Technology and Regulation

Not Hardcoding but Softcoding Data Protection

Aurelia Tamò-Larrieux*, Simon Mayer** & Zaira Zihlmann***

GDPR, privacy by design, techno-regulation, privacy engineering, softcode

aurelia.tamo@unisg.ch

simon.mayer@unisg.ch

zaira.zihlmann@unilu.ch

The delegation of decisions to machines has revived the debate on whether and how technology should and can embed fundamental legal values. In this article, we discuss the translational, system-related, and moral issues raised by implementing legal principles in software. While our findings focus on data protection law, they apply to the interlinking of code and law across legal domains. These issues point towards the need to rethink our current approach to design-oriented regulation and to prefer 'soft' implementations, where decision parameters are decoupled from program code and can be inspected and modified by users, over the 'hard' embedding of such parameters into opaque pieces of program code.

1. Introduction

With more smart devices guiding us through our daily activities comes the quest to ensure that these technologies reflect the fundamental values of the society they are embedded in. Smart products like social robots can sense their environment, weigh various options against each other, and act upon their decision-making.[1] The key question thus becomes how options within the decision-making process are balanced and whether those decisions can take the legal environment into account.

The automatic adaptation of code to the legal parameters set out in law raises fundamental questions. A rich literature on techno-regulation and hardcoding or hardwiring data privacy exists, upon which this article builds.[2] Whether the encoding of law appears as part of the 'solution space'[3] or part of a problem, depends also on what legal field one is analyzing (e.g., Intellectual Property rights and Digital Rights Management systems, privacy by design).[4] What is clear is that law in writing vs. law in code can have very different properties, i.e., act differently upon society, thereby raising *systemic* and *moral* issues.

While interdisciplinary research groups have been active in addressing *translational* challenges of interlinking code and law,[5] philosophers and legal scholars have debated the merits and limitations of such initiatives. Seminal research has been conducted among others by Ronald Leenes, who has disentangled techno-regulatory initiatives originating from state and non-state regulators;[6] Mireille Hildebrandt, who has coined the term 'Ambient Law' which more broadly strives to integrate legal protection into the design of technology;[7] Karen Yeung, who analyzes the different effects of legal prohibition vs. techno-regulation on moral agency suggesting that the partial erosion of moral

1 George A. Bekey, 'Current Trends in Robotics: Technology and Ethics' in Patrick Lin, Keith Abney and George A. Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT Press 2012) 17.

2 Lee Bygrave, 'Hardwiring Privacy' in Roger Brownsword, Eloise Scotford and Karen Yeung (eds), *The Oxford Handbook of Law, Regulation and Technology* (OUP 2017) 754; Bert-Jaap Koops and Ronald Leenes, 'Privacy regulation cannot be hardcoded. A critical comment on the 'privacy by design' provision in data-protection law' (2014) 28(2) *International Review of Law, Computers & Technology* 159; Ugo Pagallo, 'On the Principle of Privacy by Design and Its Limits: Technology, Ethics and the Rule of Law' in Serge Gutwirth and others (eds), *European Data Protection: In Good Health?* (Springer 2012) 343; Karen Yeung, 'Can We Employ Design-Based Regulation While Avoiding Brave New World?' (2011) 3(1) *Law, Innovation and Technology* 1.

3 Urs Gasser, 'Recoding Privacy Law: Reflections on the Future Relationship Among Law, Technology, and Privacy' (2016) 130(2) *Harvard Law Review Forum – Law, Privacy & Technology Commentary Series*.

4 Bygrave, 'Hardwiring Privacy' (n 2), 755.

5 Cf. e.g., Ronald Leenes and others, 'ENDORSE. Deliverable D2.5 Legal Requirements' (2011) https://cordis.europa.eu/docs/projects/cnect/3/257063/080/deliverables/001-ENDORSED25submitted.pdf (accessed 29 October 2020); Stefan Schiffner and others, 'Towards a Roadmap for Privacy Technologies and the General Data Protection Regulation: A transatlantic initiative' in *Proceedings of the Annual Privacy Forum 2018* (Barcelona, Spain, June 2018) https://people.cs.kuleuven.be/~bettina.berendt/Papers/schiffner_et_al_APF_2018.pdf (accessed 8 November 2020); Michael Birnhack, Eran Toch and Irit Hadar, 'Privacy mindset, technological mindset' (2014) 55(1) *Jurimetrics* 55.

6 Ronald Leenes, 'Framing Techno-Regulation: An Exploration of State and Non-State Regulation by Technology' (2011) 5(2) *Legisprudence* 143.

7 Mireille Hildebrandt, 'A Vision of Ambient Law' in Roger Brownsword and Karen Yeung (eds), *Regulating technologies: Legal futures, regulatory frames and technological fixes* (Hart Publishing 2008) 175; Mireille Hildebrandt, 'Legal Protection by Design: Objections and Refutations' (2011) 5(2) *Legisprudence* 223.

* Aurelia Tamò-Larrieux is an International Postdoctoral Fellow at the Law School of the University of St.Gallen.

** Simon Mayer is a Professor of Interaction- and Communication-based Systems at the Institute of Computer Science of the University of St.Gallen.

*** Zaira Zihlmann is a PhD Candidate at the Faculty of Law of the University of Lucerne.

Received 13 Aug 2020, Accepted 10 Apr 2021, Published: 15 May 2021.

freedom through technology does not have to result in overall collapse of moral foundations;[8] as well as Emre Bayamlıoglu and Ronald Leenes, who describe how data-driven decision-making that enacts regulatory orders undermines the rule of law.[9]

Guided by a concrete implementation of data protection principles in a smart product[10] and building upon literature on the failures of hard-coding privacy[11], we explore the pitfalls of bottom-up implementations of legal principles into software. This leads to a better understanding of why encoding data protection is an imperfect remedy. Sometimes, the imperfectness originates from the structure and behavior of law, sometimes from the structure and behavior of code. Our goal is to enable a differentiated discussion on those interactions in the specific field of data protection. The translational issues raised throughout the article lead to a call for action for both, the computer science and the legal community. Beyond these translational issues, we discuss systemic and moral challenges raised by design-based regulation. These challenges point to more fundamental questions on how and when we want law to be interlinked with code in a way that code regulates human and machine transactions. We argue that, to address those latter issues, we need to move towards 'softcoding' which decouples decision parameters (e.g., production rules, conditionals, thresholds) from opaque program code and thereby allows users to observe and adapt them. Softcoding does not only lead to advantages on the technology side, since it ensures that systems remain flexible to changes of the (legal) environment; it also entails that systems remain transparent, contestable, and malleable and thereby still allow for disobedience as well as control by users and judges.

This article contains three main sections. In Section 2, we start by describing the design implications of the GDPR with focus on the norm on data protection by design and default. From this overarching principle we move towards discussing hard and softcoding approaches to law as well as the technology implementations that have been proposed to comply with the principles of data protection law. This literature review situates the topic of this article into both its legal and technology contexts. Moving away from this dichotomy, Section 3 discusses why encoding data protection principles in practice is an imperfect remedy. On a meta-level, the imperfectness is grouped into *eight clusters of issues* that arise when taking a bottom-up approach to encoding data protection. Within each cluster, detailed specifications on why the interlinking of code and law does not lead to an isomorphic representation of the foundation of the law within code are discussed. Upon this basis, Section 4 describes a path forward: While in our opinion imperfectness does not equal failure nor suggests that we should abandon those approaches altogether, we emphasize the need for *more flexible, loosely coupled*, implementation approaches that allow for more transparency, contestability, and malleability. We furthermore emphasize the need for transdisciplinary experts who promote responsible technology that does not merely lead to superficial implementations of law in code but to one that preserves core tenets of our legal system. If, in the future, law becomes even more computable[12], then the need to establish clear procedural rules on how to contest hard- or softcoded provisions, ensure understandability of legally binding decisions will become key. Such challenges can only be addressed when moving beyond strictly disciplinary approaches.

2. From an Ideal to Implementations

2.1 "Yes, but..." and Other Design Implications of the GDPR

The quest to interlink law and code and create computable laws is seen in various legal fields such as in data protection law, which will be the focus of this article. As a regulation, the GDPR can be best described as a *compromise*. It is a compromise between different data protection regimes within the EU as well as a compromise between various interests that have shaped its final scope.[13] The compromise between different data protection regimes in the EU was already apparent within Directive 95/46/EC[14] (Directive), the predecessor of the GDPR. The Directive itself drew heavily from the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108)[15], which was originally signed in 1981 and later updated in 2018. Convention 108 was the initial push to a harmonized data protection approach in the EU.[16] Its main principles were incorporated and refined in the Directive and adopted within the GDPR. Convention 108, the Directive, and the GDPR all outline their 'objectives' and 'purpose' along the lines of wanting to ensure the protection of fundamental rights and freedoms of individuals with respect to their 'right to the protection of personal data'[17] and 'right to privacy.'[18] The objective of protecting fundamental rights is also what makes the application and, as will be shown, technical implementation of data protection law challenging. Fundamental rights in their core *require a balancing approach*, which from a technical perspective means that more often than not the solution will be not merely 'yes' or 'no' but a 'yes, but' or 'no, but' (i.e., its logic is defeasible). The *'yes, but'-principle* is inherent to the European data protection approach.[19]

The principles set in place within Article 5 of the GDPR set the basic

8 Yeung (n 2), 27.

9 Emre Bayamlıoglu and Ronald Leenes, 'The 'rule of law' implications of data-driven decision-making: a techno-regulatory perspective' (2018) 10(2) *Law, Innovation and Technology* 303 et seqq.

10 Kimberly Garcia and others, 'Towards Privacy-Friendly Smart Products' (2021) preprint available here https://www.alexandria.unisg.ch/262898/1/TechPaperToyRobot_Alexandria.pdf (accessed 5 April 2021). See Section 2.2 "Hard- or Softcoding Law" for further context.

11 Koops and Leenes (n 2), 159; Ronald Leenes and Federica Lucivero, 'Laws on Robots, Laws by Robots, Laws in Robots: Regulating Robot Behaviour by Design' (2014) 6 *Law, Innovation and Technology* 193.

12 We understand the term "computable" as used in social science literature as regulation processed by and through machines, while not referring to the theory of computation in computer science.

13 Cf. Ece Ö Atikcan and Adam W Chalmers, 'Choosing lobbying sides: the General Data Protection Regulation of the European Union' (2019) 39(4) *J Pub Pol* 543, 545; cf. also Jukka Ruohonen, 'David and Goliath: Privacy Lobbying in the European Union' (2019) https://arxiv.org/pdf/1906.01883 (accessed 28 October 2020).

14 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ 1995 L 281/31.

15 Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data 1981, ETS 108.

16 Eleni Kosta, *Consent in European Data Protection Law* (Nijhoff Studies in European Union Law, BRILL Martinus Nijhoff Publischers 2013) 24 with reference to Frederick W Hondius, *Emerging data protection in Europe* (Elsevier 1975) 63 et seqq.

17 Art. 1(2) GDPR and Rec. 1 referring to Art. 8(1) of the Charter for Fundamental Rights of the European Union (Charter); note that the term 'privacy' is not used any longer within the GDPR unlike its predecessor and Convention 108.

18 Art. 1(1) Directive 95/46/EC; Art. 1 Convention 108.

19 Serge Gutwirth and Paul De Hert, 'Regulating Profiling in a Democratic Constitutional State' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Springer Netherlands 2008) 279: "As a rule, personal data may be processed, provided the data controller meets a number of conditions. The rule is a 'yes, but ...' rule.".

rules for processing personal and sensitive data. They contain technical requirements, such as ensuring the integrity and confidentiality of data, as well as ones that demand a balance between the input and output, such as limiting the data collection to what is necessary to achieve a specified purpose. Any data controller must comply with the principles and demonstrate compliance with the principles.[20] The requirement of demonstrating compliance shows that there is no 'right or wrong' implementation of the principles but that their implementation must depend on the *specific case* and the involved risks.[21] In other words, because of the context-specificity multiple ways to implement the data protection principles can co-exist, with some more right or wrong where a definitive answer can only be provided when taking the circumstances, purposes, risks, and remedies into account. Article 5(2) of the GDPR also highlights the personal responsibility of the data controller to determine the adequate measures for the intended data processing.[22] Thereby, Article 5(2) 'serves as a *meta-principle*' as it does not only establish a substantive responsibility of complying with the fundamental principles but also entails a *procedural requirement* of being able to demonstrate such compliance.[23]

The principles are coupled to the *requirement of legality*.[24] The requirement of legality mandates a lawful basis for the processing of personal or sensitive data. The interplay between principles and the requirement of legality found within the GDPR are the product of the compromised approach to data and privacy protection in Europe. As the evolution of data protection law among European countries shows, the approaches in different countries (and later member states adopting the Directive) varied,[25] and to this day influence the interpretation of national courts.[26] From a design perspective such a heterogeneous landscape and understanding of data protection law has engineering implications: Either one designs a system to comply with the (internationally) highest standard of the legal requirements or product variants are built that can adapt to the local regulatory environments.

With the GDPR the focus shifted more and more towards implementing data protection through organizational and in particular technical measures.[27] The implementation of Article 25 of the GDPR introduced the concept of *data protection by design*[28] *and default* into data protection law and thereby requested data controllers to employ technical and organizational measures not only to protect personal data from attacks, leaks, or destruction but overall to ensure that the data protection principles are adhered to. Data controllers must ensure that their engineers and developers implement adequate solutions to protect personal data into their products and services.[29] Failures to include proper measures can result in high fines, as seen in Germany where a company failed to ensure the erasure of personal data of employees (e.g., salary statements, contracts, etc.).[30] Yet, the implementation of technical and organizational measures has its boundaries: The implementation must economically and technically be feasible and the relationship between the risk of the processing and the data protection by design measures set in place must be balanced. In other words, data controllers are not required to "spend a disproportionate amount of resources when alternative, less resource

20 Art. 5(2) GDPR.

21 Horst Heberlein, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten' in Eugen Ehmann and Martin Selmayr (eds), *DS-GVO: Datenschutz-Grundverordnung: Kommentar* (2nd edn, Beck'sche Kurz-Kommentare, C.H. Beck, LexisNexis 2018) 29; European Data Protection Supervisor, 'A Preliminary Opinion on data protection and scientific research' (6 January 2020) https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf (accessed 28 October 2020); Peter Hustinx, 'EU Data Protection Law: The Review of Directive 95/46/EC and the General Data Protection Regulation' in Marise Cremona (ed.), *New technologies and EU law* (The collected courses of the Academy of European Law, Oxford University Press 2017) 154; Milda Macenaite, 'The "Riskification" of European Data Protection Law through a two-fold Shift' (2017) 8(3) *European Journal of Risk Regulation* 506, 525.

22 Heberlein (n 21), 29; Art. 5(2) GDPR refers to "accountability" in the English version of the GDPR, the German wording is "Rechenschaftspflicht" and French wording "résponsabilité"; Lachlan Urquhart and Jiahong Chen, 'On the Principle of Accountability: Challenges for Smart Homes & Cybersecurity' (2020) https://arxiv.org/pdf/2006.11043 (accessed 28 October 2020) 3 et seqq.

23 Urquhart and Chen (n 22), 3 et seqq.; note that Lachlan Urquhart, Tom Lodge and Andy Crabtree, 'Demonstrably doing accountability in the Internet of Things' (2019) 27(1) *International Journal of Law and Information Technology* 1, 10 argue that Art. 5(2) GDPR must be read in conjunction with Art 24 GDPR thereby extending the requirement of (demonstrating) compliance to the whole GDPR.

24 Note that in the EU the principle of lawfulness (Art. 5(1)(a) GDPR) can be interpreted broadly or narrowly. If interpreted narrowly, fulfilling the principle of lawfulness requires establishing an adequate legal ground listed in Art. 6 GDPR. If understood broadly, lawfulness means that no other legal obligations related to the processing of data may be breached and that aside from its legal grounds according to Art. 6 GDPR must be demonstrated. Cf. on said discussion Eike Michael Frenzel, 'Art. 5 Grundsätze für die Verarbeitung personenbezogener Daten' in Boris P Paal and Daniel A Pauly (eds), *Datenschutz-Grundverordnung, Bundesdatenschutzgesetz* (Beck'sche Kompakt-Kommentare, 2nd ed. C.H.Beck 2018) 14 et seqq.

25 Viktor Mayer-Schönberger, 'Generational development of data protection in Europe' in Philip Agre and Marc Rotenberg (eds), *Technology and privacy: The new landscape* (MIT Press 1997).

26 Cf. Rebecca Wong, 'The Data Protection Directive 95/46/EC: Idealisms and realisms' (2012) 26(2-3) *International Review of Law, Computers & Technology* 229, 230; cf. Orla Lynskey, 'The 'Europeanisation' of Data Protection Law' (2017) 19 *Cambridge Yearbook of European Legal Studies* 252, 264 et seqq.

27 While the Directive 95/46/EC already obliged controllers to "implement appropriate technical and organizational measures to protect personal data" (Art. 17 Directive 95/46/EC) its focus rested predominantly on security measures. Nonetheless, courts such as the European Court of Justice (ECJ) already had indirectly required privacy-friendly modifications, such as in the Google vs. Spain decision (C-131/12) which required Google to enable de-indexation (which can be seen as a more privacy-friendly operation). Lee Bygrave, 'Article 25. Data protection by design and by default', *The EU general data protection regulation (GDPR): A commentary* (OUP 2020) 575.

28 The idea of data protection by design aligns with Article 8 of the Charter of Fundamental Rights of the EU which requires the adoption of "technical and organizational measures" to ensure "effective protection." The European Court of Human Rights (ECtHR) also embraced privacy by design ideals in its I v Finland decision. Bygrave, 'Article 25. Data protection by design and by default' (n 27), 575 and I v Finland App no 20511/03 (ECtHR, 17 July 2008) rec. 41 et seq.; Axel M. Arnbak, *Securing private communications: Protecting private communications security in EU law: fundamental rights, functional value chains and market incentives* (Doctoral Thesis, University of Amsterdam IViR 2015).

29 Mireille Hildebrandt and Laura Tielemans, 'Data Protection by Design and Technology Neutral Law' (2013) 29(5) *Computer Law & Security Review* 509, 517; cf. also Lee Bygrave, 'Data Protection by Design and by Default: Deciphering the EU's Legislative Requirements' (2017) 1(02) *Oslo Law Review* 105, 114; Fabian Niemann and Philipp Scholz, 'Privacy by Design and Privacy by Default - Wege zu einem funktionierenden Datenschutz in Sozialen Netzwerken' in Falk Peters, Heinrich Kersten and Klaus-Dieter Wolfenstetter (eds), *Innovativer Datenschutz* (Duncker & Humblot 2012) 109 et seqq.

30 Berliner Beauftragte für Datenschutz und Informationsfreiheit, 'Berliner Datenschutzbeauftragte verhängt Bussgeld gegen Immobiliengesellschaft' (5 November 2019) https://www.datenschutz-berlin.de/fileadmin/user_upload/pdf/pressemitteilungen/2019/20191105-PM-Bussgeld_DW.pdf (accessed 28 October 2020). Smaller fines have been issued based on Art. 25 GDPR in Bulgaria, Greece, Romania. For further cases see GDPR Enforcement Tracker https://www.enforcementtracker.com/# (accessed 28 October 2020).

demanding, yet effective measures exist."[31]

While the scope of Article 25 of the GDPR includes all the principles of the GDPR (i.e., meeting all the requirements of the law) and can thus be seen as a 'hollow norm,'[32] the data protection by design norm differentiates among factors that support 'extra' technical measures or that tip the balance in favor of the data subject and factors that reduce the need to implement technical measures. The former (i.e., factors supporting extra measures) include: *high risks for or impact on the data subject's rights and freedoms*, *'unreasonable' purposes*, and *sensitive context of the processing*. The latter (i.e., factors reducing the burden of implementing technical and organizational measures) include: *costs of the actual implementation of the technical measures and limited scope of the processing* (tied to the purposes of the processing and legitimacy of the purposes). While not strictly mandated by the GDPR, ways to ensure that devices comply with the principles via their software have been promoted by developers (see Section 2.3 "Machine-understandable Data Protection Law"). These approaches encode the principles into devices and try to determine ways to automatically factor in the heterogeneous requirements demands mentioned; however, this requires the creation of complex technical systems.

2.2 Hard- or Softcoding Law

In the aim of a bottom-up approach this article draws on a case study in which a toy robot prototype was developed as a (fictional) learning tool for young children.[33] By taking a toy robot as a use case, one can examine how the legal environment of such a smart product is reflected in its firmware implementation. A toy robot, as will further be elaborated below (see Section 3 "Encoding Data Protection Law: An Imperfect Remedy"), includes various data processing capabilities that challenge the fundamental principles of data protection law (e.g., privacy-sensitive sensors such as cameras, continuous processing of personal data, movable, and used by vulnerable users such as children in their private homes). The design of a toy robot prototype requires an iterative approach, starting from the technical dimensions, considering the data-protection-relevant data flows of the toy robot, and establishing a continuous feedback loop between legal scholars and computer scientists to adapt and augment the data flows of the toy robot to fit the requirements laid out by the law. Those attempts target not only the configuration of the robot itself, but also impact the decision criteria that the robot relies on and on a run-time level the adaptability of the toy robot to changed circumstances.

Privacy-by-design scholars and computer scientists working on machine-understandable data protection law seem to agree that a successful encoding of data protection principles for a given system requires (1) a general description of foundational legal principles, (2) the ability to collect information about legally relevant criteria at run time, (3) specific context- and capacity-tailored decision criteria of how the principles (1) are applied together with the criteria (2), and (4) the ability to act upon the decisions produced by (1-3) by adapting the system's behavior at run time. When designing a privacy-friendly toy robot, (1) is satisfied by building upon available ontologies[34] (see below Section 2.3 "Machine-understandable Data Protection Law"; e.g., the concept of parental consent). (2) is given when the robot obtains context data through its virtual or physical sensors (e.g., the data subject's age or the robot's current location). (3) evaluates the legal principles (from (1)) given the context data (from (2)); e.g., to determine Member State specific parental consent age limits, or information about the data subject's age). And (4) is established when the robot is able to update its procedures when circumstances change (e.g., when the robot moves to a new jurisdiction, or parental consent is not required anymore).

To better distinguish between the different components and implementations of data protection by design approaches we start by the norm addressee: While Article 25 of the GDPR binds data controllers, the implementation in particular of technical measures will rest upon the engineers and developers creating the data processing devices.[35] If developers want to configure a product that adheres to the fundamental principles of data protection law, many design decisions will have to be taken already at the time of designing the software architecture of the system and implementing its software modules and they will need to consider the advice of legal experts. For instance, determining the possible legal grounds for processing, the purposes of processing, or the possible ways and technical means to adhere to the principle of transparency, the minimization of data and limitation of storage, as well as the implementation of security principles will have to be determined when developing a smart product and implemented into the design from the beginning. However, developers can choose to design a robot that does not only reflect a single set of pre-defined purposes or legal contexts but can select among (not: decide or judge) at run time which among a multitude of different possible settings it adopts. In other words, data controllers define collections of parameters with legal implications together with heuristics that allow the robot to select one of these - in this way, the robot can - at run time - adapt to legal, contextual, and technical changes. For instance, a legal change would occur if a smart device moves from one jurisdiction to the other and the age of consent changes (e.g., from France, where the consent age is 15, to Belgium where the consent age is 13). Adapting to this change would require access to the geolocation of the device (component (2) above) in order to ensure that the robot requires a new consent (components (3) and (4)) if the age threshold has changed according to the shared understanding of legal principles (component (1) above). Or as another example, if new security standards are published a robot could automatically change its processing operations to adhere to these new standards (e.g., encryption standards) - this is referred to as "crypto-agility"[36] but follows a very similar architecture in that shared foundational assumptions need to be laid out in a machine-readable way and used as a basis for the contextual adaptation of the system's behavior at run time. Thus we see that this configuration impacts the behavior of the toy robot at run time. This does, however, not make the toy robot per se a norm addressee of the GDPR but merely is a way for data controllers, via their engineers, to ensure that their devices are tailored to local requirements in data protection law and can adapt

31 EDPB, 'Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0' (20 October 2020), at 9.

32 Aurelia Tamò-Larrieux, *Designing for Privacy and its Legal Framework: Data Protection by Design and Default for the Internet of Things* (Issues in Privacy and Data Protection, Springer International Publishing 2018) 209.

33 Kimberly Garcia and others (n 10). The toy robot roams private family rooms, taking pictures of its surroundings every few seconds and analyzing them to identify known people (typically children) within its field of view. Once a person has been identified, the robot stops to perform an educational action, such as playing a song that motivates the identified person to sing along using preselected personalized content, which would be tailored to the child's age and current interests.

34 E.g., DPV3 vocabulary, https://dpvcg.github.io/dpv/ (accessed 28 October 2020).

35 Tamò-Larrieux (n 32), 84 et seq.

36 Bryan Sullivan, 'Cryptographic Agility' (2010) available at http://media.blackhat.com/bh-us-10/whitepapers/Sullivan/BlackHat-USA-2010-Sullivan-Cryptographic-Agility-wp.pdf (accessed 20 December 2020).

over time to new requirements automatically.

This is where the distinction between hardcoding and softcoding comes in. Above, we introduced "softcoding" as the decoupling of decision parameters (e.g., production rules, conditionals, thresholds, etc.) from opaque program code. We argue that this would better enable users to understand, monitor, and adapt systems compared to the "hardcoded" implementation of regulation directly in program code. The inflexibility that this entails does not only have negative consequences regarding the *adaptivity of a system*: Assuming that a device has hardcoded rules, updating the device to for instance a changed legal landscape (e.g., from German to Swiss data protection law) would require sending in the product to upload a different variant of the software. Via softcode, these rules could instead be retrieved at run time and could even be kept up to date with current decisions and case law. In addition, we argue that the hardcoding of such rules undermines the *moral legitimacy* of systems that implement legal code in this way. The moral legitimacy would be negatively impaired because a system is not flexible nor malleable for a user or to outside circumstances. We will elaborate on this discussion further below (see Section 4 "Softcoding as a Path for More Responsiveness, Flexibility, and Transparency").

In contrast, a "softcoded" solution links executable code with regulation that is expressed - readable for humans as well as machines - in openly accessible documents. This has implications on several levels: Regarding the *architectural design of a software system* (or a cyber-physical system), it means that an explicit effort must be taken to decouple such parameters from the compiled, executable, program. Instead, the system would be designed so that the parameters are loaded, at run time, from a remote source (e.g., a publicly available knowledge base or database), where that remote source needs to be semantically aligned with the system (e.g., through a shared ontology, corresponding to component (1) above). Such a system would then be configured to adapt it to different execution contexts (e.g., different jurisdictions) by swapping this remote source while keeping the same executable code. Finally, *during operation*, the system would retrieve the decision parameters from the configured remote source and thereby adapt its execution (corresponding to components (3) and (4) above) given its context (corresponding to component (2) above). The timeliness and frequency of these retrieval operations here depend on the context and the concrete decisions that the system needs to take - in some situations, it might be sufficient to update the parameters only upon specific trigger events (e.g., a location change) while in other circumstances, regular updates might be required.

2.3 Machine-understandable Data Protection Law

To enable systems that adapt to regulation as outlined above, we first require a way to express law so that it can be interpreted by machines, corresponding to component (1) above; these machine-interpretable documents then form the basis of run-time- adaptations (components (3) and (4)) based on context data (component (2)). For several decades, researchers across the domains of computer science, information systems, and law have been working on representing legal circumstances and documents in a way that would make them automatically interpretable by machines in this way. Setting the stage for such automatic interpretations of legal documents are legal support software that cover simple extensions to text processing systems, collaboration tools for contract drafting (e.g., Beagle),[37] contract highlighting/visualization (e.g., LegalSifter)[38] and term extraction (e.g., LegalRobot)[39]. In addition to these tools, the domain of legal document analytics comprises algorithms that can be run across documents from several data sets and dictionaries and support automatic text analysis and legal text mining.[40] The *ontological* modeling of legal terms and their relationships adds the potential of better structuring and indexing information from legal documents to prepare it for more efficient searching and even for automated reasoning, in addition to providing a foundation to better understand legal terms in their context and for semantic integration[41], e.g., to contrast across (legal) domains or jurisdictions, harmonize documents, and as a bridge between technical and legal perspectives.[42] In this field, lightweight ontologies and taxonomies are used for *describing* concepts and domains while domains can also be *axiomatized* through heavyweight ontologies. This axiomatization creates a foundation for automatic problem-solving, such as fully automatic compliance checking,[43] and such automatic checks have been proposed in the context of complying with specific norms of the GDPR.[44]

To enable automatic compliance checks with the GDPR, systems require access to high-level descriptions of data processing actions (e.g., *storing* or *deletion* of data) and to machine-understandable formalizations of the relevant parts of the underlying legal basis (e.g., GDPR).[45] In addition, the software that performs the processing needs to be (automatically or manually) annotated to allow its interpretation in the context of these formalizations and thereby permit the fusing of legal and program code. A current overview of the state of the art in the domain is given by Rodrigues and his colleagues[46]; in addition, researchers have analyzed the GDPR using formal concept analysis to recover concepts, attributes, and implications with the same level of formality and rigor with which the regulation was created with the goal of supporting more GDPR-consistent systems and service design.[47] While a full axiomatization of legal documents such as the GDPR is currently out of reach,[48] it is, based on such manual analysis, possible to encode *aspects of regulations* that should

37 https://www.capterra.com/p/142807/Beagle/ (accessed 28 October 2020).

38 https://www.legalsifter.com/ (accessed 28 October 2020).

39 Sudhir Agarwal, Kevin Xu and John Moghtader, 'Toward Machine-Understandable Contracts' in *AI4J – Artificial Intelligence for Justice* (Workshop at the 22nd European Conference on Artificial Intelligence, The Hague, The Netherlands, August 2016) 5.

40 Charalabidis and others propose a range of applications of such legal text mining including parallel search across legal frameworks that are formulated in different languages, automatic assessment of the degree of transposition of national and international laws (e.g., regarding the relationship of EU Directives and national legislation), comparative analyses of connected laws, timeline analysis including the interrelation of laws and news articles, and text- or even geographically-based visualization. Cf. Yannis Charalabidis and others, 'Use Case Scenarios on Legal Text Mining', in Ben Dhaou Soumaya, Carter Lemuria and Mark A Gregory (eds), *ICEGOV2019: Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance* (Melbourne VIC Australia April 2019, Association for Computing Machinery, 2019) 364.

41 Núria Casellas, *Legal Ontology Engineering* (Springer Netherlands, Dordrecht 2011) 50.

42 Cleyton M d O Rodrigues and others, 'Legal ontologies over time: A systematic mapping study' (2019) 130 *Expert Systems with Applications* 12, 12 et seqq.

43 Rodrigues and others (n 42), 12 et seqq.

44 Piero A Bonatti and others, 'Machine Understandable Policies and GDPR Compliance Checking' (2020) https://arxiv.org/pdf/2001.08930.pdf (accessed 28 October 2020) 1 et seqq.

45 Bonatti and others (n 44), 1 et seqq.

46 Rodrigues and others (n 42), 12 et seqq.

47 Damian A Tamburri, 'Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation' (2020) 91 *Information Systems* 101469.

48 Bonatti and others (n 44), 1 et seqq.

hold unambiguously and without reference to their interpretation contexts. From a technical perspective, such systems thus softcode legal contexts that are described in a transparent way and within openly accessible legal ontologies; and we can even conceive of systems that allow users to modify which of a range of legal (and possibly even personalized) ontologies to use at run time.[49]

Within systems that encode aspects of regulation in this way, one approach towards enabling the automatic processing of contracts, policies, and law is explicit rule-based modeling. These rules are then applied to generate exact legal consequences such as obligations and prohibitions when a specific process execution is identified as part of a monitored workflow.[50] Workflow systems are thereby enabled to initiate actions only after consulting a database with regulatory clauses in order to determine active obligations; the machine-readable representations of clauses and rules however currently need to be created manually. Approaches from the Semantic Web domain, in particular ontologies and vocabularies that are defined using languages from the families of the Resource Description Framework (RDF) and the Web Ontology Language (OWL) can be used for their expressivity and to increase the interoperability of such solutions, while the limits of these standards in the context of conceptualizing the legal domain remain little explored.[51] Legal reasoning is, in principle, defeasible,[52] and it is therefore not possible to decide all juridical nuances using classical logic while formalizing the domain using a monotonic logic only is labor-intensive and might not be understandable by domain experts.[53] Manual encoding of documents by applying non-classical logics may also not scale to a full legal corpus.[54] Moreover, legal rules may conflict with each other, which is resolved through meta-rules that define priority relationships and require defeasible logics.[55] While thus both rule languages (such as LegalRuleML) and languages that correspond to description logics (such as OWL2) have been used as policy languages,[56] policy-reasoning tasks are decidable only in the latter while compliance-checking is undecidable in rule languages, or at least intractable in the absence of recursion.[57]

Researchers have thus been working on the creation of ontologies for the legal domain for several decades with the goals of establishing common and unambiguous terminology and of making the domain accessible to automated processing.[58] Description models of a wide variety of types and on many different abstraction levels have been created. Generally, the manual development of ontologies by knowledge engineers and with the support of domain experts starting from concepts of the target domain is referred to as top-down ontology development and is distinguished from bottom-up approaches where ontologies are extracted by mapping from underlying data sources (e.g., legal documents).[59] In the legal domain, top-down approaches include the *Legal Knowledge Interchange Format* (LKIF) and its *core ontology of basic legal concepts*[60] that is arranged in three clusters: *legal-action*, *legal-role*, and *norm*. To give a concrete example, the norm cluster defines concepts such as *Contract*, *Decree*, and *Treaty*; it then expresses that documents of type *Contract* bear at least one entity of type *Norm* that are held by agents of type *Natural_Person* or *Legal_Person* towards some *Thing* (e.g., an action) that is normatively qualified (i.e., allowed or disallowed).[61]

For applying such an ontology in a practical application, it needs to be complemented with a more specific legal domain ontology and with a formalization and vocabulary of the underlying argumentation and reasoning which represents the structure and dynamics of argumentation that shall be applied.[62] In other words, these models are typically only loosely coupled with the actual legislation text which makes it difficult to verify whether they are effective[63] and accurate with respect to their representation of law. Consequently, there is a lack of practical adoption and the body of academic work is criticized, for instance regarding specific omissions that constrain practical usage.[64] Together with the challenges around the rule-based modeling of the legal domain discussed above, there has thus also not been an instantiation of LKIF and LegalRuleML at scale or used for formalizing or annotating the content of a legal corpora either automatically or manually.[65] To overcome this gap between research and practice, recent work targets the design of semantic systems that can be used to express legal circumstances in *specific domains* (e.g., to express legislative obligations[66]) and often coupled to *specific use cases*. Only then are these connected to more abstract knowledge models—in the case of [67] as an extension profile that can be used to model obligations with the *Open Digital Rights Language* (ODRL).[68] While the design of such extensions is thus from the beginning informed from approaches such as ODRL and LKIF, the implementation is done in a bottom-up way, and the combined system is in addition instantiated in the form of a usable tool.[69]

3. Encoding Data Protection Law: An Imperfect Remedy

Unsurprisingly, the increased deployment of smart devices like social robots has led to an increased interest among academics in

49 Kimberly Garcia and others (n 10).
50 Alan Abrahams, David Eyers and Jean Bacon, 'An asynchronous rule-based approach for business process automation using obligations' in Bernd Fischer (ed.), *Proceedings of the 2002 ACM SIGPLAN workshop on Rule-based programming* (ACM, New York, NY 2002).
51 Rodrigues and others (n 42), 12 et seqq.
52 Juan B Carlos, 'Why is Legal Reasoning Defeasible?' in Arend Soeteman (ed.), *Pluralism and Law* (Springer, Dordrecht 2001).
53 Rodrigues and others (n 42), 12 et seqq.
54 Guido Governatori and others, 'Norm Modifications in Defeasible Logic' in Marie-Francine Moens and Peter Spyns (eds), *Legal Knowledge and Information Systems, JURIX 2005: Eighteenth Annual Conference* (IOS Press 2005) 13 et seqq.
55 Marcello Ceci, 'Combining Ontologies and Rules to Model Judicial Interpretation' in *Proceedings of the RuleML@ECAI 6th international doctoral consortium* (Montpellier, France, August 2012) 2.
56 Bonatti and others (n 44), 1 et seqq.
57 Bonatti and others (n 44), 1 et seqq.; Piero A Bonatti, 'Datalog for Security, Privacy and Trust' in Oege de Moor and others (eds), *Datalog reloaded: First international workshop, Datalog 2010, Oxford, UK, March 16 - 19, 2010, revised selected papers* (Lecture Notes in Computer Science vol 6702. Springer 2011).
58 Rodrigues and others (n 42), 12 et seqq.
59 Biralatei Fawei and others, 'A Semi-automated Ontology Construction for Legal Question Answering' (2019) 37 *New Gener. Comput.* 453.
60 Rinke Hoekstra and others, 'The LKIF Core Ontology of Basic Legal Concepts' in Pompeu Casanovas and others (eds), *Proceedings of the 2nd Workshop on Legal Ontologies and Artificial Intelligence Techniques* (Stanford, CA, USA 2007) 43 et seqq.
61 The LKIF core ontology is available at https://github.com/RinkeHoekstra/lkif-core (accessed 28 October 2020).
62 Ceci (n 55), 2.
63 Sushant Agarwal and others, 'Legislative Compliance Assessment: Framework, Model and GDPR Instantiation' in Manel Medina and others (eds), *Privacy Technologies and Policy: 6th Annual Privacy Forum, APF 2018, Barcelona, Spain, June 13-14, 2018, Revised Selected Papers* (Security and Cryptology vol 11079, Springer International Publishing 2018) 131 et seqq.
64 Agarwal and others (n 63), 131 et seqq.
65 Fawei and others (n 59), 453 et seqq.
66 Agarwal and others (n 63), 131 et seqq.
67 Agarwal and others (n 63), 131 et seqq.
68 https://www.w3.org/TR/odrl/ (accessed 28 October 2020).
69 Cf. Agarwal and others (n 63), 131 et seqq. for GDPR compliance assessment.

the interaction between regulation and social robots.[70] Leenes and Lucivero differentiate between four scenarios: First, the ability of *law to regulate the design of a robot*, second, the *ability of a robot to regulate user behavior* through its design, third, the *ability of law to regulate the effects of a robot's behavior*, and fourth, the *ability of code to regulate a robot's behavior*.[71] Encoding data protection as enshrined in the GDPR focuses in particular on the first and last category mentioned by Leenes and Lucivero: Ensuring that the external and internal design of a robot complies automatically with the fundamental principles of data protection law (e.g., transparency about the data gathering, limitation of data processing practices, deactivation of functionality upon lacking user consent). Thereby, encoding data protection regulates the potential privacy implications and effects of a social robot and thus the impact this robot has on user behavior (e.g., a privacy-friendly robot might increase user comfort, while a privacy-invasive one may lead to chilling behaviors). As mentioned above (see Section 2.2 "Hard- or Softcoding Law") the design process ideally will not only lead to configuring robots with the data protection principles in mind but also constructing devices that at run time can adapt to contextual changes.

As described in Section 2 "From an Ideal to Implementations", while remedies to encode data protection have been proposed, they have encountered various obstacles. In the following, we map the issues that arose in the implementation of the data protection principles in a social robot[72] and refer to other research projects and literature highlighting similar difficulties.[73] While our findings stem from an investigation on the implementation of data protection by design and thus focus on data protection law, they apply to legal code across legal domains. In fact, different examples[74] of encoding of law can be found which show that, depending on the characteristics of the legislation at hand (e.g., ones involving calculations, relying on machine-readable factual information, involving compliance with processes),[75] the difficulties arising in implementing the law into the design vary (see Section 4 "Softcoding as a Path for More Responsiveness, Flexibility, and Transparency"). The difficulties arise in particular when dealing with balancing norms rather than procedural ones (or muddy norms instead of crystal norms[76]). Former norms are more vague and open to interpretation. Here we see difficulties that arise from the need to come up with assumptions (e.g., de facto hierarchies), 'solve' conflicts within the law, determining how to deal with balancing tests and legitimacy criteria, generalize legal terms to encode them, and disentangle connected norms. Moreover, the lack of automatic access to machine-readable documentation and the difficulties of assessing risk complicate the implementation of law into code. Lastly, we discuss the business implications and potential constraints to encoding data protection principles.

3.1 Encoding Assumptions

Encoding data protection implies coming up with solutions when the law is silent, vague, and ambiguous.[77] Doing so requires relying on assumptions, even when those may be well founded and documented in the literature. In that sense, law indulges in the luxury (and, sometimes, necessity) of staying vague, but code cannot.[78] Nonetheless, if no clear case law in favor of one or the other interpretation exists in a general manner, even the most well-argued assumption remains debatable and defeasible. One example that illustrates this difficulty arises when encoding the principle of lawfulness: The purpose of the processing determines the legal ground, which in turn must be established before the processing occurs. Thus already the choice of the legal ground becomes dependent on other characteristics of the processing that are determined at the design stage. In addition, as will be explained below, since no hierarchy of legal grounds can be found within the law or case law, developers will be motivated to create a de-facto normative hierarchy, which ultimately is subjective and imposed by system designers and engineers.

According to the Article 29 Working Party (WP29), the data controller must *determine which lawfulness ground is the most appropriate in a given scenario*. Not all the processing can thus be justified by consent but only instances in which consent is the appropriate lawfulness ground. This provision by the WP29 has been criticized.[79] But case law has made clear that the choice of the appropriate legal basis is key and an *inappropriate ground for processing leads to fines and inability to claim other legal grounds at a later point of time*.[80] One could interpret the WP29 opinion and the cited case law as such that if other lawfulness grounds than consent are applicable, those need to be given priority in the design and implementation process. In other words, a data controller needs to first check whether data can be processed on other legal grounds than consent given its current context, and if that is not the case require consent of the data subject. But of course, such an interpretation is highly controversial,[81] and depending

70 Leenes and Lucivero (n 11), 198; Bibi van den Berg. 'Robots as Tools for Techno-Regulation' (2011) 3 *Law, Innovation and Technology* 319; Christoph Lutz and Aurelia Tamò, 'RoboCode-Ethicists' in *Proceedings of the 2015 ACM Web Science Conference* (Oxford, United Kingdom, June – July 2015).

71 Leenes and Lucivero (n 11), 198.

72 Kimberly Garcia and others (n 10).

73 Leenes and others (n 5); Koops and Leenes (n 2), 159; Leenes and Lucivero (n 11), 193.

74 Cf. for examples e.g., the OECD Working Papers on Public Governance, 'Cracking the code: Rulemaking for humans and machines' (2020) available at https://www.oecd-ilibrary.org/governance/cracking-the-code_3afe6ba5-en (last accessed 20 December 2020).

75 Cf. findings of New Zeland LabPlus in 2018 https://www.digital.govt.nz/dmsdocument/95-better-rules-for-government-discovery-report/html (accessed 8 November 2020).

76 A term coined by Carol M Rose, 'Crystals and Mud in Property Law' (1988) 40 *Stanford Law Review* 577.

77 Cf. Leenes and others (n 5), 28 elaborating on the vagueness, open texture, and ambiguity of law.; cf. also on delineating the scope of data requirements Koops and Leenes (n 2), 163.

78 We note that the law often remains vague for good reasons; we do not mean to disesteem these reasons, but note that the vagueness creates an obstacle to the encoding of law.

79 Winfried Veil, 'Einwilligung oder berechtigtes Interesse? – Datenverarbeitung zwischen Skylla und Charybdis' (2018) 71(46) Neue Juristische Wochenschrift 3337, 3338.

80 EDPB, 'Company fined 150,000 euros for infringements of the GDPR' (31 July 2019) https://edpb.europa.eu/news/national-news/2019/company-fined-150000-euros-infringements-gdpr_en (accessed 28 October 2020) Hellenic DPA fines PWC reason is that the company asked for consent for the processing of data, yet this was seen as an inappropriate legal ground as the processing was covered by another legal ground that was not mentioned to the employees. This decision shows that reversing the legal ground is not readily possible, as the infringement has consequences with respect to the data that has been processed without appropriate legal ground.

81 Even the WP29 seems to have contradicting options on said matter. Cf. Article 29 Working Party, 'Guidelines on consent under Regulation 2016/679' (WP 259 rev.01, 10 April 2018), at 3 in conjunction with 23 and Article 29 Working Party, 'Opinion 8/2014 on the on Recent Developments on the Internet of Things' (WP 223, 16 September 2014), at 15 where the WP29 states that "Consent (Article 7(a)) is the first legal basis that should be principally relied on in the context of the IoT, whether by device manufacturers, social or data platforms, devices lenders or third party developers". The contradiction between the two opinions has not been addressed in the literature. The recent case of the Hellenic DPA (see above footnote 80) shows however clearly that appropriate lawfulness grounds are necessary.

on the interpretation (and national understanding of data protection law as a whole), different approaches could be proposed.[82] One argument to give priority to other legal grounds prior to resulting to consent is the following: Both consent and data processing necessary for the performance of a contract are based on the idea that a user/data subject gives consent to a specific action or manifests an intent to enter into a (contractual) relationship with the data controller. Yet, in particular consent is inherently linked with problems with respect to its efficacy to provide control over data processing.[83] Thus, legal grounds that are not affected (as much) by cognitive biases discussed in the literature shall be given priority. These grounds are based on a legislative process or have been established by case law. In any case these grounds are tied to a democratically established process, which arguably should give them more weight. That being said, the resulting engineering implications are to *determine a hierarchy for testing legal grounds* (e.g., (1) processing based on a legal obligation; (2) processing based on legitimate interests (3) processing necessary for the performance of a contract, and (4) processing based on consent). While such an interpretation enables taking the purpose into account (e.g., in case of processing of data in an employment situation to pay benefits to employee, the first legal ground in the hierarchy could be fulfilled; e.g., in the case of processing for marketing purposes, the fourth legal ground would be fulfilled) to automatically test for a legal ground, the result in practice might be that the de-facto hierarchy set by developers will lead to relying on an inadequate legal ground, as a decision must be taken in order to proceed.

This obstacle sheds light on a difficulty that often arises when trying to embed data protection into the design: The vagueness of the law and potential syntactic ambiguity complicates and potentially impedes such endeavours. In our opinion, while vagueness is acceptable when dealing with balancing tests and legitimacy criteria within the principles, determining the adequate legal ground has a procedural element to it which projects some sort of hierarchy and thus requires more specific guidance – not only for engineers but also for lawyers. In the end, we see here how European data protection law, which in itself is a compromise between different approaches in the member states, fails to reconcile these different approaches to its fullest, which become apparent when trying to embed data protection into the design.

3.2 'Solving' Conflicts in the Law

In addition, *tensions between different principles* need to be addressed and practically feasible processes of how to solve those tensions need to be devised when designing for compliance with data protection law.[84] Clear mechanisms on how to resolve such have not been widely discussed in the literature yet are necessary in order to determine the legal code implications thereof.

In particular, a conflict between the principle of accuracy and data minimization has been raised in the literature.[85] The principle of accuracy aligns with the interests of the data controller, who has an interest in having accurate and up-to-date data.[86] The principle of accuracy is also linked to data security by means of requiring the integrity of the data (i.e., that the data is maintained as it was originally collected) as well as its trueness and veracity.[87] However, even originally correct data that has not been changed can become inaccurate after a certain time has elapsed, as the principle of accuracy is context-dependent.[88] In fact, the principle of accuracy exists not as a stand-alone principle, but as a connecting principle. The 'connecting' aspect of accuracy can for instance be seen in the ECJ's Google vs. Spain decision that ultimately links accuracy of data to the fairness principle, by stating that out-of-context information can lead to unfair decisions or judgements.[89] By that token though, the principle of accuracy does not seem to be much in conflict with the principle of data minimization (which in turn is interlinked with the principle of purpose and storage limitation).[90] A core design feature under the GDPR is to process only the (minimum) data necessary to achieve a specified purpose. This implies also to limit the storage of the data to only that data that is necessary to achieve said purposes. These principles set the data controller under pressure to be able to justify why certain data is being collected, processed, and kept, and thereby strongly decreases the data controller's incentives to keep unnecessary data. In fact, from a technical perspective the principle of accuracy and data minimization can be encoded, for instance by implementing expiration dates on data processing operations.[91]

Another aspect that conflicts with the data minimization principles is the fact that the controller bears the burden of proof that valid consent was obtained when relying upon that legal ground.[92] This

82 The German literature seems to typically praise consent as the ultimate means to establish informational self-determination. Cf. e.g., Marie-Theres Tinnefeld and Isabell Conrad, 'Die selbstbestimmte Einwilligung im europäischen Recht' (2018) 9 *Zeitschrift für Datenschutz* 391, 392; Dirk Heckmann and Anne Paschke, 'Art. 7 Bedingungen für die Einwilligung' in Eugen Ehmann and Martin Selmayr (eds), *DS-GVO: Datenschutz-Grundverordnung: Kommentar* (2nd edn, Beck'sche Kurz-Kommentare, C.H. Beck, LexisNexis 2018) 9. However, other scholars from Germany seem to have a more critical stance, cf. Stefan Ernst, 'Die Einwilligung nach der Datenschutzgrundverordnung' (2017) 3 *Zeitschrift für Datenschutz* 110, 110.

83 Chris J Hoofnagle, Bart van der Sloot and Frederik Z Borgesius, 'The European Union general data protection regulation: what it is and what it means' (2019) 28(1) *Information & Communications Technology Law* 65, 80; Daniel J Solove, 'Privacy Self-Management and the Consent Dilemma' (2013) 126 *Harv. L. Rev.* 1880, 1883 et seqq.; cf. Elettra Bietti, 'Consent as a Free Pass: Platform Power and the Limits of the Informational Turn' (2020) 40(1) *Pace Law Review* 310.

84 Koops and Leenes (n 2), 166 et seq.; Leenes and Lucivero (n 11), 211 et seqq.

85 Cf. Erik Zouave and Jessica Schroers, 'You've been Measured, You've been Weighed & You've been Found Suspicious - Biometrics & Data Protection in Criminal Justice Processing' in Ronald Leenes, Rosamunde van Brakel and Serge Gutwirth (eds), *Data protection and privacy: The Internet of Bodies* (Computers, privacy and data protection 2018) 9; cf. Pagallo (n 2), 343; cf. Michael Veale, Reuben Binns and Jef Ausloos, 'When data protection by design and data subject rights clash' (2018) 8(2) *International Data Privacy Law* 105.

86 Thomas Hoeren, 'Big Data and Datenqualität – ein Blick auf die DSGVO' (2016) 10 *Zeitschrift für Datenschutz* 459; Gloria Gonzáles Fuster, 'Inaccuracy as a privacy-enhancing tool' (2010) 12(1) E*thics of Information Technologies* 87; the principle of accuracy is also a guiding principle in the OECD 1980 and now 2013 Guidelines.

87 Hoeren, (n 86), 459 with reference to the ISO Standard 5725-1: 1994.

88 Cf. Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, Case C-131/12, [2014] (ECLI:EU:C:2014:317), at para. 93.

89 Cf. Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, Case C-131/12, [2014] (ECLI:EU:C:2014:317); Rolf H. Weber and Simon Henseler, 'Regulierung von Algorithmen in der EU und in der Schweiz: Überlegungen zu ausgewählten Regulierungsthemen' (2020) 28 *Zeitschrift für Europarecht* 31.

90 Data minimization relies on the purpose for which the data is being processed as it requires that only data that is absolutely necessary to achieve said purpose is being processed; storage limitation can be seen as a form of data minimization as it requires erasing data that is no longer necessary for achieving a stated purpose.

91 Bart Custers, 'Click here to consent forever: Expiry dates for informed consent' (2016) 3(1) *Big Data & Society* 1.

92 Cf. Art. 7(1) GDPR. EDPB, 'Guidelines 05/2020 on consent under Regulation 2016/679 Version 1.1' (4 May 2020), at 22; Article 29 Working Party,

requires data controllers to 'store the declaration of consent together with the name of the data subject or another reliable identifier (email address, etc.) and the time of the consent ('timestamp')' as long as the processing activity persists.[93] Moreover, whenever data controllers target children (e.g., connected toys), the controller must ensure that parental consent is obtained when the data subject is below a certain threshold.[94] Although the GDPR does not demand the controller to verify the age of the child, it might be inevitable in practice, given that mechanisms for age confirmation can easily be circumvented.[95] Different age verification mechanisms exist, yet it is likely that all of them put at risk the privacy of children by requiring the collection of additional personal data.[96] Moreover, where a device processes data continuously or periodically, it is possible that during this time the child may exceed the age threshold and parental consent is no longer required, but the consent of the child himself. Since relying on parental consent after the child has reached the respective age makes the processing unlawful, the data controller is likely to record not only the declaration of consent together with the name of the data subject or another reliable identifier (as seen above), but also the child's date of birth, in order to ensure that the system can obtain the child's own consent once the child reaches the respective age threshold.[97]

Another conflict in the law can be found in the prospective element of transparency. The wording of Article 22 of the GDPR stipulates a right of the data subject to object to specific forms of automated decision-making practices, yet the article prohibits such practices unless explicit consent is provided. Unsurprisingly, this has triggered a debate on whether it qualifies as a right or as a prohibition all together. This conflict remains unresolved in the literature, as arguments in favor of a right[98] and in favor or a prohibition[99] can be found. Also a historical analysis cannot resolve this conflict, as some member states had interpreted the former Article 15(1) DPD as a prohibition (while others had not).[100] While the WP29 - and the EDPD - however seem to agree that despite the wording as a right Art. 22(1) of the GDPR and its position within Chapter III of the GDPR should qualify as a prohibition,[101] also arguments in favor of an individual right are abundant. Especially, the fact that the information duties of data controllers listed in Articles 12 to 14 of the GDPR include a requirement to mention if automated decision-making occurs (which would not be needed if no such decision-making would be allowed) point towards an individual right.[102]

3.3 Dealing with Legitimacy

Another difficulty arises whenever the law refers to legitimacy criteria and balancing of competing interests. An example thereof is determining when the legal ground of *legitimate interests*, which is only applicable in the context of businesses and users,[103] can be applied. Data controllers may argue – in line with the WP29 statement – that sometimes they 'temporarily need to perform some facial recognition processing steps precisely for the purpose of assessing whether a user has provided consent or not as a legal basis for the processing. This initial processing (i.e., image acquisition, face detection, comparison, etc.) may in that case have a separate legal basis, notably the legitimate interest of the data controller to comply with data protection rules. Data processed during these stages must only be used for the strictly limited purpose to verify the user's consent and should therefore be deleted immediately after.'[104] But this statement does not exempt from an assessment of the reasonable expectations of a data subject at the time of the collection which is based on the relationship with the controller.[105] The reasonable expectation relates to the 'foreseeability and acceptance from the side of the data subject of the processing operation. While the foreseeability needs to be articulated objectively (clear, timely, and transparent information notice, justified for the purposes it serves, etc.) by the data controller, the acceptance of the data subject can also be implied (otherwise, we would refer to 'consent').'[106]

'Guidelines on consent under Regulation 2016/679' (n 81), at 20.

93 Sebastian Dienst, 'Lawful processing of personal data in companies under the General Data Protection Regulation' in Tobias Kugler and Daniel Rücker (eds), *New European general data protection regulation: A practitioner's guide: ensuring compliant corporate practice* (C.H. Beck; Nomos; Hart 2018) 99; cf. EDPB, 'Guidelines 05/2020 on consent under Regulation 2016/679 Version 1.1' (n 92), at 23.

94 Cf. Art. 8(1) GDPR.

95 Article 29 Working Party, 'Guidelines on consent under Regulation 2016/679' (n 81), at 25 et seq.; Eleni Kosta, 'Article 8. Conditions applicable to child's consent in relation to information society services', *The EU general data protection regulation (GDPR): A commentary* (OUP 2020) 360 et seqq.

96 Unicef, 'Children's Online Privacy and Freedom of Expression' (May 2018), https://issuu.com/unicefusa/docs/unicef_toolkit_privacy_expression?e=29613278/60947364 (accessed 29 October 2020) 15; cf. Article 29 Working Party, 'Guidelines on consent under Regulation 2016/679' (n 81), at 27.

97 Cf. Koops and Leenes (n 2), 165 with respect to the Dutch implementation of the DPD; cf. also Kosta, 'Article 8. Conditions applicable to child's consent in relation to information society services' (n 95), 361 et seq.

98 Wulf Kamlah, 'Art. 22 DSGVO' in Kai-Uwe Plath (ed.), *DSGVO/BDSG Kommentar* (3rd edn, Dr. Otto Schmidt 2018) 4; Anton Vedder and Laurens Naudts, 'Accountability for the use of algorithms in a big data environment' (2017) 31(2) *International Review of Law, Computers & Technology* 206, 213 et seq.

99 Cf. e.g., Isak Mendoza and Lee Bygrave, 'The Right Not to Be Subject to Automated Decisions Based on Profiling' in Tatiani Synodinou and others (eds), *EU Internet Law: Regulation and Enforcement* (Springer 2017), 86 et seq.; Lee Bygrave, 'Minding the Machine v.2.0: The EU General Data Protection Regulation and Automated Decision Making' in Karen Yeung and Martin Lodge (eds), *Algorithmic Regulation* (Oxford University Press 2019) 246; Frederike Kaltheuner and Elettra Bietti, 'Data is power: Towards additional guidance on profiling and automated decision-making in the GDPR' (2018) 2(2) *Journal of Information Rights, Policy and Practice* 1, 10 et seq.; Eike Mario Martini, 'Art. 22. Automatisierte Entscheidungen im Einzelfall einschließlich Profiling' in Boris P Paal and Daniel A Pauly (eds), *Datenschutz-Grundverordnung, Bundesdatenschutzgesetz* (Beck'sche Kompakt-Kommentare, 2nd ed. C.H.Beck 2018) 29; Guido Noto la Diega, 'Against the Dehumanisation of Decision-Making – Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information' (2018) 9(1) *JIPITEC* 3, 17.

100 Bygrave, 'Minding the Machine v.2.0: The EU General Data Protection Regulation and Automated Decision Making' (n 99), 6.

101 Cf. Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (WP251 rev.01, February 2018), at 19. Potentially, Article 11 of the GDPR, which exempts data controllers from having to comply with individual rights but excludes Article 22 from this exemption indicates thereby that a difference between individual rights (Art. 15-20 GDPR) and Art. 22 of the GDPR exists. This could be taken to mean that Art. 22 has to be treated differently from individual rights. Cf. also Maja Brkan, 'Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond' (2019) 27 *International Journal of Law and Information Technology* 91, 99.

102 Lee Bygrave, 'Article 22. Automated individual decision-making, including profiling', *The EU general data protection regulation (GDPR): A commentary* (OUP 2020) 531.

103 Rec. 47 excluding the applicability of this legal ground in the case of state and citizens.

104 Article 29 Working Party, 'Opinion 02/2012 on facial recognition in online and mobile services' (WP 192, 22 March 2012), at 5.

105 Cf. Rec. 47; Irene Kamara and Paul de Hert, 'Understanding the Balancing Act Behind the Legitimate Interest of the Controller Ground: A Pragmatic Approach' (Brussels Privacy Hub Working Paper, August 2018) https://brusselsprivacyhub.eu/BPH-Working-Paper-VOL4-N12.pdf (accessed 28 October 2020) 10.

106 Kamara and de Hert (n 105), 17.

Determining when processing can be based on legitimate interests, and taking the reasonable expectations into account, is not a trivial task; and encoding of this process is even more complicated. In fact, to do so requires conducting three tests that are interlinked with one another and lead to an overall balance of interests. To put it differently, the balance of interests test, which is the final step out of three, necessitates two prior steps: a legitimacy of interests test and adequacy test.[107] The *legitimacy test* requires a proof of a legitimate interest by the data controller.[108] According to the WP29, legitimate interests of data controllers must be real and present interests that the data controller has articulated. In other words, future interests, i.e., ones that depend on the fulfilment of a future condition or expectation, are not sufficient. The WP29 also notes that the 'concept of 'interest' is closely related to, but distinct from, the concept of 'purpose'.'[109] While a purpose relates to any aim of the data processing, the interests relate to the broader stake the controller has in the processing and the benefit the controller derives from that processing.[110] An interest is not considered to be legitimate 'where the processing is not genuinely necessary for the performance of a contract but rather relates to the ancillary use of data and is achieved through terms unilaterally imposed on the data subject.'[111] The GDPR mentions examples of legitimate interests such as preventing fraud and direct marketing[112] and ensuring network and information security.[113] Those interests are likewise mentioned by the WP29.[114] In case law, different legitimate interests have emerged: In Case C-708/18[115] in which the court had to determine the legitimacy of installed video surveillance in the common parts of a building, the court weighed the interests in the protection of the property and the health and life of co-workers against the right to privacy. The court saw the data processing as legitimate as it argued that the data controller had no other means available that were less invasive to ensure the mentioned interests. In a similar case[116] the court followed the same argument. In another decision,[117] the court acknowledged that the interests of 'a third party in obtaining the personal information of a person who damaged their property in order to sue that person for damages can be qualified as a legitimate interest.'[118] In a recent case, a Dutch court overturned the Dutch DPA's restrictive interpretation of Article 6(1)(f), according to which commercial interests cannot be legitimate interests. Following the European Data Protection Board's guidelines, the court instead found that purely commercial interests are legitimate interests, provided they are real and not speculative.[119] In contrast to those cases acknowledging a legitimate interest as a legal ground, in its famous *right to be forgotten* decision, the ECJ argued that purely economic interests of the search engine provider in not de-indexing certain information are not legitimate interests.[120] The *adequacy test* looks at whether the processing is indeed necessary to achieve the interests or if less intrusive means would be available.[121] The case law above also illustrates how adequacy/necessity are context- or case-dependent. Lastly, the *balancing test* takes into account the impact of the data processing on the data subject.[122] This requires an assessment that takes the positive and negative (potential) consequences into account.[123] While positive consequences can include the interests of the data controller, those interests can overlap with those of the broader community (e.g., freedom to conduct business, of information, science). Negative consequences include potential adverse effects such as emotional impacts and chilling effects.[124] As such emotional and behavioral impacts are difficult to predict, caution should be employed when arguing such consequences let alone codifying them. Nonetheless, based upon the WP29 Opinion on legitimate interests,[125] some criteria are mentioned that more likely tip the balance in one direction or the other: For instance, if sensitive data such as biometric data is being processed, more severe negative consequences are assumed,[126] likewise in case of data being 'publicly disclosed or otherwise made accessible to a large number of persons, or whether large amounts of personal data are processed or combined with other data.'[127]

The question of legitimacy does not only arise with respect to finding

107 Cf. also Autorité de protection des données, 'Recommandation n°01/2020 du 17 janvier 2020 relative aux traitements de données à caractère personnel à des fins de marketing direct' (17 January 2020) https://www.autoriteprotectiondonnees.be/publications/recommandation-n-01-2020.pdf (accessed 28 October 2020) on these three tests.

108 Kamara and de Hert (n 105), 12.

109 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (WP 217, 9 April 2014), at 24.

110 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), at 24.

111 Róisín Á Costello, 'The Impacts of AdTech on Privacy Rights and the Rule of Law' (2020) *Technology and Regulation* 11, 17 with reference to Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), 16.

112 As mentioned in Rec. 47 GDPR.

113 Rec. 49 GDPR.

114 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), at 25 stating: "conventional direct marketing and other forms of marketing or advertisement", "unsolicited non-commercial messages, including for political campaigns or charitable fundraising", "prevention of fraud, misuse of services, or money laundering", "physical security, IT and network security", or "processing for research purposes (including marketing research)".

115 *TK v Asocia ia de Proprietari bloc M5A-ScaraA*, Case C-708/18 [2019] (ECLI:EU:C:2019:1064).

116 *František Ryneš v Ú ad pro ochranu osobních údaj*, Case C-212/13 [2014] (ECLI:EU:C:2014:2428), at para. 34.

117 *Valsts policijas R gas re iona p rvaldes K rt bas policijas p rvalde v R gas pašvald bas SIA 'R gas satiksme'*, Case C-13/16 [2017] (ECLI:EU:C:2017:336), at para. 30 and the case-law cited.

118 *Productores de Música de España (Promusicae) v Telefónica de España SAU, Promusicae*, Case C-275/06 [2008] (ECLI:EU:C:2008:54), at para. 53.

119 Hunton Andrews Kurth LLP's Privacy and Cybersecurity practice, 'Dutch Court Overturns DPA Fine on Legitimate Interest Legal Basis' (1 December 2020) https://www.huntonprivacyblog.com/2020/12/01/dutch-court-overturns-dpa-fine-on-legitimate-interest-legal-basis/ (accessed 21 December 2020); Ady Nieuwenhuizen, 'Judge overturns Dutch DPA GDPR fine' (26 November 2020) https://www.fieldfisher.com/en/insights/judge-overturns-dutch-dpa-gdpr-fine (accessed 21 December 2020).

120 *Google Spain SL, Google Inc. v Agencia Española de Protección de Datos (AEPD), Mario Costeja González*, Case C-131/12 [2014] (ECLI:EU:C:2014:317), at para. 81.

121 Cf. Rec. 39 GDPR.

122 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), at 36 et seq.

123 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), at 37.

124 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), at 37; Moritz Büchi and others, 'The chilling effects of algorithmic profiling: Mapping the issues' (2020) 36 *Computer Law & Security Review* 105367.

125 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109).

126 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), at 38-39.

127 Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (n 109), at 39.

the adequate legal ground but is a question that is at the core of data protection law. In particular, the principle of purpose limitation states that each processing of data must occur for legitimate purposes. It could be argued that the purposes of processing are legitimate if the processing is lawful according to Article 6 of the GDPR.[128] While this seems reasonable for processing that occurs for the purpose of complying with a legal obligation or to protect vital interests, making the legitimacy of a purpose depending on consent seems less reasonable. In particular because of the failings noted in the literature with respect to consent (e.g., failures with respect to accepting terms that are not read, biases of individuals and inability to calculate long-term risks vs. short-term benefits, others). These failures show that the term *legitimate* must likely be understood more broadly, as *in accordance with the law*. According to the WP29 it should include not only primary and secondary legislation but all forms of written law, principles, and case law.[129] In addition, also codes of conduct and ethics and 'the general context and facts of the case' as well as social and technical changes must be taken into account if they affect the legitimacy of a given purpose over time.[130]

3.4 Generalizing Legal Terms

Many aspects encountered within the data protection law cannot today be expressed in a machine-readable way, meaning that depending on the principle at hand case-by-case considerations are key. This is also true for principles for which there is a rich (and evolving) case law and which ultimately require updates as to the factors that courts took into consideration. This results in decisions that are based on the (partially subjective) weighing of different options, and can lead to (un)intended *generalizations* and delineations.[131]

An example thereof is the interpretation of the term 'fairness'. Fairness, transparency, and lawfulness are all closely linked to one another. This link is already apparent in Article 5(1)(a) of the GDPR which ties the concepts together. In other words, formally speaking the concept of fairness can be seen as the middle ground on a spectrum between lawfulness and transparency, providing a link between both concepts. As such a *middleman*, the ideal of fairness is linked to the concept of lawfulness when fairness reflects procedural fairness; and linked to the concept of transparency when fairness reflects 'fairly transparent' processing. Aside from this, fairness in itself must also be understood as 'effect-based' wanting to mitigate imbalances that lead to vulnerability and discrimination.[132]

Understanding fairness as more aligned with lawfulness means to interpret it as *procedural fairness*.[133] This procedural fairness implies a balanced approach with respect to weighing competing interests against each other. What speaks in favor of understanding fairness as procedural fairness is that in some translations of the GDPR in languages of EU member states the term 'fairness' is translated as a term relating closer to lawfulness.[134] On the one hand, fair balancing means taking the context into account in order to prevent unjust 'outcomes' or 'impacts.' On the other hand, procedural fairness requires implementing guiding procedural rules.

The GDPR refers in numerous articles and recitals to 'fair and transparent' processing.[135] This demonstrates the strong link among fairness and transparency and is linked to the information duties as the data subject must have actual knowledge of the main characteristics of the processing of his or her personal data.[136] While the ECJ has interpreted the concept of fairness as a sort of requirement of transparency in the case of the processing of personal data when public authorities transfer data to other authorities,[137] such an interpretation is also possible within the private sector. As Clifford and Ausloos conclude, the court's reasoning in these cases was to provide protection against asymmetric relationships, even in cases where the sharing of data is not malevolent (i.e., instances in which the controller is not trying to deceive a data subject).[138] Interestingly, the ICO and CNIL likewise understand the term 'fairness' as a means to rebalance asymmetric relationships, among others by means of providing more transparency about the underlying data processing.[139]

While aligning the meaning of fairness with lawfulness and transparency would mean with respect to the engineering implications that the provisions of lawfulness and transparency would need to be followed through (with all mentioned caveats), the term fairness

128 Whether or not one agrees with this argument will depend also on whether the term 'lawfulness' is understood broadly or narrowly. Cf. footnote 24 above.
129 Article 29 Working Party, 'Opinion 03/2013 on purpose limitation' (WP 203, 2 April 2013), at 20.
130 Article 29 Working Party, 'Opinion 03/2013 on purpose limitation' (n 129), at 20.
131 Cf. here Koops and Leenes (n 2), 163.
132 Gianclaudio Malgieri, 'The Concept of Fairness in the GDPR: A linguistic and contextual interpretation' in Mireille Hildebrandt and others (eds), FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (Barcelona Spain January 2020, Association for Computing Machinery New York, NY, United States). Note that the link between fairness and non-discrimination can already be found within the Resolutions of the Council of Europe on the protection of privacy in electronic data banks from 1973 and 1974, cf. Council of Europe, Committee of Ministers, Resolution 73 (22) on the protection of privacy of individuals vis a vis electronic data banks in the private sector; Council of Europe, Committee of Ministers, Resolution 74 (29) on the protection of privacy of individuals vis a vis electronic data banks in the public sector referring both to "unfair discrimination".
133 Malgieri (n 132), 157 with reference to Damian Clifford and Jef Ausloos, 'Data Protection and the Role of Fairness' (2018) 37 *Yearbook of European Law* 130, 140 et seqq.
134 Malgieri (n 132), 157.
135 Rec. 39, 60, and 71 and Art. 13, 14, and 40 GDPR.
136 Cf. Rec. 60 GDPR stating "The principles of fair and transparent processing require that the data subject be informed of the existence of the processing operation and its purposes. The controller should provide the data subject with any further information necessary to ensure fair and transparent processing taking into account the specific circumstances and context in which the personal data are processed. Furthermore, the data subject should be informed of the existence of profiling and the consequences of such profiling. Where the personal data are collected from the data subject, the data subject should also be informed whether he or she is obliged to provide the personal data and of the consequences, where he or she does not provide such data. That information may be provided in combination with standardised icons in order to give in an easily visible, intelligible and clearly legible manner, a meaningful overview of the intended processing. Where the icons are presented electronically, they should be machine-readable." Cf. Jef Ausloos, Michael Veale and René Mahieu, 'Getting Data Subject Rights Right' (2019) 10(3) *JIPITEC* 283, 283.
137 Malgieri (n 132), 157 with reference to Smaranda Bara and Others v Casa Na ional de Asigur ri de S n tate and Others, Case C-201/14, [2015] (EU:C:2015:638); Opinion of Advocate General Campos Sánchez-Bordona delivered on 17 October 2018 (1); Deutsche Post AG v Hauptzollamt Köln, Case C-496/17, [2019] (ECLI:EU:C:2019:26).
138 Clifford and Ausloos (n 133), 140 et seq.
139 Information Commissioner's Office, 'Big data, artificial intelligence, machine learning and data protection Version 2.2' https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf (accessed 28 October 2020) 19 et seqq.; Michael Butterworth, 'The ICO and artificial intelligence: The role of fairness in the GDPR framework' (2018) 34(2) Computer Law & Security Review 257, 257 et seqq.; CNIL, 'Algorithms and artificial intelligence: CNIL's report on the ethical issues' (25 May 2018) https://www.cnil.fr/en/algorithms-and-artificial-intelligence-cnils-report-ethical-issues (accessed 28 October 2020).

includes also another—own—dimensions, *the mitigation of imbalances and prevention of discriminatory practices*.

Technical measures must be implemented that prevent data processing practices that might lead to discriminatory effects.[140] From a technical perspective the question remains what sort of technical measures are adept to discover discriminatory effects and mitigate them. The discovery and mitigation is a tricky if not impossible task because EU courts have interpreted and applied non-discrimination law heterogeneously.[141] It has therefore been claimed that the concept of fairness understood as non-discrimination cannot be implemented into automated systems: 'While numerous statistical metrics exist in the technical literature, it is not possible to reliably capture a European conceptualization of discrimination which is, by definition, contextual.'[142] This statement seems to focus in particular on cases of indirect discrimination where context matters most. In cases of direct discrimination (based on protected categories) non-context-related categories will be decisive.[143] While contextuality and flexibility of non-discrimination law and its interpretation is advantageous for many reasons (e.g., ensuring that the individual case receives the attention it deserves, that contextual factors such as time and relationships are reflected in the decision, that conflicting rights are balanced against each other, others), at the same time the contextuality of said laws renders their technical implementation impossible.[144]

These findings with respect to the technical implementation of fairness understood as the prevention of non-discriminatory practices lead to the conclusion that even if technical tools working towards fairness—in the use case for instance software that ensures the same accuracy rate of recognition of children faces irrespective of their ethnicity—can be employed, such tools will never fully be able to adhere to the fairness principle.[145] Taking the example of facial recognition, this is thus currently not possible, and it is likely that no system will ever be able to adhere to the principle.

3.5 Disentangling Connected Requirements

Requirements under the law are often connected across documents and domains. However, encoding them in a feasible and transparent way requires disentangling these dependency chains. One example thereof is the user-focused principle of transparency,[146] which includes two different elements, a prospective (incl. the continuous ability to have access to prospective information)[147] and a retrospective one.[148] While the former is an active information duty, the latter is more reactive and its scope has triggered a lively academic debate in particular on the establishment of a right to explanation[149] and the qualification of Article 22 of the GDPR[150] (see also Section 3.2 "'Solving' Conflicts in the Law").

The prospective information duty under the GDPR is *active*, meaning that the data controller must actively inform the data subject in an easily accessible manner (e.g., by way of a direct link, QR codes, SnapTags, NFC, dashboard) about the ongoing data processing. The burden of finding the information does not rest on the data subject.[151] To this end, the WP29 introduced the concept of push notice (i.e., just-in-time information notices) and pull notices (e.g., through a dashboard with the possibility to obtain further information).[152] From a design perspective it is key to avoid information overload,[153] which is why a layered approach to complying with the prospective information duty can be useful.[154] In itself, the prospective element contains multiple requirements which each trigger not only an individual implementation but one that puts each element into its bigger context.

One key information element is to *facilitate exercising individual rights* under Articles 15 to 22 of the GDPR.[155] Making use of one's individual rights does not require a specific motive; Curiosity about one's personal data being processed by a smart product must suffice to trigger an obligation of the data controller to provide said information.[156] A dashboard solution facilitates fulfilling this requirement and has been

140 This can be read into Rec. 71 GDPR explicitly mentions the use of technical measures to ensure that the processing does not lead to discriminatory effects.

141 Sandra Wachter, Brent Mittelstadt and Chris Russell, 'Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI' https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3547922 (accessed 28 October 2020) 5 et seq.

142 Wachter, Mittelstadt and Russell (n 141), 5.

143 For further elaboration on the problem of antidiscrimination doctrine in the context of automated systems, see, e.g, Raphaële Xenidis and Linda Senden, 'EU Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination' in Ulf Bernitz and others (eds), *General Principles of EU law and the EU Digital Order* (Kluwer Law International 2020), 151 https://ssrn.com/abstract=3529524 (accessed 28 March 2021); Frederik J. Zuiderveen Borgesius, 'Strengthening legal protection against discrimination by algorithms and artificial intelligence' (2020) 24(10) *The International Journal of Human Rights* 1572; Philipp Hacker, 'Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law' (2018) 55(4) *Common Market Law Review* 1143.

144 Wachter, Mittelstadt and Russell (n 141), 5 et seq.; cf. Hacker (n 143), 1146.

145 See Emre Kazim, Jeremy Barnett and Adriano Koshiyama, 'Automation and Fairness: Assessing the Automation of Fairness in Cases of Reasonable Pluralism and Considering the Blackbox of Human Judgment' https://ssrn.com/abstract=3698404 (accessed 28 March 2021).

146 Cécila de Terwangne, 'Article 5. Principles relating to processing of personal data', *The EU general data protection regulation (GDPR): A commentary* (OUP 2020) 314.

147 See here Article 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (WP 260 rev.01, 11 April 2018) at 10.

148 Frenzel (n 24), 21; Heike Felzmann and others, 'Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns' (2019) 6(1) *Big Data & Society* 1, 2.

149 Cf. on the subject: Bryan Casey, Ashkon Farhangi and Roland Vogl, 'Rethinking Explainable Machines: The GDPR's 'Right to Explanation' Debate and the Rise of Algorithmic Audits in Enterprise' (2019) 34(1) *Berkeley Technology Law Journal* 143; Lilian Edwards and Michael Veale, 'Slave to the algorithm? Why a "right to an explanation" is probably not the remedy you are looking for' (2017) 16(1) *Duke Law and Technology Review* 18; Margot E Kaminski, 'The Right to Explanation, Explained' (2019) 34(1) *Berkeley Technology Law Journal*; Andrew D Selbst and Julia Powles, 'Meaningful information and the right to explanation' (2017) 7(4) *International Data Privacy Law* 233; Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a right to explanation of automated decision-making does not exist in the general data protection regulation' (2017) 7(2) *International Data Privacy Law* 76.

150 Cf. e.g., Mendoza and Bygrave (n 99), 86 et seq.; Bygrave, 'Minding the Machine v.2.0: The EU General Data Protection Regulation and Automated Decision Making' (n 99), 246; Kaltheuner and Bietti (n 99), 10 et seq.; Martini (n 99), 29; Noto la Diega (n 99), 17.

151 Article 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (n 147), at 8.

152 Article 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (n 147), at 20 et seq.

153 Centre for Information Policy Leadership, 'Recommendations for Implementing Transparency, Consent and Legitimate Interests under the GDPR' (17 May 2017) https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_recommendations_on_transparency_consent_and_legitimate_interest_under_the_gdpr_-19_may_2017-c.pdf (accessed 28 October 2020) 2.

154 Ausloos, Veale and Mahieu (n 136), 286.

155 Art. 12(2) and 13(2)(b) GDPR.

156 Ausloos, Veale and Mahieu stating that individual rights are "intent-agnostic/motive-blind", Ausloos, Veale and Mahieu (n 136), 305 with reference to case law of national courts.

suggested by data protection authorities as well as scholars.[157] While a dashboard allows individuals to make use of their rights, such an action must trigger a predefined technical action in the background.[158] These actions will have to depend on the categories of data being processed. In fact, if a data subject consents only to a fully data-minimized processing (e.g., only locally stored data without third party or data controller access), making use of individual rights may become obsolete following Article 11 of the GDPR. From a design perspective, the exemption of Article 11 of the GDPR introduces a sort of hierarchy, as the provision indicates that the principle of data minimization must be given priority even if that means not being able to then fulfil individual rights. In many instances though, smart devices will rely on data processing of the data processor (e.g., use of external facial recognition software). Here, encoding data protection encounters technical constraints. In the concrete case of machine-learning-based facial recognition for instance, erasing the uploaded training data is possible, but erasing or rectifying inferences by machine-learning algorithms with respect to the classification is not feasible in general. Such 'unlearning' has become a topic of research in the machine-learning community,[159] however no satisfying approaches that can be applied generally exist to-date. In addition, similar to differential privacy systems, machine unlearning will imply trade-offs between the performance of a learning system and its unlearning ability.

3.6 Lack of Automatic Access to Relevant Structured Information

Prospective transparency duties extend to providing data subjects with information about what data is transferred to third countries and what adequacy measures are set in place to do so. Here, in a first step, one has to determine where (regarding geographical location) data flows when it 'leaves' a smart device. This becomes an issue when external processing is involved; The data sharing agreement should state where data is being processed in order to enable to determine automatically if the data is stored and processed in a country that falls under the 'adequacy decision list'[160] of the Commission. This list could also be automatically parsed by a computer system at regular intervals—in its current form with the help of heuristics that extract the individual country names from the list. It would, however, be desirable if regulatory information such as this was published in a machine-readable format and, ideally, would be linked to open data sources such as Wikidata that already contains representations of sovereign nations (e.g., representing the country Switzerland[161]). When data is not stored or processed in such an 'adequate' country or by a certified company, a device has to check whether binding corporate rules are in place that contain 'enforceable data subject rights and effective legal remedies for data subjects.'[162] These can take the form of standard contractual clauses adopted by the Commission,[163] which would be 'attached' to the data sharing agreements.[164] By means of Natural Language Processing (NLP) the agreements could be searched for such addendums and classified as such in order to provide a user with that information. Yet, this does not equal actual reading the agreements but merely provides for a fast way to check whenever data is processed in a country outside the adequacy decision list, if standard contractual agreements were signed. This would however require storing machine-processable representations of the contracts, which might often not be the case. One, albeit manual, possibility is to create and attach these documents in machine-readable formats (e.g., based on ODRL or LKIF, see Section 2.3 "Machine-understandable Data Protection Law") — this information could then be presented to users in a similar way to the transparency interface.[165]

While other approaches to fulfil this requirement exist, such as approved codes of conduct pursuant to Article 40 of the GDPR or certification mechanisms, the multitude of options complicates the technical codification of double checking whether this information requirement must be fulfilled and, if so, what information must be provided. Furthermore, to date, no standard format or mechanisms are established that could be used to implement automatic compliance checks of corporate rules or certificates and publication of which corporate rules or certificates that prove compliance with the GDPR. In other words, the four options to prove compliance if there are no adequacy decisions—namely binding corporate rules, use of standard contractual clauses, corporate rules,[166] or certifications—would require multiple additional steps and relying on information provided by the companies employing them and data protection authorities that are not easily available.

The challenges point also to policy-making needs: If encoding data protection in the spirit of Article 25 of the GDPR should become reality (or at least initiatives building towards it encouraged), measures that enable the extraction of relevant information is key. This requires an effort not only from data controllers, but also from data protection authorities to work towards standardizations and open-access of information that is published as machine-readable structured data

157 Cf. e.g., Article 29 Working Party, 'Guidelines on transparency under Regulation 2016/679' (n 147), at 10; cf. also Information Commissioner's Office, 'Guide to the General Data Protection Regulation (GDPR)' https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf (accessed 28 October 2020) 90; cf. Philip Raschke and others, 'Designing a GDPR-Compliant and Usable Privacy Dashboard' in Marit Hansen and others (eds), *Privacy and identity management: The smart revolution: 12th IFIP WG 9.2, 9.5, 9.6/11.7, 11.4, 11.6/SIG 9.2.2 International Summer School, Ispra, Italy, September 4-8, 2017; revised selected papers* (IFIP Advances in Information and Communication Technology vol 526. Springer 2018).

158 Note that bystanders, whose image data is processed based on legitimate interests, do not have access to the dashboard needed to obtain information about the processing. To facilitate the information access and align with the principle of transparency, a visible QR code could be included onto the device's surface leading a bystander to further information about how data about non-users are being processed. This should take into account the concrete consent settings for the device in question which are stored by the data controller: thereby, bystanders would be informed about the concrete processing that their data undergoes.

159 Lucas Bourtoule and others, 'Machine Unlearning' (2020) https://arxiv.org/abs/1912.03817 (accessed 28 October 2020).

160 European Commission, 'Adequacy decisions' https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en (accessed 28 October 2020).

161 https://www.wikidata.org/wiki/Q39 (accessed 28 October 2020).

162 Art. 46(1) GDPR.

163 Art. 46(1)(c) and (d) in conjuncture with Art. 93(2) GDPR.

164 Note that according to the ECJ's Schrems II decision the standard contractual clauses remain valid but it must be determined on a case by case basis whether in a particular transfer of data setting the clauses are legitimate. Cf. *Data Protection Commissioner v Facebook Ireland Ltd, Maximillian Schrems*, Case C-311/18, [2020] (ECLI:EU:C:2020:559), at para 134 and 149.

165 Bonatti and others (n 44), 1 et seqq.

166 In accordance with Art. 47 GDPR. Here, too, the verification that those are in place is not a straightforward issue that can easily be programmed. Technically, the simplest solution would be to verify whether the competent supervisory authority approved the corporate rules that have to fulfil a catalogue of requirements set out in Article 47(2) of the GDPR. However, this requires knowing which authority is in charge of approving the binding corporate rules of the external party and having said authority publish (and regularly update) a list elaborating which corporate rules it approved.

and thereby can directly be used by systems.

3.7 Dealing with Risk

The GDPR has intensified the debate on how to classify risks that occur with respect to the data protection rights of individuals.[167] On a macro-level two understandings must be differentiated: A broader interpretation of the risk-based approach applies the concept on both, compliance and enforcement of the GDPR; A more narrow understanding, applies it as an obligation targeted at data controllers.[168] On a more micro-level two further understandings of the risk-based approach must be differentiated: The WP29 approach separating between risks and compliance,[169] and Gellert's argument to understand risks as 'compliance risk.'[170] Even if only focusing on a micro-level, taking a risk-based approach requires differentiating between these two understandings. While the WP29 approach seems confusing and goal oriented (by acknowledging the need for flexibility as well as the danger of a risk-based approach for fundamental rights),[171] Gellert's approach relies upon the scalability notion of compliance.[172] He argues that two elements of risk must be differentiated: First, the event-element of a compliance risk is the lack of compliance altogether, and second, the consequence-element of compliance risk which is the resulting risk to the data subject's rights and freedoms.[173] On a meta-level, these interpretations show the challenges of dealing with risk when designing or even encoding data protection principles.

Even when dealing with the data security principle, where the measures that are specified in the law align with the technical understanding of how to keep data confidential, integer, and available at all times,[174] specifying the risks is not a trivial task. While the alignment of technical and legal objectives enables a more straightforward implementation of technical measures to achieve 'legal' aims, it remains difficult to automatically assess the internal and external risks and corresponding redress mechanisms. In fact, two steps are required to determine the engineering implications of the principle of data security. In a first step, the (external and internal) *risks of each data flow* including storage must be discussed. The risk will depend on the sensitivity of the data processed. For instance, biometric data (such as facial attributes) are more sensitive than other data. Therefore, the impact for the data subject if such data is exposed in a security incident is higher than for other data. In a second step, the *redress measures*, and the extent to which they minimize the outlined risk in the first step, must be described. Such measures include the erasure of training data after the training or only storing network credentials in an encrypted format and only for as long as they are required. With respect to the encryption format, future technological developments (also with respect to decryption schemes) must be taken into consideration.[175] This requirement leads to the responsibility to keep the system up to date—as mandated by Article 32 of the GDPR—which in turn implies a constant update of the recommended level of encryption according to established industry guidelines.

From a business perspective, a conscious weighing of strategic, user experience, and legal aspects (and risks) becomes necessary, which is hard to automate. It requires the data controller to balance the overhead in the design and implementation of the system, a possibly inferior user experience, and strategic business implications against the assumption of compliance risks and the overhead of properly managing collected information (e.g., secure storage, provisioning of data access to users, others). These decisions however need to be taken on a per-use-case, per-product, or even per-processing-purpose basis.

3.8 From Smartness to Dumbness?

An overly strict encoding of data protection principles – meaning that the necessity of much of the processed data is questioned and thus rejected – might lead to an overall reduction of the smartness of a device. In the extreme, this results in the design of a system that is unable to easily restore user passwords, thus undermining the positive perception of a product by users for the sake of maximizing the minimization of data collection. While such an extreme maximization of the data minimization principle goes against the inherent balancing notion of the GDPR, it is true that such an interpretation of the principle of data minimization and storage limitation can preclude several features of a product that are heralded as some of the prime advantages of 'digitalized business models.' For instance, if a device's location is not disclosed, the data controller cannot track its products (e.g., for supply-chain optimization). And if a device does not upload any diagnostics data, said data cannot be used by the data controller for product improvements or pre-emptive software updates or hardware repairs which might endanger the security of data and users; this also undermines rental and leasing business models. These modifications thus turn a 'smart device' into a more 'traditional' product. Moreover, an engineering decision to host the configuration dashboard locally instead of relying on a remote dashboard (i.e., a Website) to configure the system would lead to more complicated setups and higher cost on the side of the data controller while deteriorating the user experience. There are also *strategic* implications for the data controller: For example, the supplier of a smart device will need to weigh between its ambition to become independent of third-party facial recognition services (by storing uploaded images and using them to improve its own algorithms) and strict adherence to data minimization and storage limitation. The adherence to GDPR thus will require the data controller to find a balance between business aspects (e.g., ability to become independent of third-party services; ability to deliver an optimal user experience; ability to implement digital business models; others) and the legal risk and responsibility it assumes. This might, for some data controllers, lead to a 'minimal

167 Raphaël Gellert, 'Understanding the notion of risk in the General Data Protection Regulation' (2018) 34(2) *Computer Law & Security Review* 279, 279 et seq.; Lina Jasmontaite and others, 'Data Protection by Design and by Default: Framing Guiding Principles into Legal Obligations in the GDPR' (2018) (4)2 *European Data Protection L Rev* 168, 180 et seq.

168 Macenaite (n 21), 515.

169 Article 29 Working Party, 'Statement on the Role of a Risk-Based Approach in Data Protection Legal Frameworks' (WP 218, 30 May 2014), at 2.

170 Gellert (n 167), 284.

171 According to the WP29, individual data protection rights should be granted regardless of the level of risks of the processing and fundamental principles "should remain the same, whatever the processing and the risks for the data subjects." At the same time however, the WP29 also acknowledges that the fundamental principles are always applied in a context and are thus "inherently scalable." Moreover, the WP29 acknowledges that there are "different levels of accountability obligations depending on the risk posed by the processing in question." This statement is however again followed by a "but", as "controllers should always be accountable for compliance with data protection obligations including demonstrating compliance regarding any data processing whatever the nature, scope, context, purposes of the processing and the risks for data subjects are." Article 29 Working Party, 'Statement on the Role of a Risk-Based Approach in Data Protection Legal Frameworks' (n 169), at 3.

172 Gellert (n 167), 281 et seq.

173 Gellert (n 167), 282.

174 Tamò-Larrieux (n 32), 186 et seq.

175 Gerald Spindler and Philipp Schmechel, 'Personal Data and Encryption in the European General Data Protection Regulation' (2016) 7(2) *JIPITEC* 163, 172 with reference to Rec. 26 GDPR.

data design' where, in addition to the data that is absolutely required to leave the device for fulfilling its purpose (i.e., the images required for facial recognition), only information necessary for recording user consent is uploaded to the data controller, and might thus either undermine or in the extreme impede the data controller's business model.

In the extreme, data controllers could be motivated to only transiently process data in the hope that this qualifies as anonymous from the very beginning on (at the point of collection). The difficulties of achieving a state of full anonymization have been well documented, with various studies showing the identifiability of alleged anonymized data.[176] The GDPR though does not mandate full anonymity to fall outside its realm but a state of anonymization that is *not likely* to be reversed. To achieve this, one needs to not only look at the data itself (including the anonymized data), but also consider other resources that would reasonably enable re-identification.[177] This approach to anonymization under the GDPR has been criticized to overlook part of the risks of re-identification which are not only related to the data and resources available for identification but also depend on the motivation of the adversary to re-identify data, the security of the infrastructure in place, and the potential for mistakes that would lead to a disclosure allowing for re-identification.[178]

As mentioned, one measure that has been debated in the literature as a means to obtain anonymized data is *transient data processing*, i.e., technologies that merely sense their environment and process data ephemerally without storing it.[179] The legal reasoning that is key in this debate is the relative approach interpretation to personal data established by the ECJ.[180] In fact, a strict application of this approach would likely mean that transiently processed personal data that cannot be retrieved will fall outside the scope of the GDPR. Yet, following other interpretations of the term 'personal data,' such as the WP29 arguments that personal data can be established by purpose or result,[181] transient data processing may very well be covered under the GDPR.[182]

In any case, with respect to machine learning, transient processing can only relate to the raw data and not the learning aspect. Any transient data processed by machine learning algorithms influences the algorithms (the machine 'learned' something from it) and this derived or learned data (or parameters) is permanently kept within the system without the option to easily erase such derived data and undo its effects on the trained model.

Transient processing of the raw data can be combined with *local processing* such as image recognition with a pretrained local model as for instance Google's Inception-v3. If data can only be accessed by the owner of the device, it is questionable whether protection in this case is necessary. Similarly, the French Data Protection Authority (CNIL) argued that biometric data processing within smartphones falls under the *household exemption* if the biometric device is incorporated within a smartphone that only locally stores biometric templates of a user (e.g., fingerprints) and prevents the biometric data from being accessed from outside.[183] CNIL calls such a device an 'enclave' or 'sealed box.'[184] This reasoning does then not require an extensive analysis of whether personal data is being processed, but merely an analysis of whether data can be accessed from 'outside.' CNIL issued some rules for such technology to fall under the household exemption, such as: A user must use a device *privately*; the user has the *choice* to decide whether his or her data is being processed within the device (i.e., there must be alternative ways of unlocking a device in the case of biometric authentication); the data can by no means be *shared* with the outside (i.e., also external bodies cannot override this function); the stored data is *encrypted* by state of the art cryptographic algorithm and key management; and all technical solutions are *technically reliable*, i.e., the system is trustworthy.

These discussions show that encoding the principle of data minimization to its fullest can – depending on the design – result in avoidance of falling within the scope of the GDPR.[185] Yet, ephemeral processing of data also results in a reduction of the smartness of devices. How to balance these two aspects will depend on the context and purpose of processing. We see multiple examples where reduction of smartness

176 E.g., Latanya Sweeney, Akua Abu and Julia Winn, 'Identifying Participants in the Personal Genome Project by Name' (Data Privacy Lab, IQSS, Harvard University. White paper, 2013) https://privacytools.seas.harvard.edu/files/privacytools/files/1021-1.pdf (accessed 28 October 2020); Alexandra Wood, David O'Brien and Urs Gasser, 'Privacy and Open Data Research Briefing' https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2842816 (accessed 28 October 2020); Article 29 Working Party, 'Opinion 05/2014 on Anonymisation Techniques' (WP 216, 10 April 2014).

177 With respect to anonymized data scholars have debated when encrypted data can be considered anonymous data. According to Spindler and Schmechel, encrypted data might only be anonymous data if only the data subject him or herself has access to the decryption key (but not in scenarios where the data controller still has access to both). The authors argue that in instances where the data controller does not have access to the decryption key, illegal attacks could still occur, yet that those do not have to be taken into account when determining if data is personal or anonymous. Cf. Spindler and Schmechel (n 175), 172 with reference to Opinion of Advocate General Campos Sánchez-Bordona, delivered on 12 May 2016, Case C-582/14 – *Patrick Breyer v Bundesrepublik Deutschland*. Cf. Rec. 26 GDPR.

178 Mark Elliot and others, 'Functional anonymisation: Personal data and the data environment' (2018) 34(2) *Computer Law & Security Review* 204, 205 et seqq. with further references.

179 Cf. Damian George, Kento Reutimann and Aurelia Tamò-Larrieux, 'GDPR bypass by design? Transient processing of data under the GDPR' (2019) *International Data Privacy Law* 285; Peter Davis, 'Facial Detection and Smart Billboards: Analysing the 'Identified' Criterion of Personal Data in the GDPR' (2020) *University of Oslo Faculty of Law Research Paper No. 2020-01* https://ssrn.com/abstract=3523109 (accessed 28 October 2020) 1; Maša Gali and Raphaël Gellert, 'Data protection law beyond identifiability? Atmospheric profiles, nudging and the Stratumseind Living Lab' (2021) 40 *Computer Law & Security Review* 105486.

180 *Breyer*, Case C-582/14, [2016] (ECLI:EU:C:2016:779). Note that the Breyer decision did not fully exclude the possibility of following an absolute approach. A vagueness that has been criticized by scholars.

181 Article 29 Working Party, 'Opinion 4/2007 on the concept of personal data' (WP 136, 20 June 2007), at 10; *Peter Nowak v Data Protection Commissioner*, Case C-434/16, [2017] (ECLI:EU:C:2017:994), at para 35 where the court argues that inferences about an individual are personal data as such information "by reason of its content, purpose or effect, is linked to a particular person."

182 Article 29 Working Party, 'Opinion 3/2012 on Developments in Biometric Technologies' (WP 193, 27 April 2012), 19 in which the WP29 states "it is not important to identify or verify the individual but to assign him/her automatically to a certain category." However, the WP29 does not mention whether such a categorization still involved the processing of personal data, "nor does it appear that the WP29 was cognisant of smart billboards that process data ephemerally"; cf. Davis (n 179), 11 et seqq.

183 CNIL, 'Biométrie dans les smartphones des particuliers: application du cadre de protection des données' (24 July 2018) https://www.cnil.fr/fr/biometrie-dans-les-smartphones-des-particuliers-application-du-cadre-de-protection-des-donnees (accessed 28 October 2020). We acknowledge that the ECJ has not decided on said issue and has traditionally taken a restrictive approach to interpreting the household exemption, cf. e.g., Urquhart and Chen (n 22) with further references. The ECJ has clearly stated that if data remains accessible to an unrestricted number of people or concerns public spaces this will not fall under the household exemption.

184 CNIL, 'Biométrie dans les smartphones des particuliers: application du cadre de protection des données' (n 183).

185 George, Reutimann and Tamò-Larrieux (n 179), 285 et seqq.

and even accuracy and traceability does not hinder achieving meaningful purposes (e.g., the COVID-19 tracking app based on D3PT,[186] or differential privacy models implemented by Google and Apple[187]). We believe that *leading by example* plays a crucial role in the field of legal code. It is however no surprise that DP3T and differential privacy models have emerged in academia. They require a time-consuming process and close collaboration between technical and legal researchers which are less likely to occur in companies that are driven by economic competition. The interdisciplinary collaboration though is central to these successes. The adoption of such technologies by states and companies shows their significant merit and demonstrates that 'imperfect remedies' might lead to good enough technology that balances different needs.

A path forward includes learning from these attempts to embed privacy protection into the design of technology and moves towards responsible technology by design. Achieving this requires a broader understanding and approach towards legal code and thinking about a *softer way of encoding legal principles* in a form that permits flexibility, transparency, and contestability.

4. Softcoding as a Path for More Responsiveness, Flexibility, and Transparency

As discussed in Section 2.3 "Machine-understandable Data Protection Law", the quest to encoding legal principles into software is not new and is currently gaining traction also outside of academia. From an industry standpoint, this would for instance enable more flexible variant management (e.g., when the same hardware is shipped to different legislations together with its firmware) and for facilitated adaptation of products to end users. The creation of machine-executable legal norms can also bring automation benefits to governments, for instance when aspects of regulation that include simple logic reasoning or mathematical operations are encoded. This is the case with New Zealand's Rates Rebate Act.[188] There, the government's Service Innovation Lab (LabPlus) wanted to rewrite the Rates Rebate Act (a tax rebate for low-income homeowners) in order to respond faster to citizen requests. To do so they first created pseudocode, which is still human-readable text but with defined consistent terminology. This pseudocode was then implemented as machine-executable instructions in the Python programming language. The LabPlus team stated in their final report that such an implementation is feasible for processes-oriented regulation (like the Rates Rebate Act) that involves 'factual information to determine application, eligibility, entitlement,' and prescribes a 'process that is used repeatedly' and one that 'can be delivered digitally.'[189] Similar initiatives can be found world-wide, with for example the OECD issuing a recent working paper on 'Rules as Code' which likewise promotes the creation of machine-consumable law.[190] In addition, researchers have even started to experiment with machine-learning systems that attempt to forecast decisions in case law.[191] How promising those attempts are, remains to be seen.[192]

The examples show that even if legal scholars lament the imperfectness of the interlinking of code and law these instruments are being created and deployed. In that sense, it is not a matter of 'whether' design-based regulation should be employed, but much more on '*how*' we want it to be developed.

While the issues in Section 3 "Encoding Data Protection: An Imperfect Remedy" are mostly of translational nature, they point to two further clusters of challenges: *System-related* and *moral* ones. Addressing the challenges also means taking into account the different ways legal code can be implemented.

4.1 How Softcoding Mitigates Some of the Translational Challenges

The *translational challenges* show that law is more than just written text. It is constantly interpreted, adjusted to a specific context, and adapts over time. However, this is not true for all legal provisions either: The law is not vague in every aspect. Moreover and from a data controller's perspective, regulators could – if a need arises – be more precise, and even publish aspects of regulation (e.g., encryption standards, tax rebates calculations, or lists of countries that are considered safe to transfer data to) in a machine-readable way so that this information can be readily consumed by software and acted upon. What that means is that translational issues should be resolved by taking steps towards the middle ground and asking what norms can and cannot - and should and should not - be made more amenable.

While softcode does not help per se to deal with translational issues (e.g., how to ensure that no generalizations are projected into the code, no assumptions are made on how to interpret the law, etc.), it allows for systems to be more transparent, malleable, and responsive. Such decoupling thus enables a system to adapt over time to its regulatory environment; enabling change is an important aspect to deal with translational issues, in particular in light of how interpretations of law may change over time. The system's higher responsiveness that derives from the decoupling of major decision parameters through softcoding would thus simplify the updating of the system and thereby reduce the probability that the system remains non-compliant.

4.2 How Softcoding can Address System-Related and Moral Challenges

On a broader perspective, *system-related challenges* arise with respect to *who* should be in charge of developing code that adapts to its legal environment and how transparent such code is made to the public. The New Zealand example shows clearly a collaboration effort and involvement of the government to achieve a machine-executable Rates Rebate Act. Other initiatives, like the one the authors are

186 Cf. https://github.com/DP-3T/documents (accessed 28 October 2020).
187 Cf. https://developers.googleblog.com/2019/09/enabling-developers-and-organizations.html and https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf (accessed 28 October 2020).
188 https://www.digital.govt.nz/dmsdocument/95-better-rules-for-government-discovery-report/html (accessed 8 November 2020).
189 Service Innovation Lab (LabPlus), 'Better Rules for Government, Discovery Report' (March 2018) 27 https://www.digital.govt.nz/dms-document/95-better-rules-for-government-discovery-report (accessed 8 November 2020).
190 OECD Working Papers on Public Governance, 'Cracking the code: Rulemaking for humans and machines' (2020) available at https://www.oecd-ilibrary.org/governance/cracking-the-code_3afe6ba5-en (accessed 20 December 2020).
191 Cf. Kevin D Ashley, 'A Brief History of the Changing Roles of Case Prediction in AI and Law' (2019) 36(1) *Law in Context* 93, 103 et seqq.
192 E.g., in Estonia the idea of implementing AI judges was raised. However, no official information on the success or failure of this project can be found. A news article on said topic dates back to 2019: Eric Niller, 'Can AI Be a Fair Judge in Court? Estonia Thinks So' (*Wired*, 25 March 2019) https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/ (accessed 11 November 2020). Another example is the CaseCruncher Alpha, an artificial intelligence that became famous through a challenge where it was able to predict the outcome of cases with greater accuracy than the lawyers involved: Rory Cellan-Jones, 'The robot lawyers are here - and they're winning' (*BBC News*, 1 November 2017) https://www.bbc.com/news/technology-41829534 (accessed 8 November 2020).

following in an implementation of data compliant code is based on open-source software and decoupled, standardized legal vocabularies and ontologies and can thus in principle be held under scrutiny by users and judges alike. Yet, companies will likely not promote open-source legal code initiatives. As Herbert Burkert said already in 1997 with respect to privacy-enhancing technologies (PETs): 'PET design must be open to participatory elements. This implies designing PETs and implementing them in social systems must involve those whom these enhancements are supposed to serve.'[193]

Adopting a softcoding approach, for instance by coupling code with openly accessible ontologies that render regulation machine-readable, the data controller opens the possibility to let the end user fine-tune compliance settings of smart products, thereby increasing transparency and participation. It is even conceivable that individual agents in the society create and publish carefully crafted alternative legal ontologies that subclass a legal domain's legislation and might go beyond it (or might selectively ignore aspects of it to enable disobedience, see below). Like-minded individuals could then further develop and share these documents and point their own smart products towards them.

In addition, softcoding could help to address *moral challenges* that arise predominantly because of the lack of engagement or choice of an individual when confronted with techo-regulation. Mireille Hildebrandt talks here about a lack of buffer between the rules and the one who is ruled; in her own words: 'Rather, under the Rule of Law the legal system acts as a buffer between ruler and ruled, creating the possibility to contest state-authority in an appeal to a court that is in fact supported by the authority of the state (the paradox of the Rechtsstaat).'[194] The crucial functionality represented by a buffer is the preservation of the option of (civil) disobedience.

The ability to disobey is fundamental to moral agency. Moral agency requires the freedom to act and vulnerability with respect to the consequences one suffers if one breaks the rule.[195] Freedom to act can be impaired by legal code; yet does not have to. Karen Yeung describes three scenarios using the same road safety technology: Code that automatically stops a car at red lights. The scenarios then differ by the goals three individuals are trying to pursue: A *criminal minded-person*, who wants to cross a road at red to hurt others; a *person* who masters self-restraining most of the time but sometimes does cross at a red light; and a *Good Samaritan* who wants to cross at red to help someone else in an emergency situation. Yeung shows that the criminally-minded person still has agency to harm others in other ways; that the self-restraining person loses physical agency but not moral one (even though that person will not get praise for abiding the law without the legal code); and that the Good Samaritan has to determine other means to achieve her or his goal, but can still be seen as morally praiseworthy independent on the action he or she chooses (i.e., other means or riding the car to the hospital despite the red lights).

The discussion shows that the moral challenges should not be described with broad brushstrokes. To the contrary, they require a nuanced discussion. Softcoding approaches must be open and flexible enough to preserve the possibility for disobedience. Design-based regulation should thus not lead to 'regulation by technology'[196] but what we could call '*regulation nudged by technology*.' Such as speeding cameras that nudge individuals to comply with the speed limit while driving, technology can by default nudge individuals to comply with the rule yet allow for informed disobedience as well as contestability of those parameters (e.g., speeding to ensure that a woman in labor gets to the hospital in due time and contesting the rule due to an emergency situation). While the default value can be compliance, it must be made easy to modify the technology to - in certain instances - not comply with the rule.

In contrast to hard-coded legislation, softcoding approaches that couple a system with a default legal ontology that can be replaced by the user preserve the ability of the individual, and of society, to exert civil disobedience. The Good Samaritan from Yeung's example would be able to point her car at an ontology that does not regiment it into stopping at a red light, or one where this behavior can be overridden by the user. In principle, she could also create such a version of the machine-readable regulation herself or together with others, and publish it. Such folksonomy-based approaches would thereby pave the way to keep society in the loop.[197]

4.3 Calling for Transdisciplinary Experts

Lastly, while the literature to encoding data protection principles has proposed both, bottom-up and top-down approaches,[198] we believe that bottom-up approaches, which require legal, implementation, and business strategy teams to engage in interdisciplinary communication and collaboration are more fruitful and enable meta-deliberation processes that are much needed in the field of legal (soft) code. In contrast to top-down approaches, iterative and bottom-up approaches encourage a deeper cross-disciplinary understanding and creative solution finding. This aligns more with the reality that open-text legal documents bring along such as ambiguity that leaves room for case-by-case interpretation by legal professionals who need to interpret facts of a case given subjective words or phrases and in the context of national and international legislation that might be connected to the investigated text corpus through opening clauses. While interdisciplinary collaboration is the starting point, we believe that there is a need to train transdisciplinary experts that that can 'deal with emerging value conflicts'[199] arising from the deployment of new technologies. Such transdisciplinary experts should be equipped with tools and strategies to resolve value conflicts and promote the design of responsible technology.

5. Conclusion

Neither hardcoding nor softcoding of regulation into software systems and cyber-physical systems are perfect. In contrast to hardcoding, where regulation is hard-wired into code at a given time and cannot be easily adjusted when regulation changes, softcode attempts to tie code to regulation through loose coupling. This can be accomplished for instance by means of ontologies that are publicly accessible and interpretable by users. Yet, no matter whether regulation is soft- or hardcoded, various issues remain: The need to encode

193 Herbert Burkert, 'Privacy-Enhancing Technologies: Typology, Critique, Vision' in Philip E. Agre and Marc Rotenberg (eds), *Technology and Privacy: The New Landscape* (MIT Press 1997) 125, 135.

194 Hildebrandt, 'Legal Protection by Design: Objections and Refutations' (n 7), 236.

195 Yeung (n 2), 9 et seqq.

196 Cf. Hildebrandt, 'Legal Protection by Design: Objections and Refutations' (n 7), 247 et seq.

197 Cf. on the idea and implementation of society-in-the-loop Iyad Rahwan, 'Society-in-the-loop: Programming the algorithmic social contract' (2018) 20 *Ethics and Information Technology* 5.

198 Cf. Section 2.3 "Machine-understandable Data Protection Law"; cf. e.g., Jaap-Henk Hoepman, 'Privacy Design Strategies' in *29th IFIP International Information Security Conference* (Marrakech, Morocco, June 2014); Seda Gürses, Carmela Troncoso and Claudia Diaz, 'Engineering Privacy by Design Reloaded' in *Amsterdam Privacy Conference* (Amsterdam, The Netherlands, 2015).

199 Lutz and Tamò (n 70).

assumptions because of a lack of clarity in the law, to resolve conflicts within legal norms, and to generalize terms in order to ensure compliance remain critical problems that arise. The advantage of softcode with respect to those issues is only that the system can be improved and changed over time to adapt to new legal circumstances (e.g., court decisions that have clarified legal terms and solved specific conflicts).

These issues are of translational nature, but go beyond the pure translation of law into code as they trigger systemic and moral issues as well. Systemic issues arise from a lack of transparency and the actors involved in the creation of legal code. While here, too, softcode provides some remedies, depending on how legal code is created (based on deterministic or more probabilistic decision-making systems) and by whom (state-driven initiatives vs. industry-driven ones), the opacity of legal code will remain. However, softcode would open the possibility of creating transparency tools that would enable developers and also laypersons to inspect the legal code that drives their products. Furthermore, moral issues are triggered by the lack of engagement between the ruler and the one who is ruled. Crucially, this lack of engagement can curtail civil disobedience which is key to allow social change within a society. With softcode, and the civil disobedience that it can guarantee on the individual and societal levels through folksonomy-enabled meta-disobedience, these moral issues can in principle be overcome.

Overall, the findings within this article point thus to the need for a broader yet more nuanced discussion. Future research needs to map and investigate the current designed-based regulation deployment and initiatives, their effect on individuals and society at large, their openness, the architectural decoupling of implementations and legal code, the involved decision-making (deterministic vs. probabilistic approaches), and the actors involved in the design of legal code. To do so requires not only expertise in computer science and law but calls upon the expertise of multiple disciplines within the social science community.

Acknowledgements

This work was supported by the Hasler Foundation (project #19086). The authors thank Damian George for his helpful comments on earlier versions of this article as well as Emre Bayamlıoglu and Ronald Leenes for their insightful comments during the Digital Legal Talks.