"Data Spaces": Data Structures as a Question of Law

KAI VON LEWINSKI[1]

Big data and large databases are resources of power, both in the political and the economic context. The law has so far responded to this new challenge almost exclusively with the means and instruments of data protection law. However, (traditional and current) data protection law has a specific micro-perspective since it primarily regulates individual data processing in relation to a single "data subject". Structural aspects, on the other hand, are almost not addressed at all by current data protection law. It is remarkable to what detail the processing of personal data relating to the gender of a person is regulated, but there are no general legal requirements for the (non-)representation of gender attributes (male, female, or any other) in databases—neither for individuals, nor for groups, and not for society as such.

Data structures are a social power resource although it is not in the research focus yet. This is not only true for databases, but also for taxonomies and formats on which they are based. Those who have the power of definition over data structures decide on whether and how the content is represented in such data structures; and no legal regulatory regime exists so far for this informational power yet. There are no rights to data structures, and there are hardly any obligations regarding the structuring of data, at least not explicitly. Informatic system analysis shows that by the means of structure, type and format decisions with regard to data, "data spaces" may be created and established. And these "data spaces" represent social reality in a specific way, and at the same time "data spaces" do have an impact on social reality. "Data spaces" are means of power which limit the possibilities of others. This demands for legal containment.

It should be the aim of research and politics to address the lack of a legal description of the power emanating from data spaces and to develop a new approach to their legal containment: How can power emanating from data structures be translated into legal categories (beyond the insofar unsuitable data protection law)? What legal instruments exist with regard to this power constellation? Which legal instruments are missing, and what could they look like?

[1] Prof. Dr., University of Passau, Researcher at the HIIG 2013/14.—This text has very much gained from the discussions with *Jörg Pohle* (HIIG) whom I have to thank very much for his computer science ("Informatik") perspective and input.

SOCIAL SIGNIFICANCE OF DATA STRUCTURES, TYPES AND FORMATS

The question of the structure of data and represented data classes is not only a question of theoretical interest. Formal or other restrictions, for example on database entries, can lead to such results that personal names are not mapped according to social reality, at least not according to the identity of the person concerned; more complicated personal names do not easily fit into data fields, especially if foreign letters (umlauts and others) and spellings have to be transcribed. Fiscal budgets and public registers represent only a limited cut-out of reality with a more or less narrow perspective. The same applies to statistics in general.

In addition, the layout of a database can have an effect beyond its original purpose. Every operationally used database forms a "data space" (here not used in its mathematical or informatic meaning) which then interacts with the social reality: It is the purpose of every set of data (database) that it connects data to the person or entity it represents. Such, a representation of features means controllability and the possibility of linking, whereas a non-representation results in non-regulation and the impossibility of linking—or even to the representation of non-existence. An abrogation of the personal status "gender" would, for example, make it impossible to link individuals or groups to their gender; ultimately, this would mean in practice that gender equality programs would no longer be possible, because one cannot distinguish men from women anymore. Another example has been (in Germany until recently) the restriction of data base entries in public resident registers to the binary gender order (male/female) which left existing third genders unrepresented and unrepresentable and consequently administratively non-existing.

Inclusion and exclusion of attributes in data structures have a social effect as well as an individual one. Anyone who can and may decide on the representation of sections of social reality in the logical space of data structures also influences the associated representation of informatic systems in (and into) society. For example, non-existent categories in (sectorial) planning mean that these aspects cannot be taken into account, whereas existing categories produce their (own) relevance.

Whoever determines such a space of possibility, whoever has the power to define storable and then reflected reality, controls a resource of power. In this respect, "data spaces" are also spaces of domination. It is the task of jurisprudence and law to describe and, if necessary, to limit them.

LEGAL BLIND SPOTS AND REGULATORY GAPS

Today's data protection law is narrowed to a concrete close-up perspective and is therefore blind regarding structural concentrations of power. The early data protec-

tion discourse[2], which was strongly linked to the legal informatics of the time in terms of personnel and methods, certainly had had this perspective. The later data protection jurisprudence lost this structural perspective—until today. The (German) Federal Data Protection Acts (Bundesdatenschutzgesetz, BDSG) since 1977 and the EU General Data Protection Regulation (GDPR) focus almost exclusively on the individual. Register law discusses data structures purely descriptively as well[3]. And the findings in computer science, which certainly analyses the social potential of data structures[4], are not translated into legal solution categories.

The existing (European) database law (Database Directive 96/9/EC) refers only to the (intellectual property) right to a database as a whole[5]. Consequently, database law is structurally blind in the sense that it does not address the specific structure of databases, but only a specifically structured data stock. Register law contains only very specific regulations which do not take into account the aspect of concentration of power through structural and format decisions. And data protection law has—as already mentioned—a narrowed and individualistic perspective on data processing and is in particular procedural but does neither address nor even know the category of data structures, or data pools, or data power ("Datenmacht").

In up-and-coming areas like data law and algorithms law as well as in the field of big data, two main approaches are discussed to limit informational and data power: on the one hand, the regulation of algorithms as such and, on the other hand, the limitation of the processing of the underlying data[6]. The former is primarily dealt with in the context of protection against discrimination and in duties to algorithm transparency, the latter is discussed in connection with data protection law (esp. data minimization) or intellectual property figures (so-called "data ownership"). In addition to the regulation of computation itself (algorithm regulation in a narrower

[2] In particular *Steinmüller*, EDV und Recht, 1970; *Simitis*, Informationskrise des Rechts und Datenverarbeitung, 1970; *Kerkau*, Automatische Datenverarbeitung (ADV) – Kybernetik in Rechtswissenschaft und Praxis, 1970; *Dammann*, Datenbanken und Datenschutz, 1974.

[3] For Germany cf. *Krafka*, Einführung in das Registerrecht, 2nd ed. 2008.

[4] Cf. *Stachowiak*, Allgemeine Modelltheorie, 1973; *Stachowiak* (ed.), Modelle – Konstruktion der Wirklichkeit, 1983; *Mahr*, Die Informatik und die Logik der Modelle, Informatik Spektrum 2009, pp. 228–249; *Desrosières*, Die Politik der großen Zahlen, 2005, esp. chap. 8; most recently *Guagnin/Pohle*, Welt -> Modell -> Einschreibung -> Welt', fiffkon18, 2018 (https://media.ccc.de/v/fiffkon18-10-welt_-_modell_-_einschreibung_-_welt).

[5] Comprehensive for the law making process *Indranath Gupta*, Footprints of Feist in European Database Directive: A Legal Analysis of IP Law-making in Europe, 2017; from a German perspective *Conrad/Grützmacher* (eds.), Recht der Daten und Datenbanken im Unternehmen (= Festschrift J. Schneider), 2014.

[6] Cf. the research project „Algorithmenkontrolle als Regulierungsaufgabe" at the University of Speyer by *Mario Marini* et al. (http://www.foev-speyer.de/en/research/digitization/data-driven-performance-of-public-sector-tasks/algorithm-control-as-regulatory-task.php?p_id=1904).

sense) and the regulation of data as such (conventional data protection; intellectual property law), the structure of data bases must also become a subject of regulative considerations. This is a third legal dimension in the field of data law, which has not yet been examined in detail.

This clearly shows a blind spot in legal science and legal informatics. An analysis of the existing legal regulations shows that so far no regulation addresses the problem of data structures as a power resource as such, let alone interactions between IT system and social reality.

PERSPECTIVES AND FIELDS OF FURTHER RESEARCH

For a better social understanding of information society, legislation and courts should become aware of relevant constellations of inclusion and exclusion in "data spaces", be it the name which is now only computer-compatible but incorrectly written, be it the consideration of a category relevant to planning but not provided for in planning law, be it a third gender.

A first research step would be a comprehensive collection of relevant regulations and regulatory perspectives. Since—as mentioned above—data protection law and intellectual property law do not offer the full answer, register law could turn out to be a productive source which has received only little scientific attention yet. Another important field of reference will be tax law, rules for the public budget and accounting law as well as planning law.

Additionally, the early years of legal informatics (esp. in Germany, the so-called "Rechtsinformatik" of the 1970s), should be re-visited and re-read. Researchers in these years had described the interaction and retroactivity of informatic and social systems with the means of the (meanwhile forgotten) system-theoretical "classical" legal informatics[7] (initially also known as "legal cybernetics" ["Rechtskybernetik"])[8]. This approach might add the perspective, that data structures and taxonomies form "data spaces" in which and through which power can be exercised and which constitute a means of control.

If it turns out to be true that jurisprudence has a blind spot regarding "data spaces" and data structures, it has to be filled with light. Such light might come from

[7] E.g. *Podlech*, Information – Modell – Abbildung – Eine Skizze, in: Steinmüller (ed.), Informationsrecht und Informationspolitik, 1976, pp. 21–24; *Harbordt*, Computersimulation in den Sozialwissenschaften 1 – Einführung und Anleitung, 1974; *Dammann*, Datenbanken und Datenschutz, 1974; retrospectively *Heibey*, Zu den Anfängen der informatischen Wirkungsforschung: Die Theorie der Informationsveränderungen, in: Garstka/Coy (eds.), Wovon – für wen – wozu. Systemdenken wider die Diktatur der Daten (= Gedächtnisschrift Steinmüller), 2014, pp. 131–144; comprehensively to the early years of the "Rechtsinformatik" *Gräwe*, Die Entstehung der Rechtsinformatik, 2011.

[8] Cf. *Simitis*, Rechtliche Anwendungsmöglichkeiten kybernetischer Systeme, 1966; *Suhr*, Zur Einführung: Recht und Kybernetik, JuS [Juristische Schulung] 1968, pp. 351–353; *Haft*, Nutzanwendungen kybernetischer Systeme im Recht (Diss. jur. Gießen) 1968; *Podlech*, Rechtskybernetik – Eine juristische Disziplin der Zukunft, in: Erdsiek (ed.), Juristen-Jahrbuch 10 (1969/1970), pp. 157 et seq. – The term „Rechtskybernetik" came out of use soon.

constitutional law. An innovative idea might be whether an unwritten fundamental right to identity[9] would include a right to self-determined representation in "data spaces". Other (and more conventional) constitutional law categories should be considered, adapted and adopted as well, such as the rule of law. These fundamental values can then be reflected and mirrored back on the existing legal provisions to outline applicable solutions for the legal practice.

To put it into a nutshell: The outlined new legal approach tries to understand structures of "data spaces" and the format of data sets not merely as immaterial property but as an informational resource of social power.

[9] *Kieck*, Der Schutz individueller Identität als verfassungsrechtliche Aufgabe, 2019.