ISSN 2807-2812

The Regulation of Personal and Non-Personal Data in the Context of Big Data

Daniar Supriyadi[a*]

[a] Tilburg Institute for Law, Technology and Society, Netherlands
*corresponding author: dnr.spy@gmail.com

ARTICLE INFO

Article history
Received: July 15, 2022.
Revised: February 13, 2023.
Accepted: February 14, 2023.

Keywords
Big Data;
Electronic;
Regulations;

ABSTRACT

Data protection laws provide minimum protections for personal data, as well as facilitate the free flow of such data, by setting out principles and rules for legitimate data processing. In the big data context, personal data may not be as easy to distinguish as in traditional data processing, and that makes policy-makers and businesses turn to the identifiability concept: in other words, what data are personal. This research is based on doctrinal legal research on the legal theory (concepts, rules, and principles) concerning data protection in the EU and Indonesia. The results of the research show that the understand such paramount terminology in data protection law, relevant factors are presented to assess the direct or indirect identification of a natural person. In the EU data protection law, the test entails, for example, risk-based measures and technological development, whereas Indonesian law on data protection has not yet established such assessments. Data within big data operations traditionally falls under the scope of data protection laws only if it discloses the private life of individuals, such as names or other civil identities, but without further conditions to ascertain whether the data can be indirectly identified with an individual.

1. Introduction

Big data emerged seven years ago – relatively recently – as the technology and analytics applications related were developed around 2010. In the big data era, data analytics have been utilised as a supporting system for improving decision making in the public and private domains.[1] In developing countries, like Indonesia, the application of big data analytics is still in its infancy.[2] Within this

[1] Lu Zhang, '"Personal Information of Privacy Nature'"under Chinese Civil Code', *Computer Law and Security Review*, 43 (2021) https://doi.org/10.1016/j.clsr.2021.105637

[2] Yuanxin Li and Darina Saxunová, 'A Perspective on Categorizing Personal and Sensitive Data and the Analysis of Practical Protection Regulations', *Procedia Computer Science*, 170 (2020), 1110–15 https://doi.org/10.1016/j.procs.2020.03.060

context, public authorities are beginning to employ big data analytics for media monitoring, meaning that big data is used as a supporting instrument in realising better public decision making. These days, society lives in technological turbulence led by information technology – so-called hyperhistory – which subsequently moulds society into a 'risk society.'[3]

In a risk society, (i) society no longer tolerates risks, which leads it to attempt to anticipate and prevent unfavourable situations, and (ii) data are the driver of change within the society, which increases data collection to support a know-everything mentality.[4] Hence, governments' use of media monitoring to deduce societal concerns may begin to be used to help support decision making, while previously governments were apparently unable to derive insights directly from society because of the immaturity of information technology. Furthermore, with regard to evidence-based policies and improving the accuracy of predictions for legal and policy solutions, media monitoring has become important to generating better predictions in reading the society.[5]

In that vein, the Indonesian government launched an official website last year (www.idb.kominfo.go.id) designed to collect national issues from mass media (media monitoring) to support the decision-making process. The website was created and managed by the Directorate of Information and Public Communication in the Ministry of Communication and Informatics (Kominfo), which engages in media monitoring on a daily basis to analyse the contents of media sources that involve government policies and programs.[6] It creates summaries of national news (headlines) gathered from numerous media sources, which are then presented by subject, such as the economy, societal welfare, political news, and security.[7] Kominfo operates as a supporting scheme for success in satisfying national aspirations, the performance of the executive cabinet, and

[3] Cayetano Valero and others, 'Analysis of Security and Data Control in Smart Personal Assistants from the User's Perspective', *Future Generation Computer Systems*, 2023 https://doi.org/10.1016/j.future.2023.02.009

[4] Catherine Oksas and others, 'Perspectives of Peripartum People on Opportunities for Personal and Collective Action to Reduce Exposure to Everyday Chemicals: Focus Groups to Inform Exposure Report-Back', *Environmental Research*, 212.December 2021 (2022) https://doi.org/10.1016/j.envres.2022.113173

[5] Chiara Acciarini and others, 'How Can Organizations Leverage Big Data to Innovate Their Business Models? A Systematic Literature Review', *Technovation*, In Press.May 2022 (2023), 102713 https://doi.org/10.1016/j.technovation.2023.102713

[6] Tijs van den Broek and Anne Fleur van Veenstra, 'Governance of Big Data Collaborations: How to Balance Regulatory Compliance and Disruptive Innovation', *Technological Forecasting and Social Change*, 129.September 2017 (2018), 330–38 https://doi.org/10.1016/j.techfore.2017.09.040

[7] Hui Na Chua, Jie Sheng Ooi, and Anthony Herbland, 'The Effects of Different Personal Data Categories on Information Privacy Concern and Disclosure', *Computers and Security*, 110 (2021), 102453 https://doi.org/10.1016/j.cose.2021.102453

accelerating the dissemination of information about national policies and programs, as well as for assisting in evidence-based policy decisions.[8]

In private domains, e-commence actors utilise big data analytics for various purposes with regard to their services. Smaller firms often use a third-party data analytics service to track their consumers' preferences. For example, an online webshop with limited resources may conclude an agreement with a third party which undertakes data analytics in social media to provide information on their clients' market preferences.[9] Big firms apparently engage such analytics to enhance the interoperability of data for the decision-making process and to enhance their corporate governance and communication capacities. Meanwhile, smaller businesses are eager to utilise big data analytics, but their major obstacle is their lack of competent data scientists. Therefore, these companies often collaborate with data-mining third parties.[10]

The major goal of private firms is to earn profits by providing goods and services, satisfying customers and other stakeholders by providing value while developing and sustaining a competitive edge. When businesses use big data analytics to gain insights from their customers, by monitoring or tracking using website cookies, for example, they certainly bear some responsibility to ensure privacy and personal data protection.[11] Inadequate compliance with privacy and data protection rules may hinder, or even forestall, the application of big data. This gives rise to the question of at what point big data analytics in the business sector is lawful and what matters should be incorporated into the big data platform, in particular pre-conditions (*ex ante*), data processing rules, and ex-post data operations.[12]

Historically, the adoption of big data has been led by the business sector in order to fulfill its needs and facilitate the efficient data flow for better decision making; the public sectors then started to use big data to help them serve their

[8] Yuncheng Shen and others, 'Personal Big Data Pricing Method Based on Differential Privacy', *Computers and Security*, 113 (2022), 102529 https://doi.org/10.1016/j.cose.2021.102529

[9] Tianqi Liu and others, 'Technologies for Removing Pharmaceuticals and Personal Care Products (PPCPs) from Aqueous Solutions: Recent Advances, Performances, Challenges and Recommendations for Improvements', *Journal of Molecular Liquids*, 374 (2022), 121144 https://doi.org/10.1016/j.molliq.2022.121144

[10] Kamyar Hasanzadeh and others, 'A Context Sensitive Approach to Anonymizing Public Participation GIS Data: From Development to the Assessment of Anonymization Effects on Data Quality', *Computers, Environment and Urban Systems*, 83.April (2020), 101513 https://doi.org/10.1016/j.compenvurbsys.2020.101513

[11] Galia Marinova, Aida Bitri, and Marsida Ibro, 'The Role of Women in the Digital Age: Balancing Professional and Personal Challenges during the COVID-19 Pandemic', *IFAC-PapersOnLine*, 54.13 (2021), 539–44 https://doi.org/10.1016/j.ifacol.2021.10.505

[12] Zuzanna Warso, 'There's More to It than Data Protection-Fundamental Rights, Privacy and the Personal/Household Exemption in the Digital Age', *Computer Law and Security Review*, 29.5 (2013), 491–500 https://doi.org/10.1016/j.clsr.2013.07.002

citizens and overcome national challenges.[13] In both the public and private spheres, excluding law enforcement or judicial matters, data about a person which can distinguish him or her from another person must be processed under the regulatory obligations of data protection laws, which emphasise the nature of the data, particularly 'personal data,' rather than the users of the data (i.e. data controllers, data processors, and recipients).[14] Hence, it is appropriate to review how public and private actors approach compliance with the rules and principles of protecting personal data. Even though the use of data in the era of big data in these two sectors may differ, the nature of data is similar, given a majority of data is generated by people. In other words, people are the producers and consumers of data.[15]

Compliance with the legal provisions of privacy and data protection can be determined by national laws concerning personal data protection. In the beginning this concept centred on safeguarding the confidentiality of data, but the development of automated data processing has given significant influence to the increased protection of the rights of the individual. In this vein, the notion of 'privacy' becomes foundational and is reconceptualised in terms of 'control' over personal information. Likewise, in the EU, data protection is inscribed as a fundamental right.[16] The first international convention on data and privacy, Convention 108, designates 'data protection' as corresponding to the right to privacy, and emphasises the 'fair information practice' doctrine. Accordingly, in 1995, the Data Protection Directive imported into EU law the idea that 'data protection' serves as (informational) privacy (Article 1(1) of Directive 95/46/EC).[17]

Furthermore, the EU provides a distinct right to 'personal data protection' inscribed in Article 16(1) of the Treaty on the Functioning of the European Union (TFEU), as introduced by the Lisbon Treaty 2007, which establishes the principle that everyone has the right to the protection of personal data concerning them. Likewise, the EU Charter also refers to 'personal data protection,' rather than 'data

[13] Yue Liu, 'User Control of Personal Information Concerning Mobile-App: Notice and Consent?', *Computer Law and Security Review*, 30.5 (2014), 521–29 https://doi.org/10.1016/j.clsr.2014.07.008

[14] Luca Bolognini and Camilla Bistolfi, 'Pseudonymization and Impacts of Big (Personal/Anonymous) Data Processing in the Transition from the Directive 95/46/EC to the New EU General Data Protection Regulation', *Computer Law and Security Review*, 33.2 (2017), 171–81 https://doi.org/10.1016/j.clsr.2016.11.002

[15] Anna Konstantinovna Zharova and Vladimir Mikhailovich Elin, 'The Use of Big Data: A Russian Perspective of Personal Data Security', *Computer Law and Security Review*, 33.4 (2017), 482–501 https://doi.org/10.1016/j.clsr.2017.03.025

[16] Joshua Yuvaraj, 'How about Me? The Scope of Personal Information under the Australian Privacy Act 1988', *Computer Law and Security Review*, 34.1 (2018), 47–66 https://doi.org/10.1016/j.clsr.2017.05.019

[17] Gianclaudio Malgieri and Bart Custers, 'Pricing Privacy – the Right to Know the Value of Your Personal Data', *Computer Law and Security Review*, 34.2 (2018), 289–303 https://doi.org/10.1016/j.clsr.2017.08.006

protection'.[18] For almost twenty years after 1995, the 1995 Data Protection Directive set out rules and principles of personal data processing in the EU. However, the directive will be repealed on 25 May 2018 by a new legal instrument: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, which addresses the protection of natural persons with regard to the processing of personal data and the free movement of such data, (General Data Protection Regulation, or GDPR).[19]

In general, data must be securely protected, including the general confidential information of businesses or governmental documents, intellectual property, healthcare information, personal financial information, and security information. However, this paper primarily focuses on the protection of data, in particular, personal data, and these kinds of data may fall under the scope of protection of data protection law on condition these data constitute the personal information of a natural person.[20] Hence, the analysis is in line with the DGPR, which articulates personal data protections. However, in the era of big data, data analytics with machine learning create difficulty in ascertaining whether data are personal or non-personal. Measures may be taken as guidance to better define the grey boundaries of the scope and limits of data that should be regarded as personal.[21]

Big data challenges data protection principles, in particular the fairness of the processing and what criteria or methods to use to evaluate whether data are personal or not (identifiability). In that regard, the legal provisions set out in Indonesian law seem to simplify the legal definition of personal data, which entails 'any information inherently attached to an individual and [that] can be identified directly or indirectly' (Article 1(2), the Regulation of the Ministry No. 20 of 2016 on personal data protection in the electronic system).[14] The GDPR provides a better response to the development of advanced data analytics, including machine learning, and has foreseen some of the technological turbulence surrounding personal data processing.[22]

---

[18] Mark Elliot and others, 'Functional Anonymisation: Personal Data and the Data Environment', *Computer Law and Security Review*, 34.2 (2018), 204–21 https://doi.org/10.1016/j.clsr.2018.02.001

[19] Vagelis Papakonstantinou and Paul de Hert, 'Big Data Analytics in Electronic Communications: A Reality in Need of Granular Regulation (Even If This Includes an Interim Period of No Regulation at All)', *Computer Law and Security Review*, 36.November 2015 (2020), 105397 https://doi.org/10.1016/j.clsr.2020.105397

[20] Guan Zheng, 'Trilemma and Tripartition: The Regulatory Paradigms of Cross-Border Personal Data Transfer in the EU, the U.S. and China', *Computer Law and Security Review*, 43 (2021), 105610 https://doi.org/10.1016/j.clsr.2021.105610

[21] Georgios Georgiadis and Geert Poels, 'Towards a Privacy Impact Assessment Methodology to Support the Requirements of the General Data Protection Regulation in a Big Data Analytics Context: A Systematic Literature Review', *Computer Law and Security Review*, 44 (2022) https://doi.org/10.1016/j.clsr.2021.105640

[22] Bart Custers and Gianclaudio Malgieri, 'Priceless Data: Why the EU Fundamental Right to Data Protection Is at Odds with Trade in Personal Data', *Computer Law and Security Review*, 45 (2022), 105683 https://doi.org/10.1016/j.clsr.2022.105683

In the future, both regulations may become much more relevant ("rejuvenate") by responding to extensive critiques and clarification issues in order to match the actual practice of data processing, or it could successfully safeguard the rights of individual to control their data in terms of access to and rectify it.[23] This creates reasonable doubts as to whether the existing data protection law is appropriately addressing and evaluating current practices, in order to benchmark future practices. The question that follows is: how does big data analytics present challenges to data protection laws, and how do the regulatory obligations in such laws delineate personal and non-personal data? Does the law sufficiently address personal data protection in the era of big data, the Internet of Things, cloud computing, machine learning, and artificial intelligence (AI)? Perhaps the primary objective of big data is to derive new insights to predict outcomes and behaviour based on the enormous volumes of data collected from a large number of sources. Each data source, in turn, typically contains data that relates to numerous data subjects. Data protection laws try to have the best approach to conform to the potential challenges of big data systems, with regards to security, speed, interoperability, and analytics capabilities. [24]

The primary research problem of this thesis focuses on the legal challenges of big data analytics for public and private domain applications, in particular a decision-making process and a marketing strategy.[25] The analysis will focus on answering the question of what data are personal in the context of big data, and in particular the identifiability concept of personal data. According to data protection laws, personal data refers to any information of an identified or identifiable natural person, especially by reference to an identifier such as a name, an identification number, location data, an online identifier, or another identity of that natural person.[26] Such personal information often becomes a primary discussion in cyber security and privacy, specifically big data security and data protection. Scattered nonpersonal data may also come very close to the level of protection for personal data, as the nature of the data can reveal the identity of a natural person in some cases. Therefore, in light of big data analytics, revisiting the concept of personal data should be encouraged within the legal framework of data protection in Indonesia and the EU.

2. Research Method

This research is based on doctrinal legal research on the legal theory (concepts, rules, and principles) concerning data protection in the EU and Indonesia. This research is explanatory (explaining the law), hermeneutical (interpretation, argumentation), and

[23] Tuulia Karjalainen, 'The Battle of Power: Enforcing Data Protection Law against Companies Holding Data Power', *Computer Law and Security Review*, 47.August 2018 (2022), 105742 https://doi.org/10.1016/j.clsr.2022.105742

[24] Yu li Liu and others, 'Privacy in AI and the IoT: The Privacy Concerns of Smart Speaker Users and the Personal Information Protection Law in China', *Telecommunications Policy*, 46.7 (2022), 102334 https://doi.org/10.1016/j.telpol.2022.102334

[25] Zhao ge LIU, Xiang yang LI, and Xiao han ZHU, 'Scenario Modeling for Government Big Data Governance Decision-Making: Chinese Experience with Public Safety Services', *Information and Management*, 59.3 (2022), 103622 https://doi.org/10.1016/j.im.2022.103622

[26] Jinglin Jiang and others, 'Deciphering Big Data in Consumer Credit Evaluation', *Journal of Empirical Finance*, 62.August 2020 (2021), 28–45 https://doi.org/10.1016/j.jempfin.2021.01.009

evaluative (analysing whether rules work in the given situation, or whether they are in accordance with desirable moral frameworks, legal principles, and societal aims). Part of the analysis in relation to the research question uses supporting disciplines, namely law and technology. The author provides a comparison of rules, cases, principles, and conceptual frameworks of legal doctrine between the EU and Indonesia. This research elaborates the research problem within a theoretical framework using relevant legal data, notably normative and authoritative sources. Normative sources include statutory texts, treaties, general principles of data protection law and privacy, and the like. Authoritative sources constitute case law and scholarly legal writing (literature). This research will be conducted using a problem-based approach: assembling facts, identifying legal issues, analysing problems with a view to searching for potential solutions, and arriving at a tentative conclusion.

3. Results and Discussion

Big Data Analytics and Its Impact on Data Protection Law

'Big data' is a broad term which covers almost all forms of data processing operations, some of which have well-defined and identified, while others may still be opaque concepts. Such data analytics have been developed and utilised in diverse sectors in various forms and for various purposes, not merely for online database search engines and Internet social media platforms.[27] In scientific and financial circles, big data includes everything from the meteorological data of weather stations to the market data of financial exchanges around the world. The etymology of 'big data' has been traced to the mid-1990s, when it was first introduced by John Mashey, retired former Chief Scientist at Silicon Graphic, to refer to the handling and analysis of massive data sets.[28]

Today, many companies refer 'big data' when they collect and process data about *people,* specifically their customers. This data is useful to better sell products, target marketing efforts, or make better products by collecting valuable customer data from internal and external sources. These sources include social media, email, customer feedback, call records, transactional systems, content management, expert opinions, wikis, support labelling, CRM (customer relationship management) systems, supply chain and fulfilment, DBMS (database management systems), and other external sources, collectively known as an 'enhanced 360° view of the customer.' At this point, companies know more about customers and can advertise more effectively. For instance,

[27] Cheng yong Liu and others, 'Analysis of Beijing Tianjin Hebei Regional Credit System from the Perspective of Big Data Credit Reporting', *Journal of Visual Communication and Image Representation*, 59 (2019), 300–308 https://doi.org/10.1016/j.jvcir.2019.01.018

[28] S. Antusch and others, 'Intentional Action and Limitation of Personal Autonomy. Do Restrictions of Action Selection Decrease the Sense of Agency?', *Consciousness and Cognition*, 88.January (2021), 103076 https://doi.org/10.1016/j.concog.2021.103076

insurance companies have been digging into online data for years and 'mining' Facebook to identify risky people.[29]

Thus, although there is no consensus definition of big data, it has been utilised by some companies to gain much more knowledge of their consumers. Big data may be viewed as property, as a public resource, or as an expression of an individual entity that makes possible unexpected discoveries, innovations, and advancements in quality of life. The definition of big data is contingent on the opinion of the computer scientist, financial analyst, or entrepreneur pitching the phrase.[30] In December 2016, PR Newswire announced Global Strategic Business Report, which found that the leading players in big data are IBM, SAP, Oracle, HPE, Palantir, Splunk, Accenture, and Del. IBM has invested in big data analytics since 2005, and in 2014 it concluded roughly 40,000 data analytics engagements. The IBM Watson Foundation, for instance, has become a key differentiator in the market and provides its customers with the core big data analytics capabilities, leading toward the next era of computing, cognitive or machine learning.[31]

IBM, an information technology company, defines big data through four characteristics, *volume*, *variety*, *velocity* and *veracity*, within the concept of the 'big data era.' This era is defined as a world that is changing through (i) *instrumentation*, in which people are seeing (capturing) more things and storing them (e.g. 'datafication' and 'digitalization'); (ii) *interconnectivity*, which means that people and things are becoming increasingly interconnected through advancements in communication technologies, and also refers to digital communication machine-to-machine; and (iii) *intelligence*, meaning that people are able to add more value by analysing more data (that may even initially seem unrelated) in order to paint a more robust picture of the issue at hand – in other words, by using the right non-traditional data processing methods, people can economically change all low-value data to high-value data.[32]

Another definition of big data was proposed by the Executive Office of The US President in May 2014. They described how most definitions of big data

[29] Jesus Silva and others, 'Privacy Preserving, Protection of Personal Data, and Big Data: A Review of the Colombia Case', *Procedia Computer Science*, 151.2018 (2019), 1213–18 https://doi.org/10.1016/j.procs.2019.04.174

[30] Sian Clancy and others, 'The Role of Personal Commitment to Integrity in Clean Sport and Anti-Doping', *Performance Enhancement and Health*, 10.4 (2022) https://doi.org/10.1016/j.peh.2022.100232

[31] Rahime Belen Saglam, Jason R.C. Nurse, and Duncan Hodges, 'Personal Information: Perceptions, Types and Evolution', *Journal of Information Security and Applications*, 66.March (2022), 103163 https://doi.org/10.1016/j.jisa.2022.103163

[32] Giovanni Rubeis, 'IHealth: The Ethics of Artificial Intelligence and Big Data in Mental Healthcare', *Internet Interventions*, 28.August 2021 (2022), 100518 https://doi.org/10.1016/j.invent.2022.100518

reflect the growing technological ability to capture, aggregate, and process an ever-greater volume, velocity, and variety of data sets. These sets are large, diverse, complex, longitudinal, and/or distributed data sets generated from instruments, sensors, internet transactions, email, video, click streams, and/or all other digital sources available today and in the future.[33] These data sets are colloquially referre to as the '3 Vs' – volume, variety, and velocity. The declining cost of collection, storage, and processing of data, combined with new sources, means that society lives in a world of near-ubiquitous data collection, resulting in an explosion of data that will drive demand for high-performance computing and push the capabilities of even the most sophisticated data management technologies. Despite the enormous amount of data, some data are 'born digital' or 'born analog'; the former emphasises the specific creation of digital data for digital use, and the latter refers to the digitalisation of the physical world. Hence, big data increases the capabilities of 'data fusion.'[34]

In the same vein, the Article 29 Data Protection Working Party of the European Commission, an advisory body of the EU, has given its opinion about the term 'big data.' They say it refers to the exponential growth of data, both in availability and in the automated use of information (or gigantic digital datasets) which are extensively analysed (hence the name 'analytics') using computer algorithms for identifying more general trends and correlations that potentially affect individuals when processing such data.[35] The data sets are large, complex, diverse, longitudinal, and/or distributed and generated from instruments, internet transactions, sensors, video, email, click streams and patterns, and the like. Big data is thus shorthand for the gathering, analysis, processing, and use of immense exploitable datasets, including both structured and unstructured digital information, that has become one of the compelling causes of data-driven businesses and innovations, whose features encompass machine learning, 3D printing, virtual reality, the Internet of Things, cloud computing, and nanotechnology.[36]

---

[33] Deborah Wiltshire and Seraphim Alvanides, 'Ensuring the Ethical Use of Big Data: Lessons from Secure Data Access', *Heliyon*, 8.2 (2022), e08981 https://doi.org/10.1016/j.heliyon.2022.e08981

[34] Petra Perner and Uwe Zscherpel, 'Engineering Applications of Artificial Intelligence: Editorial', *Engineering Applications of Artificial Intelligence*, 15.2 (2002), 121 https://doi.org/10.1016/S0952-1976(02)00027-1

[35] Ahmed Saleh Bataineh and others, 'Toward Monetizing Personal Data: A Two-Sided Market Analysis', *Future Generation Computer Systems*, 111 (2020), 435–59 https://doi.org/10.1016/j.future.2019.11.009

[36] Muharman Lubis and Dini Oktarina D. Handayani, 'The Relationship of Personal Data Protection towards Internet Addiction: Cyber Crimes, Pornography and Reduced Physical Activity', *Procedia Computer Science*, 197.2021 (2021), 151–61 https://doi.org/10.1016/j.procs.2021.12.129

In 2014, Bernard Marr, one of the top business influencers and a distinguished author on the topic of data and analytics for various publications (Forbes, HuffPost, and LinkedIn Pulse), described big data as the '5 Vs', namely volume, velocity, variety, veracity, and value. The first three of these Vs have often been analysed, so it is unnecessary to do so here, but the other Vs should be elaborated. Veracity refers to the trustworthiness of data, meaning that big data analytics allows certain types of data to be handled without losing quality and accuracy (e.g. hash tags, abbreviations, typos, and colloquial speech, as well as the reliability of content). Value may be the most important V of big data, because simply having access to big data is useless unless users can turn it into value.[37]

In December 2016, Bernard Marr added a sixth V, vulnerability, amid growing concerns about personal data processed by many commercial big data initiatives that try to track customer purchasing behaviour and retarget advertising. Most importantly, this is the case when it comes to medical and financial information.[38] Most people experience their data being misused or misplaced in some shape or form, like in cases of data breaches. Today, data processing requires higher assurances for customers that their data is safe (so-called 'data stewarship'), despite the fact that 'zero risk is unattainable' to ensure perfect or unbreakable security for personal data. This is because appropriate technical and organizational measures are far from being silver bullets. As a result, there are regulatory obligations for data breach notifications, data disaster recovery, and transparency, which are also often negotiated in ICT (cloud computing) contracts.[39]

Analytics techniques for big data can be distinguished according to the analytics' value and its maturity. There are two landscapes: *first*, analytics for decision-making outputs that are made on insights from generated data (insights for decisions), and *second*, analytic processes that are automated, namely descriptive, diagnostic, predictive, and prescriptive analytics.[40] Ronald Leenes, full professor in regulation by technology, has argued that analytics techniques, in essence, are designed for the purposes of understanding,

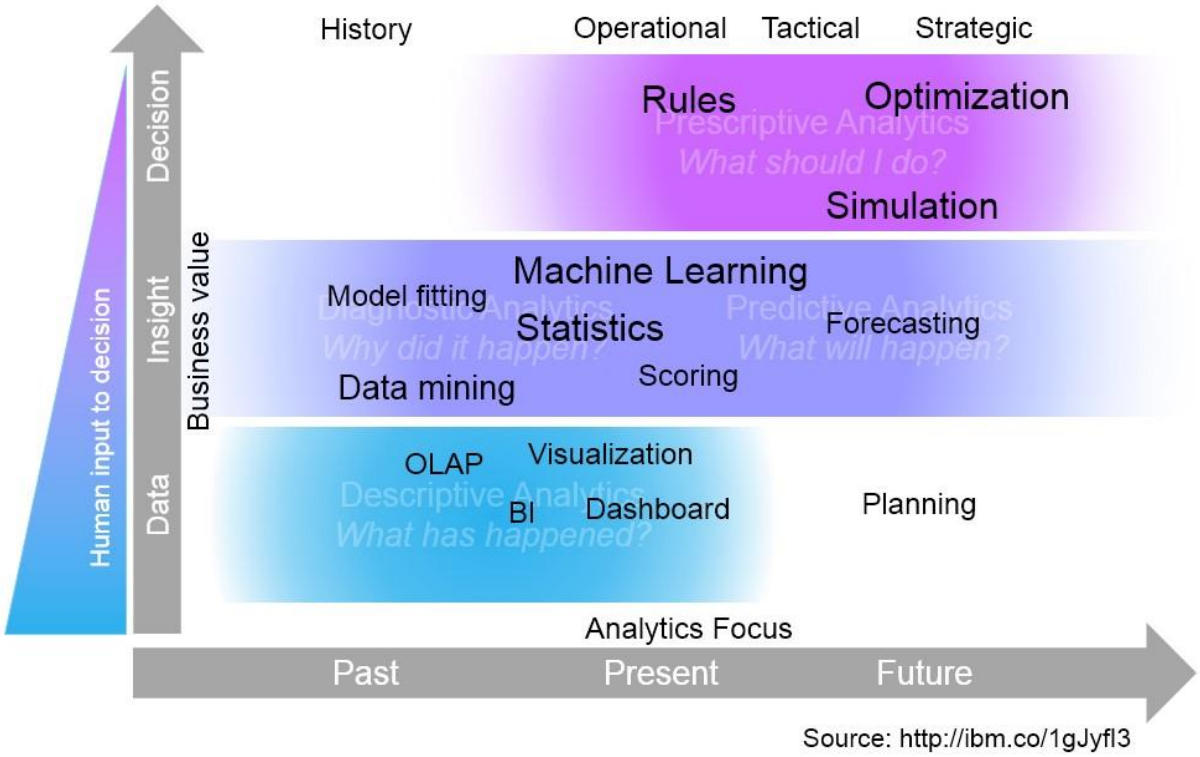
[37] Alfredo Cuzzocrea and others, 'Supporting Privacy-Preserving Big Data Analytics on Temporal Open Big Data', *Procedia Computer Science*, 198.2021 (2021), 112–21 https://doi.org/10.1016/j.procs.2021.12.217

[38] M. Giacalone and others, 'A Novel Big Data Approach for Record and Represent Compliance in the Covid-19 Era', *Big Data Research*, 27 (2022) https://doi.org/10.1016/j.bdr.2021.100290

[39] Fei Feng, Xia Wang, and Tianxiang Chen, 'Analysis of the Attributes of Rights to Inferred Information and China's Choice of Legal Regulation', *Computer Law and Security Review*, 41 (2021) https://doi.org/10.1016/j.clsr.2021.105565

[40] Gabriela Kennedy, 'Asia Pacific News', *Computer Law and Security Review*, 33.6 (2017), 896–904 https://doi.org/10.1016/j.clsr.2017.09.006

predicting (in particular, preferential, consequential, and pre-emptive prediction), and deciding sorting. It is apparent that the 'understanding' configuration of big data is analogous to 'descriptive' analytics, in that it presents the answers to *What happened?* and *What is happening now,* describing the world as it is and providing history and contemporary insights into the world in past and present times (such as in business intelligence, or BI). For example, Google Analytics (analytics.google.com) is an online analytical processing (OLAP) tool that depicts dashboards, sites usages, number of visitors, traffic sources overview, browsers, and so on.[41]



'Diagnostic' analytics breaks down patterns and trends and answers the question *Why did it happen?* based on a statistical model with key variables and relationships amid data (such as market segments, sentiment analysis, price elasticity, fraud patterns, data mining and correlation, data discovery, etc). The diagnostic model serves to determine alert validity in the Internet of Things, particularly in medical devices. Predictive analytics focuses on gaining insights into the future; presents foresights about unknown future events based on the association/correlation between certain problems or failures, according to diagnostics analytics, especially that of citizen/customer behaviour; and acts on that knowledge.[42] Big data predictions may differ from 'profiling' techniques in

[41] Shafiqul Hassan and others, 'Big Data and Predictive Analytics in Healthcare in Bangladesh: Regulatory Challenges', *Heliyon*, 7.6 (2021) https://doi.org/10.1016/j.heliyon.2021.e07179

[42] Roger Clarke, 'Can Small Users Recover from the Cloud?', *Computer Law and Security Review*, 33.6 (2017), 754–67 https://doi.org/10.1016/j.clsr.2017.08.004

the small-data world, which find common associations in the data by drawing a generalisable rule that applies to people in the groups. In big data, it is possible to identify specific individuals, which are so-called granular predictions. For example, a person with an Arabic name may not be subjected to secondary screening at an airport based on profiling, but he or she is probably specified as a terrorist according to other data, as happened in the AOL case, Netflix case, and IMDb, where anonymised personal data was successfully re-identified.[43]

Prescriptive analytics computes operational decisions (actions to be executed now) and tactical or strategic decisions in the short term and long term, respectively, according to the rules of the form *if <condition> then <action>* where 'condition' is set on predicted data to explicitly recommend the best course of subsequent actions, in particular through advanced analytics or machine learning. IBM Watson, for example, is 'cognitive technology', rather than artificial intelligence (AI) or dedicated augmenting human intelligence.[44] There are two types of AI, namely special AI (a system that performs a single task that requires human intelligence well) and general AI (a computer that has the ability to think at or beyond the human level, processing sensory data, making inferences, logical decisions, etc.). The machine learning algorithm has the ability to think at or beyond a human level and is designed to answer questions like *what should I do, what is the best achievable outcome,* and *how can we learn from and address this automatically* on the basis of data sets. In short, AI is broader concept of machine being able to carry out tasks in a way considered 'smart'.[45]

Even though machine learning may differ from special AI, it already has the ability to adapt to new environments or some type of stimuli as input, not just a single 'if <condition>'. This implementation involves the Artificial Neural Network (ANN). Advances in speech recognition, pattern recognition, and image analysis through implementation of the ANN has led to significant progress in the field of intelligence software agents and robotics. IBM Watson, general AI, is a smart data analysis and visualization service that can quickly discover patterns and meaning in data using automated predictive analytics and cognitive capabilities. Outcomes of predictive and prescriptive analytics may be processed further by automating analytics to generate automatic

[43] Zhengzheng Lin and Yanqin Jiang, 'Character Strengths, Meaning in Life, Personal Goal, and Career Adaptability among Impoverished College Students: A Chain-Mediating Model', *Heliyon*, 9.August 2022 (2023), e13232 https://doi.org/10.1016/j.heliyon.2023.e13232

[44] Ingrida Milkaite and others, 'Children's Reflections on Privacy and the Protection of Their Personal Data: A Child-Centric Approach to Data Protection Information Formats', *Children and Youth Services Review*, 129.December 2020 (2021) https://doi.org/10.1016/j.childyouth.2021.106170

[45] Cécile de Terwangne, 'Council of Europe Convention 108+: A Modernised International Treaty for the Protection of Personal Data', *Computer Law and Security Review*, 40.July 2013 (2021), 3–4 https://doi.org/10.1016/j.clsr.2020.105497

decisions and actions, such as in some automated and highly networked systems, including the smart energy grid, automated injections of certain drugs for patients, automated streetlight patterns, and other complex automated networks.[46]

Data processed in a big data platform is different from data processed in traditional data processing: a big data platform performs a (near) real-time processing of (structured and unstructured) information in great volumes and variety that is generated from multiple sources (computers, satellites, mobile devices, cameras, microphones, drones, apps, and many more), in which processing such enormous data challenges current computing technologies. That information must be transmitted, processed, and stored in a volume of petabytes or zettabytes. As an illustration, the total amount of data stored on the entire Internet in 2004 was 1 petabyte (equivalent to 100 years of all television content). By 2020, that number is expected to reach 35 zettabytes (see Figure 2.3). Accordingly, the big data platform is enabling new abilities to analyse this vast amount of data.[47]

Given this, the 6 Vs, as promulgated by Bernard Marr, are attributes of big data. Other data analysts also present other attributes for it, such as exhautibility, rationality, variability, etc. Furthermore, big data analytics allow actors and agencies to generate, use, and repurpose data sets into various contexts by means of proliferating data profiles on a person and social groups (the *polyvalent* feature). Data moves from primary to secondary uses and can be applied to a completely different purpose, not just according to present uses but also all possible ways in the future; this is the so-called 'option value' of data.[48] Big data tools, techniques, and technologies make it possible to predict the likelihood of people's behaviour, such as optimising and limiting options, opportunities, and chances in accessing healthcare, insurance, credit, and employment (the *predictive* attribute). Accordingly, from this preconceived characteristic, it is important to explicate the concept of the legal protections for individuals laid down under data protection and privacy legislation. The above analysis depicts big data in the context of information, technology, and advanced data analytics. This reflects how big data analytics is no longer

[46] Martin Mullins, Christopher P. Holland, and Martin Cunneen, 'Creating Ethics Guidelines for Artificial Intelligence and Big Data Analytics Customers: The Case of the Consumer European Insurance Market', *Patterns*, 2.10 (2021) https://doi.org/10.1016/j.patter.2021.100362

[47] Yong Wan, 'Deep Linking Does Not Constitute a "Making Available to the Public": The Perspective of Beijing Intellectual Property Court', *Computer Law and Security Review*, 33.6 (2017), 876–83 https://doi.org/10.1016/j.clsr.2017.05.013

[48] Nadezhda Purtova, 'Default Entitlements in Personal Data in the Proposed Regulation: Informational Self-Determination off the Table. and Back on Again?', *Computer Law and Security Review*, 30.1 (2014), 6–24 https://doi.org/10.1016/j.clsr.2013.12.006

attributed to traditional data analytics; rather, big data processing will soon require a higher level of data processing beyond what is offered by most data businesses today.[49]

The primary question to be answered here is whether data in the big data context are equal to data processing before the buzzword of 'big data' was introduced. The author assumes that the meaning of 'data' has not changed from the point of view of data protection laws.[50] However, it is important to elaborate the types of data that can be subjected at the starting point to clarify the blurring boundaries between personal and non-personal data. It may be the case that the use of personal medical information by public hospitals for the purpose of medical research is less prone to violating privacy and data protection laws, in contrast to the use of patient data by pharmaceutical firms.[51]

Hence, it is important to give a brief overview of the types of data in data collection and processing. In principle, information is defined in terms of *data* and *meaning*. A clear way of formulating 'information' is as a tripartite definition: information entails *data* that is *well-informed* according to rules (code, language, system) and *meaningful* to the reader. Meanwhile, data may be identical to *signals,* which are what makes possible the coding of *symbols* and the physical implementation of data (analogue, digital, printed on paper, audio, expressed in words or pictures). In general, people may refer to data as the description of something (or reality) that can be recorded, analysed, and reorganised. Such a description involves carriers of knowledge and information; in other words, both information and knowledge (insight) are communicated through data and by means of data storage and transfer devices and systems. Knowledge carried by data relates to its predictive nature, resulting in logical decisions for predictive analytics when incorporated with intelligent software patterns, such as in machine learning.

Information about the past and present provide the basis for the prediction of the future with a degree of certainty. A piece of data only becomes information and knowledge when it is interpreted by its receiver/reader (information and data are used interchangeably in this paper, unless otherwise noted). Therefore, data cannot just be recorded and retrieved, but also requires interpretation to assign meaning to it. For example, mathematics gave new meaning to data

[49] Xiaolan Yu and Yun Zhao, 'Dualism in Data Protection: Balancing the Right to Personal Data and the Data Property Right', *Computer Law and Security Review*, 35.5 (2019), 1–11 https://doi.org/10.1016/j.clsr.2019.04.001

[50] Nick Pantlin, 'European National News', *Computer Law and Security Review*, 33.6 (2017), 892–95 https://doi.org/10.1016/j.clsr.2017.10.002

[51] Kit Burden, 'EU Update', *Computer Law and Security Review*, 33.6 (2017), 884–91 https://doi.org/10.1016/j.clsr.2017.10.001

because calculating (or measuring) a number of something can be understandable without any difficulties and be six times faster when counting with Arabic numerals than with counting boards. In short, data does not entail rights, as the data are clearly not a *living being* or a person that should be afforded protection under data protection laws. However, it is realistic to say that extensive obligations and rights arise *in relation to* data.

Accordingly, data can be categorised in numerous classifications (or types). The interpretation of data categories can be philosophically determined as *relata,* which can be distinghuised according to (i) classification of data, (ii) logical types to which data belong, (iii) the kind of support required for the data implementation or representation, and (iv) the dependence of their semantics in a source or producer. Hence, there are five quite common categories of data, known as typological neutrality: *primary data, secondary data* (constituted by their absence, i.e. silence may be very informative), *metadata* (indications about the nature of some data), *operational data,* and *derivative data,* which is data that can be extracted from some data whenever the latter are used as indirect sources in search of patterns, clues, or inferential evidence about things other than those directly addressed by the data themselves.[52]

However, in the case of personal data in advanced data analytics, a common distinction is among provided, observed, derived, and inferred data. Provided data refers to data that is directly (or voluntarily) provided by an individual with his or her consent, or when he or she is fully aware of the data creation – for example, when a person registers with an online social network ('posted data'), talks to voice computing assistants (i.e. Amazon's Alexa, Apple's Siri, Google Assistant, Microsoft's Contana), applies for a loan ('initiated data'), or purchases goods ('transactional data'). Observed data means data that is gathered from a third party by means of observing and capturing it in a digital format, such as data generated by cookies, sensors (e.g. smartwatch, smartphone), and CCTV cameras. Derived data are produced from existing data using traditional data processing, such as creating annuals of customer value by aggregating all consumer records. Inferred data results from probability-based analytic processes, in the sense of finding connections and relationships between data for revealing unexpected or unknown results; this allows finding for the 'what' without knowing the 'why'. For example, knowing there is a

[52] Nynke E. Vellinga, 'From the Testing to the Deployment of Self-Driving Cars: Legal Challenges to Policymakers on the Road Ahead', *Computer Law and Security Review*, 33.6 (2017), 847–63 https://doi.org/10.1016/j.clsr.2017.05.006

correlation between a low credit rating and having more car accidents cannot reveal why this happens.[53]

Furthermore, data can also be seen as its original description, in particular, data from words, locations, and interactions, through the process of datafication and digitalisation. Datafication means putting a phenomenon or something in a quantified format that allows it to be tabulated and analysed, whereas digitalisation is the process of converting analogue information into the zeros and ones of binary code so computers can handle it.[54] Datafication allows data that has already been collected to deliver services (e.g. Google StreetView), which is opposite to the past view in which data were a by-product of a service. Words can become data; for example, the Google Book project, which began in 2004, makes digital copies of books (scanned and captured digital images – digitalisation) to be text-indexable and thus searchable. This unleashes numerous means by which that data can be used by humans for reading *and* by machines for analysis (machine-to-machine or interoperability and portability).[55]

When a location becomes data, it can be identified, recorded, tallied, analysed, and communicated in a standardised, numerical format, thus allowing companies to collect users' geo-loco data that makes targeted advertising possible based on where the person is situated or is predicted to go. An example is the Foursquare mobile app, which lets people give grades or ratings (a number of 'stars') to their favourite locations. A noncommercial use of geo-loco is a pioneering project in MIT's Human Dynamics Laboratory that is processing the large volume of data from mobile phones to make inferences and predictions about human behaviour. In practice, this so-called 'reality mining' has successfully identified people who had contracted the flu before they realised they were ill.[56]

When interactions become data, it means that datafication transforms intangible elements of people's everyday lives (relationships, experiences, and moods) into data that can be reused for another purpose; for example, Facebook

[53] Mouna Rhahla, Sahar Allegue, and Takoua Abdellatif, 'Guidelines for GDPR Compliance in Big Data Systems', *Journal of Information Security and Applications*, 61.June (2021) https://doi.org/10.1016/j.jisa.2021.102896

[54] Daniel Le Métayer and others, 'Interdisciplinarity in Practice: Challenges and Benefits for Privacy Research', *Computer Law and Security Review*, 33.6 (2017), 864–69 https://doi.org/10.1016/j.clsr.2017.05.020

[55] Mark Giancaspro, 'Is a "Smart Contract" Really a Smart Idea? Insights from a Legal Perspective', *Computer Law and Security Review*, 33.6 (2017), 825–35 https://doi.org/10.1016/j.clsr.2017.05.007

[56] D. Mendelson and D. Mendelson, 'Legal Protections for Personal Health Information in the Age of Big Data – a Proposal for Regulatory Framework', *Ethics, Medicine and Public Health*, 3.1 (2017), 37–55 https://doi.org/10.1016/j.jemep.2017.02.005

datafied relationships (Facebook's 'social graph'), Twitter datafied sentiments, and Microsoft LinkedIn datafied people's professional experiences. The metadata (information about information) of these data posted/created by internet users provide further valuable data, such as a user's language, geo-location, associated friends, each single click on a website (even more, moves of the mouse cursor), devices and apps used, and time for reading or hovering on adversarial content.[57]

From this information, advanced analytics (e.g. sentiment data analytics) may discover groups of likeminded people that help companies receive aggregated customer feedback, adapt better marketing strategies, or judge the impact of marketing campaigns. This analysis shows how digitalisation turbocharges datafication, but they are not substitutes, because the act of digitisation – transforming analogue information into a computer-readable format – by itself does not datafy. Datafication has made it possible to render attitude and sentiments, as well as human behaviour, into an analysable form that can be understood by not only humans but also machines (non-human actors).[58]

Furthermore, big data may be closely related to data reuse in order to unleash the value of big data. The reuse of data seems unlikely to safeguard data protection rights, particularly in large-scale data processing, where the minimum protections for data reuse are often overlooked. Hence, one of the strategies to encourage the value of big data and to balance it with personal data protection at the same time is to distinguish data reuse according to the perspective of data controllers and data subjects, so that the awareness and intention of data processing can be more transparent for data subjects, thereby giving them control over their data.[59]

From the data subjects' perspective, the categorisation of data uses could include data sharing, where the data subject has authorisation rights to permit or forbid disclosure of their data; data portability, involving interoperable uses of data across different systems to avoid data lock-ins; and data that is required to be erased on the basis of the data subject's wishes (related to the right to be forgotten). From the data controllers' perspective, data reuse can be treated

[57] Andria Agesilaou and Eleni A. Kyza, 'Whose Data Are They? Elementary School Students' Conceptualization of Data Ownership and Privacy of Personal Digital Data', *International Journal of Child-Computer Interaction*, 33 (2022) https://doi.org/10.1016/j.ijcci.2022.100462

[58] Richard Steppe, 'Online Price Discrimination and Personal Data: A General Data Protection Regulation Perspective', *Computer Law and Security Review*, 33.6 (2017), 768–85 https://doi.org/10.1016/j.clsr.2017.05.008

[59] Christian Pauletto, 'Options towards a Global Standard for the Protection of Individuals with Regard to the Processing of Personal Data', *Computer Law and Security Review*, 40.3 (2021), 371–81 https://doi.org/10.1016/j.clsr.2020.105433

separately as data recycling (same purposes in several rounds of data processing), data repurposing (processing data for different purposes), and data recontextualisation (interpretating data in a different contexts, e.g. health data interpreted differently by a physician and by health insurance firms). In that regard, data controllers could design their terms and conditions or privacy policies for data subjects so that the interactions between the users and producers of data over data reuse can be tracked or mapped onto the fairness conditions according to the data protection laws.[60]

In that regard, in every phase of big data adoption, by default only personal data which are necessary for each specific purpose are processed, and by design to implement data processing principles. However, advanced data analytics may blur the boundaries of the phases of data processing. In advanced data analytics, rules of data processing are transferred to technology (computer language), known as codes. Even though the advantages of technoregulation include that it is efficient and adaptive and that producers' commitment/involvement can be incorporated into the initial development, this regulatory modality often presents unfavourable outcomes: in the case of large-scale introductions, it often undermines the norms and the reasons behind the production, resulting in violations of individual autonomy or the dignity of persons. For example, even though Facebook's DeepFace provides better ways to learn about users and package their data for advertisers, it has proven controversial, as it can recognise people's names from a high-resolution photograph of a crowd.[61]

In the same vein, real-time data analytics can also easily undermine privacy and data protection. Real-time analytics is defined in microseconds, milliseconds, or seconds, or neartime in seconds. The paramount example of this technology is the Internet of Things. Enabling sensors for the datafication of an event could help to predict the customer's intent, sentiments, and behaviour by mining data and correlations, like the real-time analytics developed by Verizon. To date, there is no international standard of compatibility that exists at the macro levels of the Internet of Things.[62] Hence, not only there is no-interoperability cross-industry for data sharing, but this state of things also

[60] Václav Jane ˇcek, 'Ownership of Personal Data in the Internet of Václav Jane C', *Computer Law & Security Review*, 34.2018 (2020), 1039–52 https://doi.org/https://doi.org/10.1016/j.clsr.2018.04.007

[61] Helena Hansson and Jaap Sok, 'Perceived Obstacles for Business Development: Construct Development and the Impact of Farmers' Personal Values and Personality Profile in the Swedish Agricultural Context', *Journal of Rural Studies*, 81.September 2020 (2021), 17–26 https://doi.org/10.1016/j.jrurstud.2020.12.004

[62] Jay Pil Choi, Doh Shin Jeon, and Byung Cheol Kim, 'Privacy and Personal Data Collection with Information Externalities', *Journal of Public Economics*, 173 (2019), 113–24 https://doi.org/10.1016/j.jpubeco.2019.02.001

leaves open communications from individual devices for unauthorised access and interception. The issue becomes more problematic when the individual devices record private information about an individual, like behaviour patterns and geospatial positions. Even though the data is unstructured, it is still personal data as long as it can identify an individual or if the likelihood of identifiability of an individual is reasonably high, as the time, technology, and costs are conceivable by the users of data.[63]

As has been argued, almost all companies seem to employ big data analytics to gain insights into their customers for purposes such as targeted advertising. This particular purpose logically falls under the scope of the education phase, since information about millions of people is collected for online profiling or behavioural targeting without knowing the data subjects' names.[64] For example, Google claims it 'reaches 90% of internet users worldwide', and Facebook now has over 1.7 billion monthly active users – more than the Chinese population.[114] Today, new ways to track behaviour, habits, and personalities have emerged over the Internet (online tracking). Cookies are the most prominent form of online tracking. Other types of web tracking include server logs, web beacons, and evercookies.[65]

Legal Protection for Personal Data: Indonesian Data Protection Regime

In the EU, the primary rules of processing personal data are governed by Data Protection Directive 95/46/EC. However, this directive will be repealed by a new regulation on data protection, namely the GDPR, on 25 May 2018 (Article 94 of the GDPR). For the sake of argument, this paper will pay more attention to the GDPR rather than the 1995 Directive. However, as side notes for the analysis, some of the directive's provisions will be compared to the new provision in the GDPR. The legal definition of personal data within the (new) GDPR extends the scope of information privacy, which makes data protection rules more complex and more rigid, and thus unrealistic for big data processing. Data protection laws should be simplified, give more attention to the primary underlying principles, and not translate the legal norms directly into technical requirements.[66]

In line with the fundamental rights of the EU, the primary 'celebration' is the choice of the legal instrument, a Regulation in lieu of a Directive, as there will be

[63] Angela Daly, 'Privacy in Automation: An Appraisal of the Emerging Australian Approach', *Computer Law and Security Review*, 33.6 (2017), 836–46 https://doi.org/10.1016/j.clsr.2017.05.009

[64] Daly.

[65] Yuehua Wu, 'Protecting Personal Data in E-Government: A Cross-Country Study', *Government Information Quarterly*, 31.1 (2014), 150–59 https://doi.org/10.1016/j.giq.2013.07.003

[66] Alessandro Mantelero, 'Regulating Big Data. The Guidelines of the Council of Europe in the Context of the European Data Protection Framework', *Computer Law and Security Review*, 33.5 (2017), 584–602 https://doi.org/10.1016/j.clsr.2017.05.011

no issues with inconsistencies and lack of harmonisation. According to the table below, the bundle of data protection novelties includes, inter alia, the right to be forgotten, right to data portability, data protection impact assessments (DPIA), data protection by design and by default, a code of conduct, data protection officers (DPO), and others. The inclusion of the newcomer principle of accountability may reflect the due care that has been taken into account for future data protection requirements. The accountability principle entails at least five elements for its implementation, namely, keeping documentation, appointing DPOs, having DPIAs, developing security measures and having security policies, plus data breach notifications.[67]

At this point, if the rules set out by the GDPR are considered from a different perspective, the GDPR arguably indicates the creation of a whole new industry, complete with its own marketable services, products, and professionals (e.g. DPOs, accreditation bodies for certification and codes of conduct, impact assessments, and all relevant tasks, including new organisations and experts), required for demonstrating compliance properly. Furthermore, it is arguable that data protection laws may be incomplete, as they try to conform to the high pace of technological changes that have made legal solutions ineffective.[68]

The 1945 Constitution of the Republic of Indonesia implicitly expresses the protection of data in creating the right to the protection of personal property for citizens (data can be regarded as property in a broader sense in the Indonesian language) (Article 28G (1)). To date, specific legislation that is equal to an act (national law) with regard to data protection laws has not been adopted in Indonesia. However, the executive branch has proposed a bill for the protection of data and private information; this bill is grouped into Government Priority Legislation 2016 and is being discussed in the parliament. Nevertheless, as a case of electronic data usage, it falls under the subject of electronic information and transactions, which is primarily regulated by 'Electronic Information and Transactions Act No. 11 of 2008' (EIT Act), amended by Act No. 19 of 2016, plus its main delegated legislation: Government Regulation No. 82 of 2012, regarding Provision of Electronic Systems and Transactions (Reg. 82).[69]

---

[67] Wenjie Ding and Li Chen, 'Regulating the Use of Big Data: Justifications, Perspectives and the Chinese Way Forward', *Computer Law and Security Review*, 46.100 (2022) https://doi.org/10.1016/j.clsr.2022.105729

[68] Tao Huang and Yuan Zhao, 'Revolution of Securities Law in the Internet Age: A Review on Equity Crowd-Funding', *Computer Law and Security Review*, 33.6 (2017), 802–10 https://doi.org/10.1016/j.clsr.2017.05.016

[69] Alexander Savelyev, 'Russia's New Personal Data Localization Regulations: A Step Forward or a Self-Imposed Sanction?', *Computer Law and Security Review*, 32.1 (2016), 128–45 https://doi.org/10.1016/j.clsr.2015.12.003

Furthermore, on 7 November 2016, the Indonesian ministry of communications and informatics introduced a new ministry regulation, entitled 'The Protection of Personal Data in the Electronic System' (ministry regulation no. 20 of 2016, hereafter called Reg. 20/2016). The regulation distinguishes between 'personal data' and 'certain individual data.' The Indonesian personal data definition refers to 'certain individual data which is stored, treated, kept accurate and confidential.' Certain individual data legally means 'any information inherently attached to an individual and [that] can be identified, directly or indirectly.' The latter definition can be considered similar to what is called personal data in the EU data protection law. This kind of data should only be processed lawfully and to the extent that the data subject has given consent in a written declaration of the data subject's wishes about the processing of his or her personal data for the explicit purposes for which they are processed (Article 7 (1)). The collection of personal data should also be limited only to the relevant information and be processed accurately. If personal data is already in the public sphere, data processors may use it without the data subject's consent (Article 13).[70]

The definition of personal data is data about an individual which is stored and maintained, the correctness of which is preserved and its confidentiality protected (Reg. 82 and EIT Act). There is no legal definition for the term 'sensitive personal data,' and there is no national data protection authority for data privacy. In spite of that, there are sectoral authorities that have specific competencies in data protection, namely the Financial Service Authority and the Competence Certification Body. Furthermore, there are currently no laws and regulations concerning cookies and location data. Consent is recognised as one legitimate basis for personal data processing, as regulated by either the EIT Act or Reg. 82. [71]

Electronic system providers are obligated to establish a data centre and disaster recovery centre (Article 15 (1) of Reg. 82). Reg. 82 also regulates obligations for electronic system providers, such as providing appropriate protections and ensuring the privacy of personal data, as well as ensuring the appropriate lawful use and disclosure of personal data. According to Regulation 20/2016, privacy is defined as the freedom of data subjects to keep a secret or to disclose their personal data, unless prohibited by law. To make sense of that provision, providers of electronic systems must distinguish what data is personal, and then process it in compliance with relevant data protection laws. Hence, organisational

[70] Chenguo Zhang, '"Sampling" Is Freedom of Art: The German Federal Constitutional Court Deliberates on the Acceptability of Music Sampling in the "Metall Auf Metall" Case', *Computer Law and Security Review*, 33.6 (2017), 870–75 https://doi.org/10.1016/j.clsr.2017.06.005

[71] Sarah Marschlich and Diana Ingenhoff, 'Stakeholder Engagement in a Multicultural Context: The Contribution of (Personal) Relationship Cultivation to Social Capital', *Public Relations Review*, 47.4 (2021) https://doi.org/10.1016/j.pubrev.2021.102091

and technical measures have to be adopted for data processing, in particular for 'personal data', despite the expanding concept of personal data. It is expected that such providers already know that advances in data analytics make all data potentially personal.[72]

The legal definition of personal data may be different from one country to another, and different countries apply different rules governing personal data collection and use. In the UK Data Protection Act 1984, personal data, meant 'data consisting of information which related to a living individual who can be identified from that information (or from that and other information in the possession of the data users), including any expression of opinion about the individual but not any indication of the intention of the data user in respect to that individual.' 'Data' is defined as information recorded in a form in which it can be processed by equipment operating economically in response to instructions given for that purposes.[73]

Hence, that definition distinguishes between information based on the expression of a natural person and explicitly excludes that of the indication of intention of an individual. The identity of a natural person is limited to the expressed information, but not the mental status of a natural person, or even their preference, as long as it is kept as abstract information in mind with respect to that data subject. Currently, in EU law, in particular the GDPR, personal data refers to any information about a natural person when such information can be used to identify, or at least is capable of identifying, that person by direct or indirect means, in particular by reference to an identifier such as a name, location data, an identification number, an online identifier, or any other identity, whether physical, physiological, genetic, mental, economic, cultural, or social (Article 4 (1)).[74]

In contrast, in Indonesia, existing legislation concerning information and electronic transactions (Act No. 11 of 2008) does not provide a legal definition of personal data or its elements; nonetheless, such data is legally safeguarded by Article 26 of that legislation. As is apparent in the bill on the protection of data and personal information proposed by the government (dated 2016), personal data legally means any data on a person's life that is either identified or capable of

[72] Shujie Cui and Peng Qi, 'The Legal Construction of Personal Information Protection and Privacy under the Chinese Civil Code', *Computer Law and Security Review*, 41 (2021), 1–17 https://doi.org/10.1016/j.clsr.2021.105560

[73] Vladislav Arkhipov and Victor Naumov, 'The Legal Definition of Personal Data in the Regulatory Environment of the Russian Federation: Between Formal Certainty and Technological Development', *Computer Law and Security Review*, 32.6 (2016), 868–87 https://doi.org/10.1016/j.clsr.2016.07.009

[74] Daniela Cohen and others, 'The Role of Oxytocin in Implicit Personal Space Regulation: An FMRI Study', *Psychoneuroendocrinology*, 91.February (2018), 206–15 https://doi.org/10.1016/j.psyneuen.2018.02.036

being identified separately or combined with other information, by direct or indirect means and by electronic or non-electronic means (Article 1(1)). Such personal data includes, but is not limited to, a name, passport number, photo or video, telephone number, email, fingerprint, and DNA profile.[75]

Furthermore, in both the EU and Indonesia, legislation recognises special categories of personal data ('sensitive data'), which by its nature may pose a risk to the data subject and thus requires enhanced protections. The GDPR describes sensitive data as personal data revealing political opinions, racial or ethnic origin, religious or philosophical beliefs, and trade union membership. Additionally, processing genetic data, data concerning health or a natural person's sex life or sexual orientation, and biometric data for the purpose of uniquely identifying a natural person is only allowed with specific safeguards in compliance with rules within Article 9 of GDPR, among other conditions: (a) with the explicit consent of the subject; (b) for the purpose of carrying out authorised obligations in the fields of employment, social security, and social protection law; (c) processing only in the course of the legitimate activities of a foundation, association, or any non-profit body that has undisclosed personal data outside that body; (d) in the exercise or defence of legal claims or judicial proceedings; (e) for the purposes of preventive or occupational medicine, medical diagnosis, health or social care/treatment or its management, with reference to necessary reasons of public interest in the area of public health; (f) processing for achieving a purpose in the public interest, statistical purposes, or scientific and historical research purposes, with regard to proportionate measures in light of data protection laws and safeguarding fundamental rights and the interest of the data subjects.[76]

The Indonesian Bill on the Protection of Data and Personal Information defines sensitive data as data related to religious or other beliefs, health, mental and physical condition, sex life, personal financial data, and other data that likely to jeopardise the privacy of the data subject (Article 1 (3)). The processing of sensitive data may be allowed by written agreement in relation to (a) safety protection of the data subject; (b) for exercising rights and obligations according to employment law; (c) for purposes of medical care/treatment; (d) legal enforcement; (e) exercising the functions of lawful authorities; (f) processing sensitive data which is already in the public domain because it was publicly made so by the data subject.

---

[75] Mohammed B. Degnet and others, 'The Role of Personal Values and Personality Traits in Environmental Concern of Non-Industrial Private Forest Owners in Sweden', *Forest Policy and Economics*, 141.May (2022) https://doi.org/10.1016/j.forpol.2022.102767

[76] David Erdos, 'The UK and the EU Personal Data Framework after Brexit: A New Trade and Cooperation Partnership Grounded in Council of Europe Convention 108+?', *Computer Law and Security Review*, 44 (2022), 1–17 https://doi.org/10.1016/j.clsr.2021.105639

Given these definitions, any kind of information can be personal data provided that it relates to a person. In the judgment of the *Amann* case in the European Court of Human Rights (ECtHR), the term 'personal data' was not only limited to matters of the private sphere of an individual, but also to professional life (Para. 65). In the judgment in *Volker and Markus Schecke and Hartmut Eifert v. Land Hessen,* the Court of Justice (CJEU) held that the term 'private life' under the heading 'Right to Respect for Private Life and Family Life' in Article 8 of the European Convention of Human Rights and Fundamental Freedoms (signed in Rome on 4 November 1950), 'shall not be interpreted restrictively and that there is no reason of principle to justify excluding activities of a professional […] nature from the notion of private life' (Para. 59). In addition, according to ECtHR case law, public information can be categorised under the scope of private life when it is systematically collected and stored in files held by authorities.

Moreover, the form and appearance of personal data are not relevant to the applicability of European data protection laws and also the *ex ante* Indonesian legislation. Many types of information may contain personal data, such as information, explanations, communications, statements, opinions, signs, etc which contains a value, a meaning, or a message; things that can be seen, heard, or read; and information presented in various forms and formats through either electronic or non-electronic means. This covers written and spoken communications, images, closed-circuit television (CCTV) footage, sound, electronic information, information on paper, and a sample of human tissue.[77]

Given the legal definition of personal data, the definition of non-personal data can be formulated as data that may refer to any information relating to a non-natural person when such information does not convey any identification of a natural person, whether directly or indirectly, such as the general confidential information of businesses, statistical data, intellectual property assets (e.g. standard essential patents and trade secrets), security information, and so forth. Non-personal data also refers to anonymous information/data, namely information which does not relate to an identified or identifiable natural person, or personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. In other word, anonymisation means excluding any personal identifiers from data sets.[78]

[77] Asli Deniz Helvacioglu and Hanna Stakheyeva, 'The Tale of Two Data Protection Regimes: The Analysis of the Recent Law Reform in Turkey in the Light of EU Novelties', *Computer Law and Security Review*, 33.6 (2017), 811–24 https://doi.org/10.1016/j.clsr.2017.05.014

[78] Bingbin Lu, 'The Unique Chinese Legal Approach to Online Ad Blocking: Is It in the Right Direction?', *Computer Law and Security Review*, 33.6 (2017), 786–801 https://doi.org/10.1016/j.clsr.2017.05.012

Not all data processed by big data analytics contain personal information. Machine data (e.g. data from sensors installed in refineries and data on airport weather), for example, do not contain personal information. This data is also known as operational data, which is often landlocked in operational systems like information from manufacturing equipment, industrial machineries, or chemical boilers. Thus, big data analyses of such data do not practically pose a risk to privacy or processing personal data.[79]

GDPR states that a disproportionate effort to identify a natural person may not lead to personal data when the identification processes requires unreasonable costs and time with reference to available technological resources and development. Accordingly, there may be circumstances where an individual is 'easily identifiable' and data protection laws apply. However, with indirectly identifiable data, such as from IP address, cookies, and other online scrutiny technologies, it is difficult to ascertain whether that is personal data or not, as determined solely by the proportionate or disproportionate efforts test. However, nonpersonal data can become personal data depending on its use and retention, and the accumulation of non-personal information may possibly give rise to obtaining information on a natural person.[80]

Impact on Privacy and Data Protection in Indonesia

Outside form of big data is data gathering or tracking by means of cookies and any other web tracking apps and analytics (e.g. Google Analytics). Wealthy countries such the EU and the US initially proposed how to tackle privacy and data protection issues. Private entities have also started new business models for the techno-regulation of privacy-friendly solutions, such as Ghostery™, which allows complete control of online browsing free from trackers, and Brave Internet Browser, which automatically blocks ads and trackers. As privacy breaches were one of the first visible problems to arise at unprecedented level due to massive (personal and non-personal) data processing in both centralised and decentralised analytics servers in real time. The direct solution in response to these challenges is to strengthen data subjects' control of their personal data, taking into account the free movement of data for a fair economics market and crime prevention. Hence, commentators often suggest the standard solution of consent by means of notice

[79] Jaqueline de Godoy, Kathrin Otrel-Cass, and Kristian Høyer Toft, 'Transformations of Trust in Society: A Systematic Review of How Access to Big Data in Energy Systems Challenges Scandinavian Culture', *Energy and AI*, 5.April (2021) https://doi.org/10.1016/j.egyai.2021.100079

[80] Zhiheng Chen and others, 'Using Mobile Phone Big Data to Identify Inequity of Artificial Light at Night Exposure: A Case Study in Tokyo', *Cities*, 128.May (2022), 3–5 https://doi.org/10.1016/j.cities.2022.103803

and choice and using specifications and limits, data minimisation, and transparency to overcome potential privacy problems brought up by big data.[81]

Scholars have identified potential impacts of big data technologies on the core of data subjects' right to privacy (private life) and data protection. Processing personal data within big data platforms may be exempt from liability under data protection regulations if the data being processed is qualified as non-personal data. Nonetheless, important considerations have to be thoroughly examined, namely (a) the possibility of an incorrect claim that data is qualified as 'personal' and (b) that processing non-personal data may also constitute an unlawful infringement of the right to respect for private life. During the GDPR's negotiations in the European Parliament, an official reporter outlined issues up for negotiation, such as (i) the data subject must be informed about what happens to their data, and they must be able to continuously grant their consent whether to accept or reject to data processing; (ii) all information that may directly or indirectly pertain to a person is defined as personal data and thus needs to be protected as such; and (iii) data protection by design and by default should to be encouraged, including data minimisation and data protection-friendly pre-settings.[82]

Until a certain point, data are not considered personal data (when they are not connected to any identified or identifiable individual), but might later be rediscovered by big data analytics to be linked to an actual, precise person. Accordingly, it is necessary to fully understand (e.g. scopes and limits) identifiability and the legal concept and definition of 'personal data' in the big data era. As noted by the EDPS, legal protections formulated by laws (e.g. GDPR) are not only supposed to safeguard fundamental rights (e.g. the right to the protection of personal data, right of access to data, and right to rectify it) but also go beyond this, when data-driven technologies progressively converge with artificial intelligence.[83]

Certain challenges related to this must be given particular attention, including, transparency of the data controllers. Data controllers must legally provide to data subjects information about the processing of their data (including repurposing data processing, or what happens to their data) and their (digital) rights laid down by law (especially according to the GDPR), and this information must be presented in a clear manner. Furthermore, other protection mechanisms may be sought in anticipating unpredicted data processing practices, such as when it is

[81] Li and Saxunová.
[82] Valero and others.
[83] Oksas and others.

incorporated into cloud computing technologies and smart devices (e.g. wearable devices).[84]

Balancing the interest of data subjects against public interests and legitimate interests pursued by the controllers or by a third party. In a fair and proportionate way, the possibility of a legitimate basis for data processing must not be over-stretched at a certain level vis-à-vis any possible third-party interests. One mechanism to create a proportional and fair balance is to encourage the data minimisation concept in the big data platform. Consent in the big data platform might be problematic because the purpose limitation does not fit with big data analytics, like in the case of machine learning. From a technological and economic standpoint, it might be a bit challenging and expensive to ask for consent again for global or multilevel data processing with millions of data subjects in diverse jurisdictions and with heterogeneous interests – for example, asking data subjects' consent for transforming printed books to digital books and/or future forms of expression that are compatible with state-of-the-art smart devices. For this, the data subject is often mediated or delegated by these instruments and applications.[85]

In practice, people do not seriously spend time and intellectual effort reading privacy terms, even if they have been presented in a short, layered/structured, and clear message. There is a trade-off that cannot be ignored between practical and meaningful consent. Do data subjects truly know what they are consenting to? The answer seems sceptical these days. Special categories of data ('sensitive data'). Currently, many apps and devices gather location data and other sensitive data (e.g. relationship with a person and relatives that can usually be used as an alternative password to log into digital information services).[86]

It is important to note that non-sensitive data processing may lead to revealing sensitive and personal information when employing data mining. Automated processing. Data protection laws set rules prohibiting automated processing, unless in exceptional conditions or with specific safeguards. For example, article 11 of the GDPR states that automated individual decision making is not allowed unless, in essence, it is authorised by law, providing an appropriate safeguard for the rights and freedoms of the data subjects, and human shall interfere the process. This provision arguably only focuses on the final moment when there is such automated processing that 'significantly' affect the data subjects. The unaddressed issue is the ways in which data about a person is used to make decisions about

[84] van den Broek and van Veenstra.
[85] Chua, Ooi, and Herbland.
[86] Shen and others.

another person, and the manners in which the decision techniques are reached. Arguably, further transparency is required for such pending issues.[87]

Data protection laws achieve two objectives: (1) to facilitate the free flow of data and (2) to provide minimum protections for personal data. The prior objective is likely anchored in the achievement of an internal market, and the rationale of the latter objective is to safeguard the protection of the fundamental rights and freedoms of individuals. The analysis of data protection and big data shows the challenges and (potential) implications for issues such as the validity of consent, purpose limitation principle, anonymisation and identifiable personal data, transformative use of data, profiling developments, legitimate interests, and the enforcement of data protection rights.

Hence, the central scope of data protection laws involves protecting personal data. However, in the context of big data, the nature of data means it is not easy to evaluate whether scattered data is considered personal information (private and professional life), since it still reveals a name of an individual, as long as a natural person is distinguished from other individuals (identity and identifiability) on the basis of such data. In that regard, the likelihood of data being identifiable for individuals should be addressed in order to illuminate the nature of data, in particular whether the data are personal, as this is crucial for the purpose of data protection.

4. Conclusion

Data protection laws provide minimum protections for personal data, as well as facilitate the free flow of such data, by setting out principles and rules for legitimate data processing. In the big data context, personal data may not be as easy to distinguish as in traditional data processing, and that makes policy-makers and businesses turn to the identifiability concept: in other words, what data are personal. To understand such paramount terminology in data protection law, relevant factors are presented to assess the direct or indirect identification of a natural person. As has been discussed, the assessment constitutes the legality test, which in essence explicitly excludes illegal means for gathering additional and identifying information, and then evaluates using the 'likely reasonable test', which takes into account costs in a broad sense for means of identification. In the EU data protection law, the test entails, for example, risk-based measures and technological development, whereas Indonesian law on data protection has not yet established such assessments. Data within big data operations traditionally falls under the scope of data protection laws only if it discloses the private life of individuals, such as names or other civil identities, but without further conditions to ascertain whether the data can be indirectly identified with an individual.

[87] Li and Saxunová.

Accordingly, for that particular developing country, the forthcoming benefits of big data and the legal definition of personal data, with regard to the indirect identification of a natural person, have not been formulated in the current national law that covers data protection law.

The boundaries of personal data set out by the laws, however, likely do not suffice as a good instrument for determining what data are personal because personal information includes not only the expressed identity (identified identity) but also the state of mind, interest, and mental abilities of a natural person (identifiable identity). Therefore, the relevant soft values and open norms with regard to the identifiability concept should be rooted in the foundations and goals of data protection laws, particular safeguarding data subjects' human dignity. The soft values and open norms under the concept of identifiability should be clarified by providing detailed guidelines with artificial examples or possible scenarios that may occur in the advanced data analytics of the big data era. These guidelines should be formulated by supervisory authorities to at least make sure that sensible data protection rules and principles are fully applied in the ongoing technological development of data analytics. Henceforth, philosophically speaking, data subjects' human dignity rests in being master of their own journeys and keeping their identities and choices open. Any regulatory approach to defining personal data should be framed under that consideration, as a prescriptive guidance to ascertain the defensible answer, in the future, for questions like whether this (sub) kind of data is an integral part of data subject's human dignity, not merely narrowing or expanding the legal boundaries of personal data.

References

Acciarini, Chiara, Francesco Cappa, Paolo Boccardelli, and Raffaele Oriani, 'How Can Organizations Leverage Big Data to Innovate Their Business Models? A Systematic Literature Review', *Technovation*, In Press.May 2022 (2023), 102713 https://doi.org/10.1016/j.technovation.2023.102713

Agesilaou, Andria, and Eleni A. Kyza, 'Whose Data Are They? Elementary School Students' Conceptualization of Data Ownership and Privacy of Personal Digital Data', *International Journal of Child-Computer Interaction*, 33 (2022) https://doi.org/10.1016/j.ijcci.2022.100462

Antusch, S., R. Custers, H. Marien, and H. Aarts, 'Intentional Action and Limitation of Personal Autonomy. Do Restrictions of Action Selection Decrease the Sense of Agency?', *Consciousness and Cognition*, 88.January (2021), 103076 https://doi.org/10.1016/j.concog.2021.103076

Arkhipov, Vladislav, and Victor Naumov, 'The Legal Definition of Personal Data in the Regulatory Environment of the Russian Federation: Between Formal

Certainty and Technological Development', *Computer Law and Security Review*, 32.6 (2016), 868–87 https://doi.org/10.1016/j.clsr.2016.07.009

Bataineh, Ahmed Saleh, Rabeb Mizouni, Jamal Bentahar, and May El Barachi, 'Toward Monetizing Personal Data: A Two-Sided Market Analysis', *Future Generation Computer Systems*, 111 (2020), 435–59 https://doi.org/10.1016/j.future.2019.11.009

Belen Saglam, Rahime, Jason R.C. Nurse, and Duncan Hodges, 'Personal Information: Perceptions, Types and Evolution', *Journal of Information Security and Applications*, 66.March (2022), 103163 https://doi.org/10.1016/j.jisa.2022.103163

Bolognini, Luca, and Camilla Bistolfi, 'Pseudonymization and Impacts of Big (Personal/Anonymous) Data Processing in the Transition from the Directive 95/46/EC to the New EU General Data Protection Regulation', *Computer Law and Security Review*, 33.2 (2017), 171–81 https://doi.org/10.1016/j.clsr.2016.11.002

van den Broek, Tijs, and Anne Fleur van Veenstra, 'Governance of Big Data Collaborations: How to Balance Regulatory Compliance and Disruptive Innovation', *Technological Forecasting and Social Change*, 129.September 2017 (2018), 330–38 https://doi.org/10.1016/j.techfore.2017.09.040

Burden, Kit, 'EU Update', *Computer Law and Security Review*, 33.6 (2017), 884–91 https://doi.org/10.1016/j.clsr.2017.10.001

Chen, Zhiheng, Peiran Li, Yanxiu Jin, Yuan Jin, Jinyu Chen, Wenjing Li, and others, 'Using Mobile Phone Big Data to Identify Inequity of Artificial Light at Night Exposure: A Case Study in Tokyo', *Cities*, 128.May (2022), 3–5 https://doi.org/10.1016/j.cities.2022.103803

Choi, Jay Pil, Doh Shin Jeon, and Byung Cheol Kim, 'Privacy and Personal Data Collection with Information Externalities', *Journal of Public Economics*, 173 (2019), 113–24 https://doi.org/10.1016/j.jpubeco.2019.02.001

Chua, Hui Na, Jie Sheng Ooi, and Anthony Herbland, 'The Effects of Different Personal Data Categories on Information Privacy Concern and Disclosure', *Computers and Security*, 110 (2021), 102453 https://doi.org/10.1016/j.cose.2021.102453

Clancy, Sian, Frank Owusu-Sekyere, Jake Shelley, Annalena Veltmaat, Alessandra De Maria, and Andrea Petróczi, 'The Role of Personal Commitment to Integrity in Clean Sport and Anti-Doping', *Performance Enhancement and Health*, 10.4 (2022) https://doi.org/10.1016/j.peh.2022.100232

Clarke, Roger, 'Can Small Users Recover from the Cloud?', *Computer Law and Security Review*, 33.6 (2017), 754–67 https://doi.org/10.1016/j.clsr.2017.08.004

Cohen, Daniela, Anat Perry, Naama Mayseless, Oded Kleinmintz, and Simone G. Shamay-Tsoory, 'The Role of Oxytocin in Implicit Personal Space Regulation: An FMRI Study', *Psychoneuroendocrinology*, 91.February (2018), 206–15 https://doi.org/10.1016/j.psyneuen.2018.02.036

Cui, Shujie, and Peng Qi, 'The Legal Construction of Personal Information Protection and Privacy under the Chinese Civil Code', *Computer Law and Security Review*, 41 (2021), 1–17 https://doi.org/10.1016/j.clsr.2021.105560

Custers, Bart, and Gianclaudio Malgieri, 'Priceless Data: Why the EU Fundamental Right to Data Protection Is at Odds with Trade in Personal Data', *Computer Law and Security Review*, 45 (2022), 105683 https://doi.org/10.1016/j.clsr.2022.105683

Cuzzocrea, Alfredo, Carson K. Leung, Anifat M. Olawoyin, and Edoardo Fadda, 'Supporting Privacy-Preserving Big Data Analytics on Temporal Open Big Data', *Procedia Computer Science*, 198.2021 (2021), 112–21 https://doi.org/10.1016/j.procs.2021.12.217

Daly, Angela, 'Privacy in Automation: An Appraisal of the Emerging Australian Approach', *Computer Law and Security Review*, 33.6 (2017), 836–46 https://doi.org/10.1016/j.clsr.2017.05.009

Degnet, Mohammed B., Helena Hansson, Marjanke A. Hoogstra-Klein, and Anders Roos, 'The Role of Personal Values and Personality Traits in Environmental Concern of Non-Industrial Private Forest Owners in Sweden', *Forest Policy and Economics*, 141.May (2022) https://doi.org/10.1016/j.forpol.2022.102767

Ding, Wenjie, and Li Chen, 'Regulating the Use of Big Data: Justifications, Perspectives and the Chinese Way Forward', *Computer Law and Security Review*, 46.100 (2022) https://doi.org/10.1016/j.clsr.2022.105729

Elliot, Mark, Kieron O'Hara, Charles Raab, Christine M. O'Keefe, Elaine Mackey, Chris Dibben, and others, 'Functional Anonymisation: Personal Data and the Data Environment', *Computer Law and Security Review*, 34.2 (2018), 204–21 https://doi.org/10.1016/j.clsr.2018.02.001

Erdos, David, 'The UK and the EU Personal Data Framework after Brexit: A New Trade and Cooperation Partnership Grounded in Council of Europe Convention 108+?', *Computer Law and Security Review*, 44 (2022), 1–17 https://doi.org/10.1016/j.clsr.2021.105639

Feng, Fei, Xia Wang, and Tianxiang Chen, 'Analysis of the Attributes of Rights to Inferred Information and China's Choice of Legal Regulation', *Computer Law and Security Review*, 41 (2021) https://doi.org/10.1016/j.clsr.2021.105565

Georgiadis, Georgios, and Geert Poels, 'Towards a Privacy Impact Assessment

Methodology to Support the Requirements of the General Data Protection Regulation in a Big Data Analytics Context: A Systematic Literature Review', *Computer Law and Security Review*, 44 (2022) https://doi.org/10.1016/j.clsr.2021.105640

Giacalone, M., D. C. Sinitò, M. V. Calciano, and V. Santarcangelo, 'A Novel Big Data Approach for Record and Represent Compliance in the Covid-19 Era', *Big Data Research*, 27 (2022) https://doi.org/10.1016/j.bdr.2021.100290

Giancaspro, Mark, 'Is a "Smart Contract" Really a Smart Idea? Insights from a Legal Perspective', *Computer Law and Security Review*, 33.6 (2017), 825–35 https://doi.org/10.1016/j.clsr.2017.05.007

Godoy, Jaqueline de, Kathrin Otrel-Cass, and Kristian Høyer Toft, 'Transformations of Trust in Society: A Systematic Review of How Access to Big Data in Energy Systems Challenges Scandinavian Culture', *Energy and AI*, 5.April (2021) https://doi.org/10.1016/j.egyai.2021.100079

Hansson, Helena, and Jaap Sok, 'Perceived Obstacles for Business Development: Construct Development and the Impact of Farmers' Personal Values and Personality Profile in the Swedish Agricultural Context', *Journal of Rural Studies*, 81.September 2020 (2021), 17–26 https://doi.org/10.1016/j.jrurstud.2020.12.004

Hasanzadeh, Kamyar, Anna Kajosaari, Dan Häggman, and Marketta Kyttä, 'A Context Sensitive Approach to Anonymizing Public Participation GIS Data: From Development to the Assessment of Anonymization Effects on Data Quality', *Computers, Environment and Urban Systems*, 83.April (2020), 101513 https://doi.org/10.1016/j.compenvurbsys.2020.101513

Hassan, Shafiqul, Mohsin Dhali, Fazluz Zaman, and Muhammad Tanveer, 'Big Data and Predictive Analytics in Healthcare in Bangladesh: Regulatory Challenges', *Heliyon*, 7.6 (2021) https://doi.org/10.1016/j.heliyon.2021.e07179

Helvacioglu, Asli Deniz, and Hanna Stakheyeva, 'The Tale of Two Data Protection Regimes: The Analysis of the Recent Law Reform in Turkey in the Light of EU Novelties', *Computer Law and Security Review*, 33.6 (2017), 811–24 https://doi.org/10.1016/j.clsr.2017.05.014

Huang, Tao, and Yuan Zhao, 'Revolution of Securities Law in the Internet Age: A Review on Equity Crowd-Funding', *Computer Law and Security Review*, 33.6 (2017), 802–10 https://doi.org/10.1016/j.clsr.2017.05.016

Jiang, Jinglin, Li Liao, Xi Lu, Zhengwei Wang, and Hongyu Xiang, 'Deciphering Big Data in Consumer Credit Evaluation', *Journal of Empirical Finance*, 62.August 2020 (2021), 28–45 https://doi.org/10.1016/j.jempfin.2021.01.009

Karjalainen, Tuulia, 'The Battle of Power: Enforcing Data Protection Law against

Companies Holding Data Power', *Computer Law and Security Review*, 47.August 2018 (2022), 105742 https://doi.org/10.1016/j.clsr.2022.105742

Kennedy, Gabriela, 'Asia Pacific News', *Computer Law and Security Review*, 33.6 (2017), 896–904 https://doi.org/10.1016/j.clsr.2017.09.006

Li, Yuanxin, and Darina Saxunová, 'A Perspective on Categorizing Personal and Sensitive Data and the Analysis of Practical Protection Regulations', *Procedia Computer Science*, 170 (2020), 1110–15 https://doi.org/10.1016/j.procs.2020.03.060

Lin, Zhengzheng, and Yanqin Jiang, 'Character Strengths, Meaning in Life, Personal Goal, and Career Adaptability among Impoverished College Students: A Chain-Mediating Model', *Heliyon*, 9.August 2022 (2023), e13232 https://doi.org/10.1016/j.heliyon.2023.e13232

Liu, Cheng yong, Ling Jan Chiou, Cheng chung Li, and Xiu Wen Ye, 'Analysis of Beijing Tianjin Hebei Regional Credit System from the Perspective of Big Data Credit Reporting', *Journal of Visual Communication and Image Representation*, 59 (2019), 300–308 https://doi.org/10.1016/j.jvcir.2019.01.018

Liu, Tianqi, Chukwunonso O. Aniagor, Marcel I. Ejimofor, Matthew C. Menkiti, Kuok Ho Daniel Tang, Bridgid Lai Fui Chin, and others, 'Technologies for Removing Pharmaceuticals and Personal Care Products (PPCPs) from Aqueous Solutions: Recent Advances, Performances, Challenges and Recommendations for Improvements', *Journal of Molecular Liquids*, 374 (2022), 121144 https://doi.org/10.1016/j.molliq.2022.121144

Liu, Yu li, Luyan Huang, Wenjia Yan, Xinghan Wang, and Ruochen Zhang, 'Privacy in AI and the IoT: The Privacy Concerns of Smart Speaker Users and the Personal Information Protection Law in China', *Telecommunications Policy*, 46.7 (2022), 102334 https://doi.org/10.1016/j.telpol.2022.102334

Liu, Yue, 'User Control of Personal Information Concerning Mobile-App: Notice and Consent?', *Computer Law and Security Review*, 30.5 (2014), 521–29 https://doi.org/10.1016/j.clsr.2014.07.008

LIU, Zhao ge, Xiang yang LI, and Xiao han ZHU, 'Scenario Modeling for Government Big Data Governance Decision-Making: Chinese Experience with Public Safety Services', *Information and Management*, 59.3 (2022), 103622 https://doi.org/10.1016/j.im.2022.103622

Lu, Bingbin, 'The Unique Chinese Legal Approach to Online Ad Blocking: Is It in the Right Direction?', *Computer Law and Security Review*, 33.6 (2017), 786–801 https://doi.org/10.1016/j.clsr.2017.05.012

Lubis, Muharman, and Dini Oktarina D. Handayani, 'The Relationship of Personal Data Protection towards Internet Addiction: Cyber Crimes, Pornography and

Reduced Physical Activity', *Procedia Computer Science*, 197.2021 (2021), 151–61 https://doi.org/10.1016/j.procs.2021.12.129

Malgieri, Gianclaudio, and Bart Custers, 'Pricing Privacy – the Right to Know the Value of Your Personal Data', *Computer Law and Security Review*, 34.2 (2018), 289–303 https://doi.org/10.1016/j.clsr.2017.08.006

Mantelero, Alessandro, 'Regulating Big Data. The Guidelines of the Council of Europe in the Context of the European Data Protection Framework', *Computer Law and Security Review*, 33.5 (2017), 584–602 https://doi.org/10.1016/j.clsr.2017.05.011

Marinova, Galia, Aida Bitri, and Marsida Ibro, 'The Role of Women in the Digital Age: Balancing Professional and Personal Challenges during the COVID-19 Pandemic', *IFAC-PapersOnLine*, 54.13 (2021), 539–44 https://doi.org/10.1016/j.ifacol.2021.10.505

Marschlich, Sarah, and Diana Ingenhoff, 'Stakeholder Engagement in a Multicultural Context: The Contribution of (Personal) Relationship Cultivation to Social Capital', *Public Relations Review*, 47.4 (2021) https://doi.org/10.1016/j.pubrev.2021.102091

Mendelson, D., and D. Mendelson, 'Legal Protections for Personal Health Information in the Age of Big Data – a Proposal for Regulatory Framework', *Ethics, Medicine and Public Health*, 3.1 (2017), 37–55 https://doi.org/10.1016/j.jemep.2017.02.005

Métayer, Daniel Le, Mathias Bossuet, Fanny Coudert, Claire Gayrel, Francisco Jaime, Christophe Jouvray, and others, 'Interdisciplinarity in Practice: Challenges and Benefits for Privacy Research', *Computer Law and Security Review*, 33.6 (2017), 864–69 https://doi.org/10.1016/j.clsr.2017.05.020

Milkaite, Ingrida, Ralf De Wolf, Eva Lievens, Tom De Leyn, and Marijn Martens, 'Children's Reflections on Privacy and the Protection of Their Personal Data: A Child-Centric Approach to Data Protection Information Formats', *Children and Youth Services Review*, 129.December 2020 (2021) https://doi.org/10.1016/j.childyouth.2021.106170

Mullins, Martin, Christopher P. Holland, and Martin Cunneen, 'Creating Ethics Guidelines for Artificial Intelligence and Big Data Analytics Customers: The Case of the Consumer European Insurance Market', *Patterns*, 2.10 (2021) https://doi.org/10.1016/j.patter.2021.100362

Oksas, Catherine, Julia Green Brody, Phil Brown, Katherine E. Boronow, Erin DeMicco, Annemarie Charlesworth, and others, 'Perspectives of Peripartum People on Opportunities for Personal and Collective Action to Reduce Exposure to Everyday Chemicals: Focus Groups to Inform Exposure Report-

Back', *Environmental Research*, 212.December 2021 (2022) https://doi.org/10.1016/j.envres.2022.113173

Pantlin, Nick, 'European National News', *Computer Law and Security Review*, 33.6 (2017), 892–95 https://doi.org/10.1016/j.clsr.2017.10.002

Papakonstantinou, Vagelis, and Paul de Hert, 'Big Data Analytics in Electronic Communications: A Reality in Need of Granular Regulation (Even If This Includes an Interim Period of No Regulation at All)', *Computer Law and Security Review*, 36.November 2015 (2020), 105397 https://doi.org/10.1016/j.clsr.2020.105397

Pauletto, Christian, 'Options towards a Global Standard for the Protection of Individuals with Regard to the Processing of Personal Data', *Computer Law and Security Review*, 40.3 (2021), 371–81 https://doi.org/10.1016/j.clsr.2020.105433

Perner, Petra, and Uwe Zscherpel, 'Engineering Applications of Artificial Intelligence: Editorial', *Engineering Applications of Artificial Intelligence*, 15.2 (2002), 121 https://doi.org/10.1016/S0952-1976(02)00027-1

Purtova, Nadezhda, 'Default Entitlements in Personal Data in the Proposed Regulation: Informational Self-Determination off the Table. and Back on Again?', *Computer Law and Security Review*, 30.1 (2014), 6–24 https://doi.org/10.1016/j.clsr.2013.12.006

Rhahla, Mouna, Sahar Allegue, and Takoua Abdellatif, 'Guidelines for GDPR Compliance in Big Data Systems', *Journal of Information Security and Applications*, 61.June (2021) https://doi.org/10.1016/j.jisa.2021.102896

Rubeis, Giovanni, 'IHealth: The Ethics of Artificial Intelligence and Big Data in Mental Healthcare', *Internet Interventions*, 28.August 2021 (2022), 100518 https://doi.org/10.1016/j.invent.2022.100518

Savelyev, Alexander, 'Russia's New Personal Data Localization Regulations: A Step Forward or a Self-Imposed Sanction?', *Computer Law and Security Review*, 32.1 (2016), 128–45 https://doi.org/10.1016/j.clsr.2015.12.003

Shen, Yuncheng, Bing Guo, Yan Shen, Xuliang Duan, Xiangqian Dong, Hong Zhang, and others, 'Personal Big Data Pricing Method Based on Differential Privacy', *Computers and Security*, 113 (2022), 102529 https://doi.org/10.1016/j.cose.2021.102529

Silva, Jesus, Darwin Solano, Claudia Fernandez, Ligia Romero, and Jesus Vargas Villa, 'Privacy Preserving, Protection of Personal Data, and Big Data: A Review of the Colombia Case', *Procedia Computer Science*, 151.2018 (2019), 1213–18 https://doi.org/10.1016/j.procs.2019.04.174

Steppe, Richard, 'Online Price Discrimination and Personal Data: A General Data Protection Regulation Perspective', *Computer Law and Security Review*, 33.6 (2017), 768–85 https://doi.org/10.1016/j.clsr.2017.05.008

de Terwangne, Cécile, 'Council of Europe Convention 108+: A Modernised International Treaty for the Protection of Personal Data', *Computer Law and Security Review*, 40.July 2013 (2021), 3–4 https://doi.org/10.1016/j.clsr.2020.105497

Václav Jane ˇcek, 'Ownership of Personal Data in the Internet of Václav Jane C', *Computer Law & Security Review*, 34.2018 (2020), 1039–52 https://doi.org/https://doi.org/10.1016/j.clsr.2018.04.007

Valero, Cayetano, Jaime Pérez, Sonia Solera-cotanilla, Mario Vega-barbas, Guillermo Suarez-tangil, Manuel Alvarez-campana, and others, 'Analysis of Security and Data Control in Smart Personal Assistants from the User's Perspective', *Future Generation Computer Systems*, 2023 https://doi.org/10.1016/j.future.2023.02.009

Vellinga, Nynke E., 'From the Testing to the Deployment of Self-Driving Cars: Legal Challenges to Policymakers on the Road Ahead', *Computer Law and Security Review*, 33.6 (2017), 847–63 https://doi.org/10.1016/j.clsr.2017.05.006

Wan, Yong, 'Deep Linking Does Not Constitute a "Making Available to the Public": The Perspective of Beijing Intellectual Property Court', *Computer Law and Security Review*, 33.6 (2017), 876–83 https://doi.org/10.1016/j.clsr.2017.05.013

Warso, Zuzanna, 'There's More to It than Data Protection-Fundamental Rights, Privacy and the Personal/Household Exemption in the Digital Age', *Computer Law and Security Review*, 29.5 (2013), 491–500 https://doi.org/10.1016/j.clsr.2013.07.002

Wiltshire, Deborah, and Seraphim Alvanides, 'Ensuring the Ethical Use of Big Data: Lessons from Secure Data Access', *Heliyon*, 8.2 (2022), e08981 https://doi.org/10.1016/j.heliyon.2022.e08981

Wu, Yuehua, 'Protecting Personal Data in E-Government: A Cross-Country Study', *Government Information Quarterly*, 31.1 (2014), 150–59 https://doi.org/10.1016/j.giq.2013.07.003

Yu, Xiaolan, and Yun Zhao, 'Dualism in Data Protection: Balancing the Right to Personal Data and the Data Property Right', *Computer Law and Security Review*, 35.5 (2019), 1–11 https://doi.org/10.1016/j.clsr.2019.04.001

Yuvaraj, Joshua, 'How about Me? The Scope of Personal Information under the Australian Privacy Act 1988', *Computer Law and Security Review*, 34.1 (2018), 47–66 https://doi.org/10.1016/j.clsr.2017.05.019

Zhang, Chenguo, '"Sampling" Is Freedom of Art: The German Federal Constitutional Court Deliberates on the Acceptability of Music Sampling in the "Metall Auf Metall" Case', *Computer Law and Security Review*, 33.6 (2017), 870–75 https://doi.org/10.1016/j.clsr.2017.06.005

Zhang, Lu, '"Personal Information of Privacy Nature'"under Chinese Civil Code', *Computer Law and Security Review*, 43 (2021) https://doi.org/10.1016/j.clsr.2021.105637

Zharova, Anna Konstantinovna, and Vladimir Mikhailovich Elin, 'The Use of Big Data: A Russian Perspective of Personal Data Security', *Computer Law and Security Review*, 33.4 (2017), 482–501 https://doi.org/10.1016/j.clsr.2017.03.025

Zheng, Guan, 'Trilemma and Tripartition: The Regulatory Paradigms of Cross-Border Personal Data Transfer in the EU, the U.S. and China', *Computer Law and Security Review*, 43 (2021), 105610 https://doi.org/10.1016/j.clsr.2021.105610