of whether that person ought to be able to co-determine (via consent requirements) how such samples are used. For discussion, see eg Australian Law Reform Commission (ALRC) and Australian Health Ethics Committee (AHEC), *Essentially Yours: The Protection of Human Genetic Information in Australia*, Report No 96 (ARLC/AHEC 2003) [8.48]–[8.54]; Taylor (n 33) 116–30.

[37] Lee A Bygrave, 'The Data Difficulty in Database Protection' (2013) 35 EIPR 25.

[38] Directive 96/9/EC on the legal protection of databases [1996] OJ L77/20.

accessible by electronic or other means' (article 1(2)). It is primarily the ambiguous phrase 'data or other materials' combined with the apparent disregard for 'form' that ignites the discussion. A similar discussion could arise with respect to the term 'compilation' which is employed in certain other intellectual property legislation as an equivalent to the database concept. That legislation requires the protection of 'compilations of data or other material' largely regardless of their form.[39]

However, the *gravitas* of the questions is heightened in the context of the Database Directive as the way in which the questions are answered bears on the application of the internationally unique and controversial *sui generis* database right provided by the Directive. The right prevents the contents of a database from being extracted or reutilized—independent of any copyright protection that might otherwise apply—where 'substantial investment' has gone into the 'obtaining, verification or presentation of the contents' (article 7(1)). The right has been strongly criticized on multiple grounds. Among these are accusations that it lacks economic justification, engenders improper informational lock-up, and fails to provide adequate legal certainty.[40]

Minimal guidance on the database definition exists in the legislation, and the borders it constructs around the term 'database' are porous. The legislation fails to indicate in what sense it employs the terms 'data' and 'materials'. Both terms can connote something solid and tangible yet they can also connote purely informational items (text, images and the like). Nonetheless, the preparatory works clearly evidence databases as just having informational constituents.[41] Moreover, the first proposal for the Directive only covered electronic databases. This limitation soon became untenable for various reasons.[42] In my opinion, the limitation and the subsequent need to overcome that limitation cast light on the breadth of the database definition in article 1(2). The legislator had to adopt a definition that was sufficiently generic to (i) encompass both electronic and non-electronic collections of data, (ii) take account of technological developments, and (iii) serve the purposes of both copyright and the *sui generis* right. The reference to 'other materials' was probably inserted to neutralise the connotation of the term 'database' as a *computerized* collection of data.[43] Unfortunately, the result of this generic dream was confusingly broad terminology.

[39] See eg the World Intellectual Property Organisation (WIPO) Copyright Treaty 1996, art 5; World Trade Organization (WTO) Agreement on Trade-Related Aspects of Intellectual Property Rights 1994 (hereinafter 'TRIPS Agreement'), art 10(2); United States (US) Copyright Act 1976 (17 USC), s 101.

[40] For an extensive overview and assessment of these and other criticisms, see Estelle Derclaye, *The Legal Protection of Databases: A Comparative Analysis* (Edward Elgar 2008) especially ch 2.

[41] See Bygrave (n 37) and references cited therein.

[42] See further eg Estelle Derclaye, 'What is a Database? A Critical Analysis of the Definition of a Database in the European Database Directive and Suggestions for an International Definition' (2002) 5 J World Intellectual Property 981, 984 and references cited therein.

[43] The *Oxford Dictionary of English* (3rd edn, OUP 2010) defines 'database' as 'a structured set of data held in a computer'.

The informational focus apparent in the preparatory works has been emphasized in case law. According to the EU Court of Justice (CJEU), the database concept must be defined in light of the Directive's intended function, which is 'to encourage the development of systems performing a function of "storage" and "processing" of information, according to the 10th and 12th recitals'.[44] The Court, though, has not yet clarified the precise meaning of 'data or other materials' in article 1(2). The closest that it has come to such clarification is to hold that generating or creating data is not the same as 'obtaining' it, and that investment in the former activity thereby fails to attract application of the *sui generis* right.[45]

Distinguishing between the creation of data and the collection of existing data can be difficult. This is evidenced in the football-fixture cases where the Court differed from Advocate-General Stix-Hackel as to the classification of the drawing up of fixture lists. In the Advocate-General's view, the process involved the collection of existing data;[46] the Court held otherwise.[47] The distinction between creating and obtaining data becomes especially difficult in the context of scientific research, where natural phenomena are observed and analysed. When do those processes generate data? Some scholars argue that the natural phenomena themselves are data such that analysis of them resulting in recordings of co-ordinates, measurements, etc involves pre-existing data being collected.[48] As is made clear further on in this article, such an argument finds support in some informatics-based explications of data (though the scholars concerned do not refer to these). A decision by the Munich District Court (Landgericht) also provides support for the argument, although it is somewhat ambiguous on the point.[49] My hunch is that the CJEU would be unlikely to accept the argument. It would rather look for when the formalized representations of the natural phenomena (typically in the form of recorded

[44] Case C–444/02 *Fixtures Marketing Ltd v Organismos Prognostikon Agonon Podosfairou (OPAP) (Greece)* [2004] ECR I–10549, para 28. Recital 10 states that 'the exponential growth, in the Community and worldwide, in the amount of information generated and processed annually in all sectors of commerce and industry calls for investment in all the Member States in advanced information processing systems'. Recital 12 states that 'such an investment in modern information storage and processing systems will not take place within the Community unless a stable and uniform legal protection regime is introduced for the protection of the rights of makers of databases'.

[45] Case C–203/02 *British Horseracing Board Limited and Others v William Hill Organisation Ltd* [2004] ECR I–10415, paras 31ff; Case C–46/02 *Fixtures Marketing Ltd v Oy Veikkaus AB AB (Finland)* [2004] ECR I–10365, paras 34ff; Case C–338/02 *Fixtures Marketing Ltd v Svenska Spel AB (Sweden)* [2004] ECR I–10497, paras 24ff; *OPAP* (n 44) paras 40ff.

[46] *OPAP*, Opinion, para 74; *Svenska Spel*, Opinion, para 58; *Veikkaus*, Opinion, para 68.

[47] *Svenska Spel* (n 45) para 31. See too *OPAP* (n 44) para 47; *Veikkaus* (n 45) paras 42–44.

[48] See eg Ole-Andreas Rognstad, *Opphavsrett* (Universitetsforlaget 2009) 302–3; Matthias Leistner, '*British Horseracing Board v William Hill*' (2005) 36 Intl Rev Intellectual Property and Competition L 592, 594 (comment).

[49] Judgment of 9 November 2005, reported in *Gewerblicher Rechtsschutz und Urheberrecht* (GRUR) 2006, 225 (holding that the process of producing a database consisting of a topographic map of a natural landscape involved the collection of pre-existing data). See further Bygrave (n 37) 31.

measurements) are first made and regard that process as creating, rather than obtaining or collecting data.[50] This approach would seem to tally better with the Directive's legislative history and aims.

However, it is not inconceivable that the Court would judge certain types of biobanks as fitting within the database concept, thus emulating, in a way, the judicial line taken in the *Marper* case. The Court would probably do so if it regarded the biobank as primarily filling an informational function such that its protection, in the Court's view, could be justified against the considerations mentioned in recitals 10 and 12 of the Directive's preamble—considerations emphasized in the OPAP case.[51] A pertinent example could be a collection of blood samples for genetic testing, managed as part of a commercial enterprise (eg development and sale of new pharmaceutical products).[52]

The issue of database protection for biobanks appears not yet to have been the subject of litigation, and a more general discussion about the precise meaning of 'data or other materials' has hitherto been largely confined to the ivory tower of academia. Yet, the issue remains important in principle and in the broader regulatory context. This is especially so in light of the overprotective potential of the *sui generis* database right together with the Directive's harmonization objective. Accordingly, just as human body samples should not be assimilated into the body of data protection law in a manner akin to the *Marper* 'one liner', so should careful discussion precede any extension of database protection to biobanks.

3. *Growing Prominence of Generic Information Concepts in Law*

Do the issues identified in the two case studies exist just within the fields of data protection and intellectual property law? I argue they do not. Similar problems exist in other fields of law, although their form and significance can vary from the case studies due to differences in terminological, technological, or transactional context. Such problems do not exist across all legal fields; nonetheless, they pop up—or could pop up—in various guises in a significant number of areas. Their compass reflects growth of legislation involving the use of basic information concepts as 'front-line' statutory terms that are decoupled from a narrowly specified technological platform or medium.

[50] See too Mark J Davison and P Bernt Hugenholtz, 'Football Fixtures, Horse Races and Spin-offs: The ECJ Domesticates the Database Right' (2005) 27 EIPR 113, 115; Jens L Gaster, *Der Rechtsschutz von Datenbanken* (Carl Heymanns Verlag 1999) 37.

[51] See further Bygrave (n 37) 31–32.

[52] This assumes, of course, that the blood samples already existed, such that the 'substantial investment' required for art 7(1) protection was directed at 'obtaining' the samples (as opposed to generating them).

Examples include administrative law and criminal law. In administrative law, we see, for instance, freedom of information (FOI) statutes now frequently employing 'information' as the basic unit of disclosure, either directly[53] or more indirectly, by defining the formal primary unit of disclosure (such as 'record' or 'document') in terms of 'information'.[54] Turning to criminal law, particularly pertinent statutory provisions are those criminalizing unauthorized access to, alteration or destruction of 'data' in computer systems.[55] Basic information concepts also come to the fore in rules on freedom of expression[56] and electronic communications.[57] Additionally, in intellectual property law, basic information concepts are not only applied in the protection of databases, but also in, say, the protection of trade secrets.[58]

The above list of relevant legal fields does not pretend to be exhaustive. Further, the true operational impact of basic information concepts cannot be measured simply in terms of their employment as 'front-line' statutory terms. They also figure in other legal sources (typically preparatory works) to explain statutory provisions employing related concepts, such as 'document', 'record', or 'file'.[59] Moreover, other generic terms are sometimes inserted in legislation to fulfil the role of a basic information concept and render the law more technology-neutral.[60]

Legislative use of basic information concepts seems to have grown markedly since the late 1980s. This does not mean that they were previously absent from legislation. Nor does it mean that legislation or law more generally was not

[53] See eg the Official Information Act 1982 of New Zealand (NZ); Freedom of Information Act 2000 of the United Kingdom (UK); and Government Information (Open Access) Act 2009 of New South Wales (NSW).

[54] See eg Ireland's Freedom of Information Act 1997 which employs 'record' as the basic unit of disclosure and defines the term as encompassing 'any form . . . (including machine-readable form) or thing in which information is held or stored manually, mechanically or electronically' (s 2(1)). Cf Canada's federal Access to Information Act (Revised Statutes of Canada (RSC) 1985, c A-1) which also employs 'record' as the basic unit of disclosure but defines this as 'any documentary material, regardless of medium or form' (s 3).

[55] See eg the Council of Europe Convention on Cybercrime 2001 (hereinafter 'Cybercrime Convention'), arts 2–5; Canada's federal Criminal Code (RSC 1985, c C-46), s 430(1.1) dealing with 'mischief relating to data'; Norway's General Civil Penal Code 1902 (*Almindelig borgerlig straffelov 22 mai 1902 nr 10*), § 145(2) dealing with 'unjustified access to data' (see further below).

[56] See eg European Convention on Human Rights, art 10(1), which includes freedom 'to receive and impart information'.

[57] See eg United Nations (UN) Convention on the Use of Electronic Communications in International Contracts 2006 (not yet in force) (hereinafter 'Electronic Communications Convention' or 'ECC'), art 4(b) (defining electronic communication in terms of 'data messages'); UK Communications Act 2003, s 32(1) (defining an 'electronic communications network' as a 'transmission system for the conveyance, by the use of electrical, magnetic or electro-magnetic energy, of signals of any description') and s 32(10)(a) (defining 'signal' in terms of 'data of any description').

[58] See eg TRIPS Agreement, art 39 on the protection of 'undisclosed information' (in effect, trade secrets).

[59] See eg the preparatory works for Norway's Freedom of Information Act 2006 (*lov om rett til innsyn i dokument i offentleg verksemd 19 mai 2006 nr 16*), particularly *Ny offentlighetslov*, Norges Offentlige Utredninger (Norwegian Official Reports; abbreviated as 'NOU'), 2003 no 30, 80–81, 257 and Odelstingsproposisjon (Government Bill) no 102 (2004–05) 119–20, both of which employ basic information concepts to explain the meaning of the term 'dokument' (document) in § 4(1) of the legislation.

[60] See eg use of the term 'content' in the definition of 'document' in Directive 2003/98/EC on the re-use of public sector information (hereinafter 'PSI Directive') [2003] OJ L345/90, art 2(3). According to the Directive's preamble, the intention is to 'lay down a generic definition of the term "document", in line with developments in the information society' (recital 11).

already substantially engaged in the regulation of information. There has always been extensive legal regulation of information, though often in indirect forms focused on certain communication contexts. These contexts are (or have been) often constituted and delimited by reference to certain types of *media* (eg paper, computers), *service* (eg broadcasting, telephony), *negotiation* (eg contract), *decisional process* (eg decision-making by government agencies), or combinations of these. Many contract formation rules, for instance, deal essentially with communication and quality of information, although they are clothed in other context-specific terminology, such as 'offer', 'invitation to treat', '(mis)representation', and 'acceptance'.

The delineation of context has helped to anchor the relevant legal rules, providing them with a measure of certainty and reducing the risk of over-inclusiveness. The role played by media is particularly important in this regard. Laws are often framed so that they specifically govern particular media rather than the data and information that they carry. Even laws that are primarily aimed at regulating data or information frequently require for their application that information be embodied in a specified medium. For instance, FOI laws often apply only to data or information in some sort of documentary or recorded form.[61] In other words, they will not apply to information held simply in a person's mind. Some data protection laws also operate with such a limitation.[62] This media-specific focus of law partly reflects distaste for direct interference with purely cognitive processes; lawmakers are reluctant to engage in thought control. It also reflects the concern for legal certainty and the practicability of enforcement.

4. *Loosening Media Moorings*

The above elaboration of the role traditionally played by media in the application of law helps to throw into relief the legislative developments that are the main concern of this article. In many respects, the developments involve loosening legal rules from their media moorings. The extent of such loosening affects the extent to which the rules become vulnerable to the sorts of problems highlighted in the case studies. Consequently, not all instances in which legal rules are recast in more generic terminology will result in their possible or ready application to aspects of the organic world. This is the case, for example, with instances of electronic communications law which, despite

[61] See eg the references to Irish and Canadian legislation, n 54 above. Application of the equivalent UK legislation also presumes that the information be held in recorded form: see Freedom of Information Act 2000, s 84; Freedom of Information (Scotland) Act 2000, s 73.

[62] See eg Hong Kong's Personal Data (Privacy) Ordinance of 1995 which applies only to information that is in the form of data held in a 'document' (s 2).

adopting a broad definition of the signals being communicated, specify a particular *channel* for signal transmission, such as 'free space' or 'cable'.[63] However, we cannot assume that retention of the criteria 'electronic', 'electromagnetic', 'automated' or 'automatic' as prerequisites for applying law[64] will, in itself, necessarily steer law clear of the organic world. It is trite to observe that biological processes are both generators and conductors of electricity, and that they can be automated in the sense of them occurring without human intervention. Retention of other contextual factors may nonetheless help to prevent law being applied to biological material or processes. Thus, the UN Electronic Communications Convention makes clear that it applies only to communication of data messages in the context of 'the formation or performance of a contract' (article 4(a)).

Yet, the legislative delineation of such contextual factors often fails unequivocally to prevent the law concerned from being able to drift into the biological world. This is particularly the case when the key terms for its application are also generically formulated. We see this in the case studies. We can also discern it in criminal law and administrative law. Take, for example, the Cybercrime Convention. While its title, preamble, and explanatory report evidence that the Convention is aimed at combatting criminal activity in cyberspace—which is commonly understood as a virtual dimension and, as such, quite separate from the organic world—delineation of its scope of application is otherwise vague.[65] And the key terms upon which its application turns—'computer system' and 'computer data'—are not defined in their ordinary literal sense but in a more generic way, which might encompass certain biological processes, such as biocomputing. The Convention defines a 'computer system' to include 'any device or a group of interconnected or related devices, one or more of which, pursuant to a program, performs automatic processing of data' (article 1(a)). The terms 'device' and 'program' are not defined in the Convention. The explanatory report fails to elaborate on the term 'device', although it refers to a 'program' as 'a set of instructions that can be executed by the computer to achieve the desired result'.[66] The term 'computer data' is defined as 'any representation of facts, information or concepts in a form suitable for processing in a computer system' (article 1(b)). The sole criterion posing a substantial conceptual hindrance to the

[63] See eg the Norwegian Electronic Communications Act 2003 (*Lov om elektronisk kommunikasjon 4 juli 2003 nr 83*) s 1-5(1) which defines 'electronic communication' as 'transmission of sound, text, pictures or other data with the aid of electromagnetic signals in free space or by cable in a system for signal transmission'.

[64] The case, for instance, with the UN ECC and UK Communications Act (n 57).

[65] This vagueness extends also to what the Convention's Explanatory Report refers to as 'cyber-space' (*sic*), which is simply referred to as the 'common space' created by interconnection of 'communication and information services': Explanatory Report to the Convention on Cybercrime (Council of Europe, 8 November 2001) para 8. Note that the report 'does not constitute an instrument providing an authoritative interpretation of the Convention, although it might be of such a nature as to facilitate the understanding of the provisions contained therein': ibid para II.

[66] ibid para 23.

Convention's application to biological material or processes is the reference to 'representation'.[67] The hindrance is based on an assumption that 'data', as a 'representation', must be an abstraction and accordingly separate from the natural or biological world. I discuss the strength of that assumption towards the end of the article.

Turning to administrative law, generic information concepts are gradually usurping the place traditionally played by 'document', 'record', and the like. As indicated above, this is happening partly through replacement of the latter by the former as 'front-line' statutory terms, or the latter are retained in that role, but are redefined so that they are, in effect, reduced to their informational bones. In both cases, the rules are being significantly loosened from their traditional media moorings. At the same time, the moorings themselves are becoming looser as their specification becomes more generic. Consider, for instance, the PSI Directive, which defines 'document' to cover 'any content *whatever its medium* (written on paper or stored in electronic form or as a sound, visual or audiovisual recording)' (art 2(3); emphasis added). The definition's media specification may perhaps be tightened by the elaboration in parentheses; if so, this would make it difficult, if not impossible to consider, say, a bioprint[68] as a document for the purposes of the Directive. We cannot be sure, though, that the parenthetical elaboration forms the outer boundaries of the documentary medium here. Media specification under equivalent FOI regimes in certain other jurisdictions is even looser than under the PSI Directive. The UK Freedom of Information Act, for example, applies to information 'recorded in any form' (section 84).[69] Thus, its potential application to a bioprint is conceptually easy.

5. *Technology and Generic Dreams*

The ever-greater legal salience of basic information concepts has a complex aetiology and its full explication lies beyond the scope of this article. For present purposes—and at the clear risk of oversimplification—the development is due in considerable part to the effects of computer technology on information processing, combined with a determination to regulate these effects in line with certain values.

[67] According to the Explanatory Report, the apparent breadth of the definition is also undercut by a requirement that the data concerned must be 'put in such a form that it can be *directly* processed by the computer system' (emphasis added): ibid para 25. However, that requirement scarcely hinders application to the biological world.

[68] This denotes a replication of biological material using processes that are functionally akin to the printing of paper documents: see eg <www.explainingthefuture.com/bioprinting.html> accessed 17 April 2014. The technology is still in its infancy and its potential is shrouded in both hype and controversy: see eg <http://angelasaini.blogspot.it/2011/09/could-we-really-print-body-parts.html> accessed 17 April 2014. Nonetheless, jurists and other policy entrepreneurs would be foolish to write it off as a permanent part of science fiction.

[69] See further Philip Coppel, *Information Rights: Law and Practice* (3rd edn, Hart 2010) 333 [9-003] and references cited therein, and noting that '[t]he medium on which matter is recorded should not, in principle, impinge upon its characterisation as information'.

Both digitization (ie the conversion of data into binary digital signs) and convergence (in this context, the merger of technological-organizational platforms for data processing and delivery of services) have been particularly salient in pushing lawmakers to recast legal rules using relatively generic information concepts. Both processes have loosened information from its paper medium—an outcome often summed up in the popular, though somewhat misleading phrase of the 1980s that 'the computer has set information free'.[70] In light of this 'liberation', lawmakers have set about attempting to liberate legal rules from the boundaries of paper and sometimes other relatively specific media with fixed physical and logical 'walls'. The attempt is exemplified by the comprehensive round of law reform aimed at ensuring that the bulk of transactions carried out electronically are ascribed similar legal status to non-electronic transactions.[71] It is also exemplified in the shift in focus of data protection law from regulating systematically structured collections of personal data in the form of 'registers' or 'files'—a focus typical for data protection instruments drafted in the 1970s and early 1980s—to regulating the processing of personal data largely regardless of how the data is organized.[72]

The efforts of lawmakers here cannot simply be explained as responses to digitization, convergence, and other technological-organizational developments. The nature of their response also has an ideological dimension as manifest in their choice of particular strategies to cope with rapid technological change. These strategies embody various guiding regulatory principles, ideals, and demands. At a basic level, they embody a demand for legal sustainability or 'futureproofing'. This is a demand that laws are capable, for the foreseeable future, of meeting their goals in the face of technological flux, thereby reducing the number of times the Sisyphean stone of legal reform has to be rolled uphill.[73] The demand is so intrinsic to the lawmaking process, and so uncontroversial, that it is rarely singled out for special critique or reflection.

Linked to that demand are the guiding principles of 'technology neutrality' and 'functional equivalence'. These have developed into basic regulatory axioms, particularly for coping with changes in ICT. The principle of 'technological neutrality' has several connotations and dimensions, the essential one being a demand that legal rules (either in their formulation or effects), along with regulation more generally, ought not to favour unduly one particular

[70] See eg Jon Bing, 'Information Law?' (1982) 2 J Law and Media Practice 219.

[71] See especially the work of the UN Commission on International Trade Law (UNCITRAL) with respect to its Model Law on Electronic Commerce 1996, particularly arts 5–12, and the UN Electronic Communications Convention (ECC), particularly its preamble (para 5) along with arts 8, 9, and 12.

[72] See eg Bygrave, *Data Protection Law* (n 21) 50–52 and references cited therein.

[73] See too Ian Walden, *Computer Crimes and Digital Investigations* (OUP 2007) 61 [2.178] and Bert-Jaap Koops, 'Should ICT Regulation Be Technology-Neutral?' in Bert-Jaap Koops, Miriam Lips, Corien Prins, and Maurice Schellekens (eds), *Starting Points for ICT Regulation: Deconstructing Prevalent Policy One-Liners* (TMC Asser Press 2006) 77, 87–89.

technology over another.[74] In practice, avoiding such discrimination will often require legal rules to be (re-)drafted so that their application is not exclusively tied to a particular or narrowly defined technology.

As for 'functional equivalence', the thrust of this principle posits that what is legally possible using non-electronic means should, as a point of departure, be legally possible using electronic means—a norm that is sometimes summed up in the phrase 'what holds off-line should also hold on-line'. Like technology neutrality, functional equivalence has several connotations.[75] As a substantive regulatory guideline, its application essentially involves (i) analysing the basic functions of paper-based or other 'off-line' documents and activity, and (ii) ensuring that, where desirable, the regulatory framework permits reproduction of the equivalent functionality in the digital or on-line environment. In practice, the principle tends to be intertwined with technology neutrality and serves frequently as a goal for the latter.[76]

For the purposes of this article, the especially noteworthy aspect of these principles is that they encourage use of relatively generic information concepts in law. While this point is noted at least indirectly in some legislative materials (chiefly preparatory works) and law reform proposals,[77] it has attracted surprisingly little comment in academic literature. I am not suggesting that application of the principles always necessitates making greater use of generic information concepts in legal instruments. Yet, their logical thrust (particularly that of technology neutrality) involves replacing much of the conceptual apparatus that arose in the 'paper age' with more obviously generic concepts.

6. *Scarcity of Statutory Definitions of Information Concepts*

Although basic information concepts are increasingly employed as 'front-line' legislative terms, lawmakers often leave them either undefined or only partially defined. For instance, the term 'data' is not defined in the UK Computer Misuse Act 1990, nor is the term 'information' defined in the UK Freedom of Information Act 2000. As highlighted in the case studies and further below, numerous examples exist in other areas of law too.

In some cases, the paucity of definition might be simply the result of oversight. In other cases, it can reflect a view that clarification is unnecessary

[74] For in-depth analysis of the meanings, realism, and desirability of the principle, see Koops, ibid and Chris Reed, 'Taking Sides on Technology Neutrality' (2007) 4:3 SCRIPTed 263 <hwww.law.ed.ac.uk/ahrc/script-ed/vol4-3/reed.asp> accessed 17 April 2014.

[75] For discussion of these connotations, see Maurice Schellekens, 'What Holds Off-Line, Also Holds On-Line?' in Koops and others (n 73) 51–75; Chris Reed, 'Online and Offline Equivalence: Aspiration and Achievement' (2010) 18 Intl J Law and Information Technology 248.

[76] See too Koops (n 73) 85; Schellekens, ibid 75. However, as Reed points out (ibid), considerable obstacles lie in the way of achieving full equivalence.

[77] See eg Sunde (n 2) ch 6 and references cited therein.

because the term in question has an obvious ordinary meaning and is to be applied in accordance with that meaning. This is the case with the term 'information' in New Zealand's Official Information Act 1982. In *Commissioner of Police v Ombudsman*,[78] McMullin J of the NZ Court of Appeal stated:

> Information is not defined in the [Official Information] Act. From this it may be inferred that the draftsman was prepared to adopt the ordinary dictionary meaning of that word. Information in its ordinary dictionary meaning is that which informs, instructs, tells or makes aware.[79]

Similarly, the UK Law Commission took the view that providing a definition of 'data' in the Computer Misuse Act 1990 was unnecessary, as the term's putatively ordinary meaning as 'information or facts held on a computer' sufficed.[80]

In yet other cases, the paucity of definition can reflect a view that the term in question is incapable of being defined in a legally useful way or, indeed, at all. As noted above, some legal scholars have taken this view.[81] While not all legislators share such pessimism to the same degree, they do sometimes doubt their ability to provide a statutory definition of 'information' or 'data' that is sufficiently precise to serve a legally useful purpose. The doubt is manifest in the decision of the Norwegian Ministry of Justice to omit definitions of the terms 'data' and 'datasystem' in Norway's Penal Code 2005,[82] despite the legislation containing a raft of provisions expressly aimed at protecting various forms of data and ICT. In the Ministry's view, both terms are too hard to define with sufficient precision yet flexibility to take account of technological change.[83] Thus, the Ministry's decision is also an instance of lawmakers refraining from defining a term in order to preserve its flexibility and, thereby, to 'futureproof' it. We see this also exemplified in scepticism of lawmakers towards providing a statutory definition of the term 'computer'. The Scottish Law Commission, for instance, has opined that 'since computer technology is advancing so rapidly, any definition even if expressed in terms of function rather than construction, would rapidly become obsolete'.[84]

Fear of obsolescence is often accompanied by a fear of over-inclusiveness. Like the former, the latter fear is exacerbated by concern over technological

[78] [1988] 1 NZLR 385.
[79] ibid 402.
[80] Law Commission, *Computer Misuse* (Law Com No 186, 1989) [3.30].
[81] See references cited in n 6.
[82] *Lov om straff 20 mai 2005 nr 28* (not yet in force).
[83] Odelstingsproposisjon no 22 (2008–09) 21.
[84] Scottish Law Commission, *Report on Computer Crime* (Scot Law Com No 174, 1987) [4.17]. The English Law Commission has shared this scepticism, as have others involved in drafting computer crime legislation in the UK: Martin Wasik, *Crime and the Computer* (Clarendon Press 1991) 5, 78 and references cited therein.

development. Thus, as computer technology becomes increasingly pervasive, there is the worry that a statutory definition of 'computer' risks encompassing a range of devices that are not typically associated with or regarded as computers.[85] The fear is also discernible in the above-mentioned concern of the Norwegian Ministry of Justice about achieving definitional precision in the face of technological change. In the context of criminal law, the fear of over-inclusiveness becomes one of over-criminalization.

The scepticism towards statutory definitions tends to rest on at least two assumptions. One assumption is that a term such as 'information' or 'data' has an ordinary meaning that is sufficiently stable and uniformly understood so that lawmakers may take for granted that the term can be sensibly applied without upfront instruction from them. A second assumption, which can go hand-in-hand with the first, is that those bodies charged with interpreting and applying law—primarily the courts—are willing to construe the undefined terms in a way that keeps pace with technological development or, at least, with 'evolving common understanding of those terms modified, where necessary, by their statutory context'.[86] A variant of this assumption could be a belief that the judiciary will not do a worse job of definition than legislators will, coupled perhaps with a belief that the judiciary can fix its own definitional errors more readily than having errors fixed through the legislative system.[87]

Both assumptions are contestable, though not often contested. The remainder of this article deals largely with the first-mentioned assumption. In doing so, attention is also paid to the above-noted claim by some scholars that 'information' (and perhaps related concepts such as 'data') is too conceptually inchoate to be precisely defined, at least in law. That claim undercuts the assumption of definitional stability. It might also imply that lawmakers ought to avoid or minimise express references to basic information concepts in legal codes.

Before examining the issue of definitional stability, it is pertinent to consider briefly the assumption about courts' willingness to construe undefined terms in line with technological developments. The degree of such willingness depends on judicial culture generally along with a range of more specific factors, such as

[85] This worry arose particularly in the wake of statutory definitions of 'computer' in US legislation from the 1980s: Wasik, ibid 4–5 and references cited therein.

[86] Quoting the Australian Model Criminal Code Officers Committee of the Standing Committee of Attorneys-General, *Model Criminal Code, Chapter 4: Damages and Computer Offences* (Final Report, January 2001) 129 (where the Committee claims that such understanding, rather than statutory definitions, provides 'the safest guides to the meaning of "computer", "data", "program" and like terms').

[87] For discussion of the relative flexibility of the common law in light of technological change, see Bennett Moses (n 17) 403–5.

the prevailing approach to statutory interpretation. For example, a purposive or teleological approach to interpreting legislative terms is arguably more conducive to the ability of law to keep pace with technological change than is an approach requiring adoption of the intrinsic literal meaning of statutory language. The purposive approach is popular with many courts and jurisdictions,[88] but has had to struggle to gain favour with English courts.[89] Yet, even when it does have traction, its ability to accommodate technological change can be frustrated by other factors, such as obscurity of parliamentary intention or strict adherence to the principle that clear legal authority is required for state measures that infringe upon citizens' autonomy.

That principle is particularly weighty in the field of criminal law, where courts are traditionally hesitant to expand the scope of ambiguously formulated statutory offences to the detriment of the accused. Thus, in criminal law especially, legal provisions that have been drafted in relatively generic terms in order to keep abreast of technological developments can end up being inflexibly applied by courts.

An example of relevant judicial inflexibility comes from case law of the Norwegian Supreme Court (Høyesterett). In the late 1980s, the Norwegian Penal Code 1902 was amended to improve its ability to combat computer crime. The amendments included a new provision, making it a criminal offence to gain unjustified access to protected 'data or software which are stored or transferred by electronic or other technical means' (section 145(2)). In the case concerned, the Supreme Court was called upon to determine whether use of 'pirate decoders' to decrypt scrambled television signals sent by satellite constituted such an offence.[90] A 3–2 majority concluded in the negative, holding that the term 'data' covered neither television programmes nor signals: 'data' was rather to be understood in terms of 'computer-processed information' ('EDB-basert informasjon'), which was 'on general understanding something different than television programmes'.[91] While acknowledging the legislators' intent to give 'data' a broad meaning and that 'policy considerations' ('reelle hensyn') favoured criminal punishment of the marketing of pirate decoders, the majority held that the preparatory works failed to provide a sufficiently solid basis for construing 'data' expansively.[92] The majority line was followed shortly afterwards by the Court in a unanimous judgment.[93]

The Court's line has been criticized for failing to adopt a view of 'data' in accordance with the thrust of the preparatory works and, in effect, with what is

[88] The CJEU being an especially notable instance: see further eg Paul Craig and Gráinne de Búrca, *EU Law* (5th edn, OUP 2011) 64 and references cited therein.

[89] Stefan Vogenauer, 'A Retreat from *Pepper v Hart*? A Reply to Lord Steyne' (2005) 25 OJLS 629.

[90] Norwegian Supreme Court decision of 22 December 1994, reported in Norsk Retstidende (Norwegian Law Reports; abbreviated as 'Rt') 1994, 1610.

[91] ibid 1612–13.

[92] ibid.

[93] Rt 1995 35. In this judgment, the Court simply followed the line of the majority in Rt 1994 1610 without further elaboration.

common in informatics.[94] While the criticism has merit, it is weakened by lack of detailed guidance in the preparatory works on the intended ambit of the term in question. One must also take into account that the object of interpretation was a penal provision and that application of criminal sanctions requires clear legal authority.[95]

However, the legal propriety of the judges' view of 'data' is not of primary concern here. More important is that their view highlights that we cannot reasonably expect uniformity of views about the meaning of 'data'. The regulatory consequences of the judges' line are also important. In response to the case, provisions were added to the Penal Code dealing specifically with unauthorized access to 'television and radio signals' (§ 262(4)(a)) and 'services provided via electronic telecommunication' (section 262(4)(b)). These provisions are much less technology-neutral than section 145(2). They are, in other words, an example of generic dreams giving way to definitional daylight in a technology-specific form. They illustrate that movement from technology-specific to technology-neutral rules is not purely unidirectional: when the latter rules are not provided with requisite clarity, they can be replaced or supplemented by more precisely drawn norms. While such a result is praiseworthy for enhancing legal certainty, the expense of the legislative amendment might have been avoided had legislators first taken proper account of the ambiguity of the terminology they initially employed.

7. *Complexity of Information Concepts*

A. *Semantics of Information*

One of our most prominent contemporary information theorists, Luciano Floridi, laments that '[i]nformation is a conceptual labyrinth'[96] due to its multidimensional character. The concept denotes a *process* (the action of being informed—also often called communication), *result* (that which is imparted—also typically called knowledge, at least if it leads to a change or confirmation of beliefs) or *object* (that which is informative).[97] And in terms of the latter, consensus as to which physical objects constitute information is rough and shifting—as the above case studies show.

Compounding this multidimensionality is that information also has special meanings in particular contexts. Law provides one such example: in common law jurisdictions, 'information' has been used to denote a particular type of

[94] The strongest criticism is expressed by Knut S Selmer in his short article, 'Hva er "data"?' (1995) Lov og Rett 149.

[95] See too Gunnar Aasland, 'Noen betraktninger om rettskildespørsmål i Høyesteretts praksis' (2000) Jussens Venner 157, 164.

[96] Luciano Floridi, *Information: A Very Short Introduction* (OUP 2010) 19.

[97] Michael Buckland, *Information and Information Systems* (Praeger 1991) 3–4, 39–40.

charge or complaint presented to a court.[98] Another example is the mathematical theory of communication developed by Claude Shannon and others.[99] This employs the term 'information' to denote a quantified degree of choice regarding the selection of one symbol from a set of possible symbols within a given communication channel, the semantic value of what is communicated being of secondary relevance.[100]

Conveyance of meaning, however, usually figures prominently in common usage of the information concept. It is apposite to recall the above-noted line taken by the NZ Court of Appeal: '[i]nformation in its ordinary dictionary meaning is that which informs, instructs, tells or makes aware'. Etymologically, 'information' stems at least in part from the Latin verb 'informare', which roughly means 'give form to' (an idea), 'instruct', or 'teach'.[101] As such, conveyance of meaning is integral to information as both a process and a result. The centrality of semantic value is also increasingly stressed in informatics, where a 'General Definition of Information' (GDI) in terms of 'data + meaning' has become an 'operational standard'.[102] This is reflected in the definition of 'information' given by the International Organization for Standardization (ISO): 'the meaning assigned to data by means of conventions applied to that data'.[103]

Information usually denotes a form of semantic content in law and attendant discourse too. Law is primarily concerned with regulating human relations; therein, the production and exchange of meaning play a key role.[104] Even when law deals specifically with signal transmission, the rules tend to serve aims (such as pluralism and freedom of expression) that ultimately revolve around the semantic significance of what is transmitted.[105]

Focus on meaning pushes context and the relational nature of information to the fore: the degree to which information is generated from data or other information, along with the type of information generated, depends on the interplay of many contextual factors. This is also recognized in legal rules, perhaps most explicitly in those laying down criteria for assessing when 'data'

[98] See eg Canada's Criminal Code 1985, s 576(2) ('No criminal information shall be laid or granted . . .').
[99] See especially CE Shannon, 'A Mathematical Theory of Communication' (1948) 27 Bell System Technical J 379, 623.
[100] RVL Hartley, 'Transmission of Information' (1928) 7 Bell System Technical J 535, 536.
[101] 'information, n' (*OED Online*, OUP March 2014) <www.oed.com/viewdictionaryentry/Entry/95568> accessed 17 April 2014.
[102] Luciano Floridi, *The Philosophy of Information* (OUP 2011) 83; Floridi (n 96) 20.
[103] ISO 2382-1, *Information Technology—Vocabulary—Part 1: Fundamental Terms* (1993).
[104] See too Cahir (n 2) s 4; Mayer-Schönberger (n 2) 16–17.
[105] See eg Directive 2002/21/EC on a common regulatory framework for electronic communications networks and services [2002] OJ L108/33, preamble recital 5 (noting that while '[i]t is necessary to separate the regulation of transmission from the regulation of content', such separation 'does not prejudice the taking into account of the links existing between them, in particular in order to guarantee media pluralism, cultural diversity and consumer protection').

or 'information' becomes 'personal'.[106] Numerous instances of more implicit legal recognition exist too—eg in judicially developed rules for determining when a course of negotiations becomes a legally binding contract and what the terms of such agreement are.[107]

B. *Semantics of Data*

Like 'information', the term 'data' is pregnant with confusing definitional variation.[108] In everyday discourse in English, 'data' is often used as a synonym for 'information' generally. Judges, legislators, and legal scholars also frequently conflate the two terms—as we see, for instance, in data protection law. However, 'data' is also often used to refer principally to a particular class of information; namely facts, quantities, or conditions derived from systematic observation or experimentation.[109] This use of the term is typical for the social and biological sciences. The equation of 'data' with 'facts' occurs occasionally in law too.[110]

In informatics, however, 'data' is usually seen as a vehicle of information rather than as information *per se*. More specifically, 'data' typically denotes signs, patterns, characters, or symbols which potentially represent some object or process and, through this representation, can communicate information about that object or process.[111] This way of defining 'data' is in keeping with its etymology, 'data' being the plural form of the Latin word 'datum', which roughly means 'gift' or 'something given'. In the terminology of semiotics, data has a primarily sigmatic (representational) function, while information has primarily semantic and cognitive elements. Syntax plays a role too: in order to generate a particular type of information, data must be formed and structured in a particular manner. Integrated, therefore, in the above-mentioned 'General Definition of Information' (GDI) as explicated by Floridi is the qualification that data is 'well formed'.[112] This qualification is also reflected in the definition of 'data' given by the ISO: 'a representation of facts, concepts or instructions in a formalized manner suitable for communication, interpretation or processing

[106] See eg Data Protection Directive art 2(a) ('"personal data" shall mean any information relating to an . . . identifiable natural person . . . ; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity').

[107] See eg *Gibson v Manchester City Council* [1979] UKHL 6, [1979] 1 WLR 294; *Butler Machine Tool Co Ltd v Ex-Cell-O Corp (England) Ltd* [1977] EWCA Civ 9, [1979] 1 WLR 401.

[108] For an overview, see Machlup (n 12) 646–49.

[109] 'datum, n' *OED Online*, OUP March 2014) <www.oed.com/viewdictionaryentry/Entry/47434> accessed 17 April 2014.

[110] Thus, in a US Supreme Court case dealing with the scope of copyright under US law, 'raw data' has been described as 'wholly factual information not accompanied by any original written expression': *Feist Publications v Rural Telephone Service Co* (1991) 499 US 340, 345 (O'Connor J).

[111] See eg Floridi (n 96) 20–21; Fourman (n 4) 241; Peter Rob and Carlos Coronel, *Database Systems: Design, Implementation, and Management* (8th edn, Thomson 2009) 5–6; Chrisanthi Avgerou and Tony Cornford, *Developing Information Systems: Concepts, Issues and Practice* (2nd edn, Macmillan 1998) 115.

[112] Floridi (n 102) 84; Floridi (n 96) 20–21.

by human beings or by automatic means'.[113] Thus, some sort of formalized language (with agreed syntax and semantics) will typically form the bridge between data and information. Also implicit is that the conversion of data to information involves an increase in epistemic value. This increase is often expressed in terms of a broader 'DIKW hierarchy' (Data-Information-Knowledge-Wisdom hierarchy), which embodies an assumed linear rise in epistemic value from D (low value) through to W (high value), each stage being reached on the basis of ever-higher degrees of reflection, organization, and accumulated learning.[114]

Although 'data' and 'information' tend to be differentiated epistemically and functionally in informatics, they are also viewed as inextricably interconnected. Data, at least according to the GDI as explicated by Floridi, is a necessary constituent of information. In other words, 'information cannot be dataless'.[115] This helps bolster the commonplace conflation of the two in everyday discourse.

C. *Semantics of Communication*

The term 'communication' does not seem to present the same degree of definitional instability as 'data' and 'information', although it is inevitably afflicted by some of the ambiguity inherent in the latter. Communication is broadly understood as denoting the passing on, sharing, or exchange of information. Thus conceived, communication is more than mere transmission, but connotes an informative (information-giving) process—ie impartation. This is in keeping with its etymological roots in the Latin term 'communico', which means 'to share'. As Cherry notes, communication 'is essentially a social affair'.[116] The notion of communication as impartation, though, is not always clearly reflected in the legal use of the term. Laws sometimes employ the term to denote the mere transmission of data.[117] The term is also applied to cover transmission processes between non-human nodes.[118] Surprisingly, the term is even occasionally defined to cover situations that apparently do not involve any data transmission at all.[119]

As indicated earlier in the article, laws frequently regulate particular forms of communication rather than communication generally. In doing so, they tend to

[113] ISO (n 103).

[114] Jennifer Rowley, 'The Wisdom Hierarchy: Representations of the DIKW hierarchy' (2007) 33 J Information Science 163 and references cited therein.

[115] Floridi (n 102) 84; Floridi (n 96) 17.

[116] Colin Cherry, *On Human Communication: A Review, a Survey, and a Criticism* (3rd edn, MIT Press 1978) 4, 306.

[117] See eg Norwegian Electronic Communications Act of 2003, § 1-5(1).

[118] See eg UK Regulation of Investigatory Powers Act 2000, s 81(1) (defining 'communication' as including 'signals serving either for the impartation of anything between persons, between a person and a thing or between things or for the actuation or control of any apparatus').

[119] See UN Electronic Communications Convention, art 4(a) (defining 'communication' in terms of 'any statement, declaration . . .' and the like without specifying that transmission of the statement etc must occur).

use terms other than 'communication' to denote the specific communicative processes they are intended to regulate. Such terms can also give rise to definitional problems, particularly as they have usually been linked to specific service modes and technological platforms. Examples are 'telephony', 'telecommunication' and 'broadcasting'. Convergence has called into question the workability of the legal conceptual frameworks employed to distinguish them from each other. The concept of 'broadcasting', for instance, has been successively tested by new or converging media platforms,[120] and it has typically been redefined using more generic terminology[121] or more refined criteria.[122] The success of some of these attempts at redefinition is, at the very least, debatable.[123]

D. *Data as Artefact and Abstraction*

A question that is definitively decisive for the current application of law is whether data must be created, directly or indirectly, by human beings. Also decisive is the closely related question of whether data must be a formalized representation of objects or processes in the real world. Both questions concern not just the nature of data and, thereby, information, but also the relationship between data and information on the one hand and their media on the other. How these questions are answered helps to determine the degree to which laws that specifically refer to 'data' may apply to phenomena in the natural or biological world.

It is pertinent to consider these questions, taking as a point of departure the characterization of data and information advanced in the GDI. For the purposes of the GDI, data is regarded in its most basic form as an absence of uniformity.[124] Only through such difference is information generated. Differentiation may occur at various levels: Floridi distinguishes between lack of uniformity in the real world (*diaphora de re*—eg contrast in colours), differences between signals or physical states (*diaphora de signo*—eg electrical charges) and differences between symbols (*diaphora de dicto*—eg letters in an alphabet).[125] If all of these situations are viewed as involving types of data, it is

[120] For early analysis on this point, see Nils Kristian Einstabland, *Kringkastingsbegrepet* (TANO 1992).

[121] See eg South Africa's Electronic Communications Act 2005, s 1 (defining 'broadcasting' as 'any form of unidirectional electronic communications intended for reception by (a) the public; (b) sections of the public; or (c) subscribers to any broadcasting service, whether conveyed by means of radio frequency spectrum or any electronic communications network or any combination thereof').

[122] See eg Audiovisual Media Services Directive 2007/65/EC amending Council Directive 89/552/EEC on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the pursuit of television broadcasting activities [2007] OJ L332/27, art 1(2) (introducing a distinction between 'linear' and 'non-linear' audiovisual media services, and limiting 'television broadcasting' to the 'linear' category).

[123] See eg Neal Geach, 'Convergent Regulation for Convergent Media: An Overview of the Audiovisual Services Media Directive' (2008) 1 J Information, Law and Technology (now EJLT) paras 2.1–2.1.1 <www2.warwick.ac.uk/fac/soc/law/elj/jilt/2008_1/geach> accessed 17 April 2014.

[124] Floridi (n 102) 85–86; Floridi (n 96) 23–24. In other words, data (like information) is relational.

[125] ibid.

not difficult to see that data in terms of the GDI may take numerous and sometimes surprising forms, not just as letters on a piece of paper, 1s and 0s in a computer program, or smoke signals, but also, presumably, as a horse's head,[126] at least in certain contexts. Particularly important is that data, in the view of Floridi and some others, may also consist of a lack of uniformity in the natural environment (eg concentric rings in a felled tree trunk).[127]

A corollary of this perspective is that data may consist of differences in and between biological materials. The perspective also has implications regarding when and how data arises. In particular, it has a bearing on when data is created, as opposed to collected or obtained. As shown in the above case study on database protection, that issue is germane to the application of the Database Directive and the attendant national rules.

Such a broad view of data challenges, if not breaks with, more everyday assumptions about what the concept properly denotes. This is especially so with the example of data *in natura*. A characteristic often ascribed to data—and implicit in the above-cited definition of data given by the ISO—is that it is more or less consciously shaped with the intention of communicating or otherwise processing information. In other words, data is essentially artificial. As such, it cannot exist *in natura*.[128]

Nonetheless, we cannot simply dismiss the broader conception of data advanced by Floridi as belonging to the margins of the informatics community. Floridi's work commands considerable respect and influence.[129] Moreover, his assumption that data is a necessary constituent of information accords with the apparent fact that naturally occurring phenomena do provide us with information. They do this based on our sensory perception, filtered by and combined with previously acquired information, which has, in turn, been organized as knowledge. It is not clear whether this factor constitutes a logical flaw in the ISO's approach, as it is not clear if the ISO's definition of 'information' (or, for that matter, 'data') purports to cover all types of information (or data): if it does, there is a flaw.

Even if one accepts the ISO's view of data, the conceptual barriers to treating biological material as data are not as profound as they appear. If separated from its natural environment and structured as a sample or set of samples with the intention of providing information, such material can function at the very least analogously to data in the above-described sense. This will typically be the case with biobanks.

[126] Recall the episode in Mario Puzo's novel, *The Godfather* (1969)—and in the 1972 film of the same name directed by Francis Ford Coppola—when the severed head of a race horse is placed in the bedroom of the horse's owner, thus signalling that his life is threatened if he does not accede to a particular request of the mafia.

[127] Floridi (n 96) 32.

[128] A line advanced also in Jon Bing, 'Informasjon i informasjonsretten' (2009) 46 Ugeskrift for Retsvæsen 383, 384.

[129] Evidenced in part by the International Association for Computing and Philosophy awarding Floridi with the Covey Award for Outstanding Research in Computing and Philosophy, in 2012.

We see such a view in the UK litigation of the *Marper* case, particularly in the judgment of Baroness Hale in the then House of Lords (now Supreme Court) who stated that 'the only reason [the DNA samples] . . . are taken or kept is for the information which they contain', that the samples therefore 'are kept as information . . . and nothing else', and that '[t]he same privacy principles should apply to' the samples as to the fingerprints and profiles.[130] Academic commentary on particular national data protection laws also states that biological material may become personal data pursuant to the legislation concerned, from the time that the material is isolated and structured for the purpose of generating information that facilitates identification of individuals.[131] I see considerable sense in such a line, both conceptually and as a matter of principle,[132] even if—as pointed out at the close of the first case study—there may be difficulties in its legal operationalization.

E. *The Role of Media*

We must also not forget the media in which data inheres. As the carrier or supplier of potential information, data has a physical existence, even though this can often be relatively invisible (eg as electronic impulses). Data is always incorporated in, or attached to physical material, which thereby functions as data media. Paper, magnetic tape, plastic, and silicon are typical examples. Yet, biological material can also function as a data medium. Data may thus have an embodiment with significant natural elements, even if data putatively cannot exist *in natura*.

Moreover, data is often so tightly connected or converged with its carrier that it cannot be easily separated from it physically. Examples are ink on paper, runes and hieroglyphics on stone, and tattoos on skin. And some forms of data are inseparable from their media because the latter are the data, at least if we accept Floridi's explication of data. The horse's head mentioned above is an example in point. In these sorts of cases, we may ask where the logical boundary is between the data and the physical material.

The pertinence of that question is compounded by the scientific and technological developments outlined in the introduction. With biological material increasingly mined for, and treated as information, justifying a distinction between the former and the latter—ie between medium and message—becomes more difficult.

[130] *R (S and Marper) v South Yorkshire Police (Consolidated Appeals)* [2004] UKHL 39, [2004] 1 WLR 2196 [70]. For a similar line, see ALRC and AHEC (n 37) [8.32].

[131] The case with the German Federal Data Protection Act, according to the most authoritative academic commentary on it: see Ulrich Dammann, '§3 Weitere Begriffsbestimmungen' in Spiros Simitis (ed), *Bundesdatenschutzgesetz* (7th edn, Nomos 2011) 297, 302–3.

[132] See too Taylor (n 33) 218–19 and ch 7.

8. *Conclusions*

The analysis highlights the complexity of basic information concepts along with the controversy and confusion accompanying their meanings. Hence, the aspirations and strategies—the 'generic dreams'—promoting the use of such concepts as front-line legal terms need to be accompanied by definitional daylight in the form of appropriate guidance from lawmakers. Arguably, the most pressing dilemma that such guidance must address concerns the extent to which law should clearly distinguish between data and information on the one hand and biological material and processes on the other.

The article thus provides a corrective to single-minded pursuit of the regulatory principles of technology neutrality and, to a lesser extent, functional equivalence. It indicates that the modernization of the legal conceptual apparatus of the 'paper age' has occurred with too little consideration of the ambiguity of the new generic nomenclature in light of recent biotechnological advances. In a sense, the modernization has not been sufficiently modern. Hence, we need a recalibration of both principles so that laws intended for the informational world do not haphazardly enter the biological world. In the terminology of Bennett Moses, we need greater 'targeting' of these laws.[133]

I am not suggesting that these laws ought never to apply to biological material or processes. Rather, their application should occur only after systematic assessment of their suitability for the biological world. Neither do I suggest that generic dreams are intrinsically in tension with definitional daylight. Rather, the basic aim of such dreams—to ensure the sensible application of law in the face of technological change—can usually be met *only* if definitional issues receive thorough attention. Concomitantly, we need more rigorous analysis of the information concepts that populate generic dreams.

The definitional daylight I am calling for is a far cry from the overly precise 'tick-box' style of legislation criticized by Reed.[134] The daylight need not translate into extensive or cumbersome definitional provisions, but could simply clarify the limits of information concepts—addressing, in particular, whether or not 'information' or 'data' encompasses biological material or other tangible objects. Clarification need not occur just in legislation; it could occur in preparatory works too (assuming, obviously, that these are regarded as proper sources of statutory interpretation). As intimated above, it would be unrealistic to expect definitions to address every ambiguity. Law must inevitably live with a measure of uncertainty.[135]

The definitions typically applied within informatics provide useful points of departure for defining information concepts in law. While agreement within

[133] Lyria Bennett Moses, 'Recurring Dilemmas: The Law's Race to Keep Up with Technological Change' (2007) 7 U Illinois J Law, Technology and Policy 239, 259.
[134] Chris Reed, 'How to Make Bad Law: Lessons from Cyberspace' (2010) 73 MLR 903.
[135] See too eg Timothy AO Endicott, *Vagueness in Law* (OUP 2000) esp 190; Bennett Moses (n 132) 250ff.

informatics over the proper definition of these concepts is incomplete, it provides sufficiently firm ground on which to launch the definitional efforts of jurists and other policy entrepreneurs. Informatics teases out the numerous dimensions of information and related concepts, providing lawmakers with a menu of issues, questions, and options that can help structure their own use of these concepts—as this article itself illustrates.

This does not mean that lawmakers ought slavishly to adopt the views of information scientists, nor that *one* set of definitions ought to apply across the entire legal board. Applying a single set of definitions to cover all regulatory contexts is unrealistic and undesirable. We must ensure that the plea for conceptual rigour does not end in conceptual rigor mortis. Risk of the latter is particularly acute with a polysemantic concept such as information. Thus, allowances must be made for divergence from the advocated definitional point of departure. Yet, divergence should only occur if there are good reasons for it and lawmakers themselves are aware of and communicate those reasons.

In making their decisions on these and attendant issues, lawmakers will—and ought to—be driven partly by pragmatic concerns. One such concern is the optimal division of tasks between different legal codes (eg is biobank regulation best carried out under the auspices of data protection law?). Another concern is the optimal division of tasks between the legislature and judiciary in generating legal certainty and regulatory connection. This carries over to consideration of the propriety of providing such certainty and connection *ex ante* (primarily through legislation) or *ex post* (primarily through adjudication). The precise balance to be struck between *ex ante* and *ex post* means is difficult to determine, particularly as they have mutually related effects and are affected by other factors (such as litigation costs and propensity to risk-aversion).[136]

The thrust of my argument favours *ex ante* legislative provision of certainty for the legal-regulatory issues that are the focus of this article. It does so for the reasons given at the beginning of this final section. I fear too that the scale of the biotechnological developments outlined in the introduction is too large and too much a 'game-changer' to be handled simply through ad hoc court litigation. The developments require systematic attention from legislatures with an eye for the 'big picture' rather than the usually more narrowly focused perspectives of litigants and judges.[137]

In any case, it is vital that jurists and other policy entrepreneurs do not treat information concepts in a cavalier fashion, but that they give due respect to conceptual and terminological integrity. The legal meaning attributed to such

[136] See further eg Giuseppe Dari-Mattiaci and Bruno Deffains, 'Uncertainty of Law and the Legal Process' (2007) 163 J Institutional and Theoretical Economics 627, 647–49; Bennett Moses (n 17).

[137] See too Brownsword's reservations about relying on 'purposive interpretation' by the judiciary in the face of 'productive' regulatory disconnection: Brownsword (n 18) ch 6.

concepts is a significant factor in establishing solid terminological bases for regulation. Indeed, the process of understanding how best to apply such concepts in law ought to be seen as part of a larger endeavour to develop an appropriate 'philosophy of information' (Floridi)[138] in keeping with the age in which we live.

[138] Floridi, *Information* (n 96) 8.