

## European Perspectives on Datafication

Mihai Ciobotea

Marian Stan

*The Bucharest University of Economic Studies, Romania*

[mihai.ciobotea@gmail.com](mailto:mihai.ciobotea@gmail.com)

[marian.stan.ro@gmail.com](mailto:marian.stan.ro@gmail.com)

### Abstract

*Datafication is changing from the value of concept to the value of layer for the present world. It describes every aspect of our lives with data.*

*In this paper, the researchers put the datafication in the context of the Fourth Industrial Revolution. Also, the authors examine some aspects changed by datafication. The researchers look at several aspects (put in context with the industrial evolution) and also to the effect of datafication on several areas of human activity. By using a qualitative approach, the research examines the datafication and its various dimensions.*

*The results suggest the need for robust, transparent systems to manage the complexity of data and advocate for ethical and privacy considerations. These insights are crucial for policymakers, businesses, and academics, providing a comprehensive understanding of the legal environment, societal impacts, and innovation potential in the European Union.*

**Key words:** datafication, European Union, education

**J.E.L. classification:** C55, C8

### 1. Introduction

Datafication is defined as the process of translating things and events into data that can be added to massive databases which grow daily (Gulson, 2022, p.3). Today everybody has a mobile device and everything we do is converted into data, whether it is related to education (Breiter et al., 2016), healthcare (Ruckenstein and Dow Schüll, 2017, Kuntsman et al., 2019), border control (Bellanova and Glouftsiou, 2022) professional life, personal life or any other thing surrounding us.

Datafication refers to the transformation of various aspects of life into data, enabling detailed analysis and the creation of new value from data. The concept has been introduced by Kenneth Cukier and Viktor Mayer-Schönberger in 2013 (Biltgen et al., 2016).

In the global context, datafication has revolutionized industries, enhancing decision-making processes, and fostering innovation

The paper examines qualitatively the paradigm changes resulting from Datafication and the measurement or data gathered and saved for almost all the aspects of our lives, extensively.

The authors have looked to several aspects of the present, studying how our perspectives have been changed by data.

Almost every aspect of our life is expressed by data, and this is on an increasing trend. Nobody can be out of this system.

Although at the global level there is a substantial interest in researching the datafication topic, there is still an opportunity for additional studies that explore the European context, especially regarding the recent legal and administrative changes and their associated impact.

The implementation of the General Data Protection Regulation (GDPR) had an impact on the way the data are used and protected in European Union. This may represent a significant research topic.

The main objective of this paper is to explore the most important elements of datafication in Europe. We aim to assess how European values, policies and practices influence the datafication.



of data governance: data sharing pools (DSPs), personal data sovereignty (PDS), data cooperatives (DCs) and public data trusts (PDTs).

Bellanova and Glouftisios (2022) have analyzed the Schengen Information System (SIS II), highlighting the system's fragility and the continuous maintenance required to sustain its functionality.

Calzada (2023) introduces a classification of digital citizenship regimes such as andemic citizenship, algorithmic citizenship, liquid citizenship, metropolitan citizenship, and stateless citizenship from various European locations. The author questions the current understanding of how these regimes reshaping the practices of European locations.

Kuntsman et al. (2019) investigates the ethical and privacy aspects associated with digital health technologies such as commodification of user data and privacy issues. Considering the context of European General Data Protection Regulation (GDPR), the authors show that there is a mismatch / discrepancy between what users are told and what the apps are actually doing with the user data. Also, the authors propose a tool used to evaluate apps with a focus on the opting out feature.

In empirical research, Bibri and Krogstie (2020) has proposed a new model for sustainable cities based on data-driven technologies. The authors have integrated back casting and four case studies related to top European cities and argued that proposed model may bring the sustainable cities closer to the goals of sustainability.

Rychnovska (2021) explores how biobanking anticipates and manages threats and risks associated with the biomedical data in Europe. The author addresses the challenges of integrating big data in health and the ethical implications of datafication in biobanking.

Valdivia et al. (2022) propose a transdisciplinary methodology to study datafication at EU borders. The paper explores biometric technologies and the transparency of data practices, highlighting the power asymmetries in data governance.

The reviewed literature seems to indicate the importance of exploring alternative models of data governance, addressing ethical and privacy concerns, and leveraging data-driven technologies for sustainable development in the European Union(EU).Also, the research shows the necessity for robust systems and transparent practices in managing complexity and challenges related to datafication.

A revolution is associated with radical changes, and the fact that datafication generates an abrupt change regarding the world surrounding us, by measuring a lot of things, we can conclude it is at least a part of a revolution, if not a revolution in itself.

In a classification regarding the important economic changes in human history (Schwab, 2017, p.6), there was an “Agrarian” revolution, followed by four industrial revolutions, as shown in the table below:

*Table no. 1 Industrial Revolutions (upon data from Schwab, 2017)*

#	Revolution	Period	Description
1	Agrarian Revolution	10.000 years before	Apparition of the domestic animals; Efforts for production, transportation, communication
2	First Industrial Revolution	1760-1840	Creation of the first railroads; Invention of the steam engine, resulting in mechanical production
3	Second Industrial Revolution	Late 19 <sup>th</sup> century – beg. 20 <sup>th</sup> century	Mass production
4	Third Industrial Revolution	1960s	Start of the digital era – computers
5	Fourth Industrial Revolution	2000s	Mobile internet Stronger, little, cheaper sensors Biotech innovation

*Source:* (Schwab, 2017, p.6); adapted by the authors

The Fourth Industrial Revolution has three arguments:

- Speed;
- Dimension (volume, depth);
- Systemic impact.

Schwab (2017) considers machine learning and AI to be part of the Fourth Industrial Revolution; however, the authors consider that AI and related techniques are a next level revolution. If the Fourth Industrial Revolution is related to the digital era and data, the AI (with Machine Learning and Deep Learning) is different and refers to data usage and therefore it constituted to the next, Fifth Industrial Revolution.

Considering all the facts, the AI era is contemporary with an Environmental Revolution, mankind paying more attention to factors influencing the environment and climate change. One aspect is the energy: we try to produce more energy but renewable, clean, without greenhouse gas generation. Especially the developed countries on one hand invest in solar, wind, nuclear energy, on the other hand are shifting from classic thermal vehicles to electric vehicles, reducing the carbon emissions.

Datafication is present in every aspect of our lives. In the next part we will explore the more relevant domains, with the implication and changes brought by the Data Economy.

### 3. Research methodology

The objective for this research is to explore the role of datafication and data economy in the context of today's realities and in the context of the Industrial Revolutions (Schwab, 2017). In this regard, we explore several aspects of today economy, highlighting the most important aspects, with more focus on the EU area.

For this research, the authors have performed a literature review and document analysis. The design framework of the current study has been presented in Table 2 below.

Table no. 2 The design framework

Data Type(s)	Unit(s)	Variables	Longitudinal study (yes / no)	Themes
Qualitative: Literature review based on the query in Web of Science (WoS) database	Research papers from WoS database: Articles, Proceeding Papers, Review Articles	Main findings and themes related to datafication	No	Utilizations of datafication; impact of datafication in various domains
Qualitative: document analysis related to GDPR		GDPR directives	No	Document analysis insights; specifics of datafication in Europe

Source: Created by authors

#### Explanation

- **Data Type(s):** The study relies on qualitative data gathered from a literature review of academic articles (articles, proceeding papers, review articles) from the Web of Science (WoS) database. Additionally, it includes a detailed content analysis of relevant documents pertaining to datafication in Europe.
- **Unit(s):** The units of analysis consist of academic documents that discuss datafication in the European Union.
- **Variables:** Themes in datafication focus on identifying relevant ideas and research gaps from academic literature.
- **Longitudinal Study:** This study does not track changes over time but instead provides a snapshot analysis based on the current literature and e-learning platforms content.
- **Themes:** Utilizations of datafication; impact of datafication in education; document analysis insights; specifics of datafication in Europe.

The literature review has been conducted in Web of Science Core Collection (WoS) database based on the following criteria: 1) Articles published in English ; 2) Publications of type: articles, review articles, conference proceedings ; 3) topic “datafication” . The searches have been further refined in order to filter only the papers that refer to European context. This has been done by including a separate criteria topic “euro\*”.

Furthermore, we have used VosViewer, version 1.6.19 in order to generate a co-occurrence map of the keywords.

The document analysis has included the following:

- General Data Protection Regulation (GDPR);
- Directive on Security of Network and Information Systems (NIS Directive);
- Regulation on Privacy and Electronic Communications (ePrivacy Regulation);
- Data Governance Act;
- Data Act.

#### 4. Findings

GDPR has impacted datafication through the strict guidelines on personal data collection, processing and storage. GDPR is based on consent, transparency and accountability for personal data usage, leading to more responsible data practices.

A particular side is represented by the ePrivacy Regulation, regarding personal data and the protection of privacy in the electronic communications sector.

The NIS directive (Network and Information Systems Directive) has strengthened the security across the EU, requiring the operators and providers to have in place robust cybersecurity measures and control the incidents, to cooperate with the national authorities.

The Data Governance Act (DGA) came into effect on June 23, 2022, and after a 15-month transition period, it has been applicable since September 2023. (European Commission, 2022). This is intended to provide significant advantages to European Union (EU) citizens and businesses by having more data available and sharing across EU countries and industries.

The Data Act came into effect on January 11, 2024 (European Commission, 2024a). This complement the DGA and is intended to provide more legal clarifications regarding the data access and use (European Commission (2024b).

Datafication has many consequences, from opening new horizons of data usage, opening the gate for new powerful tools, as well as some negative things, as discrimination in more ways, given by lack of access to data-driven systems (Kennedy, 2020).

The Fourth Industrial Revolution brought with itself lots of data, bringing into existence the concept of big data. The quantity and complexity of the data increased the need of special tools to deal with it. Because of quantity and complexity, this task cannot be done without machines.

In this data driven environment, it is a need learn for both humans and machines. This leads to the concept of ‘learnification’ (Knox, 2019), which has two significant ways:

1. The training of machines
2. The nudging of human decision through digital choice architectures.

The high quantities of data generate the need to be governed effectively. There are proposed models on how to organize this governance. An example is a social science-informed conceptualization of data governance (Micheli, 2020), based on four models emerging from the actors taking part in data governance: data sharing pools, data cooperatives, public data trusts and personal data sovereignty.

Datafication is enabled in a lot of ways: any online activity (which leaves traces, kept by service providers and other organizations), commercial activity, business processes, sensors and lots of other data generators.

An interesting subject, when discussing about datafication, which should be considered also the epistemic equality. Hayes theorizes the philosophical basis of a changed ‘datafication’ process for teaching excellence (Hayes, 2020); the discourse of teaching should be allocated according to member’s relevant expertise on the subject.

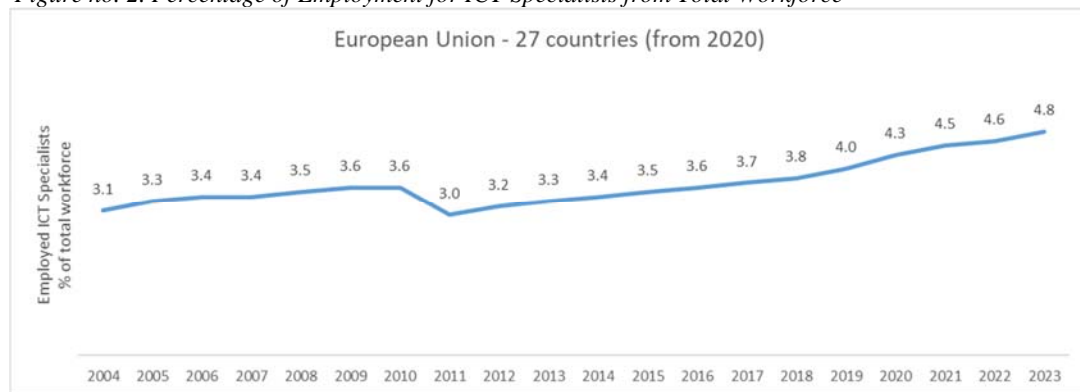
A related concept refers to the datafication of personal information. This led to the regulation of personal data (GDPR – General Data Protection Regulation), defending the personal rights and ownership for some specific data.

The increase of data volume generated the need of digital skills, which are also developing. The question is if the workforce with skills needed to operate with the data grows enough to match the data volume expansion.

The COVID-19 Pandemic specifics have created the circumstances to catalyze and therefore to accelerate some datafication processes (Nguyen, 2020), for any activity which could be moved from real life to the digital world. As a consequence, teaching/learning, back-office activities, clerk activities and many others were digitalized. It was a great time for people aiming to improve their skills.

From the next chart we can see that the COVID-19 pandemic has slightly accelerated the percentage of ICT specialists’ employment, from 4.0% to 4.3%, which can be explained by the fact that people had time to improve their digital skills during the pandemic.

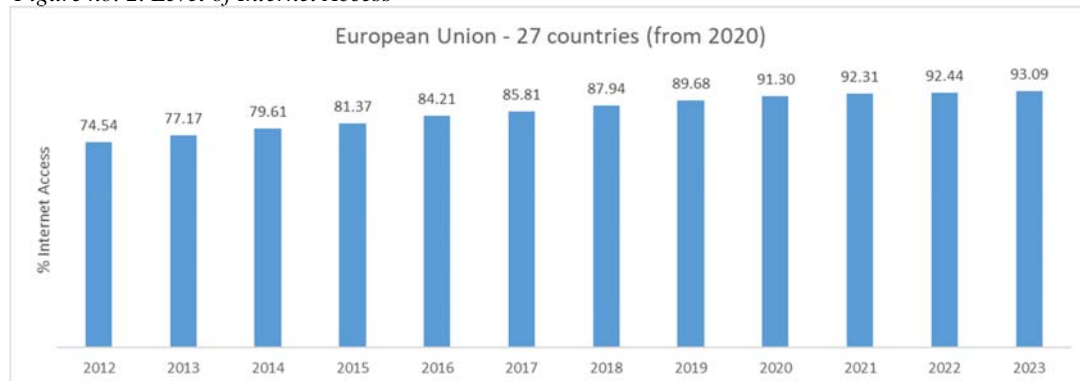
Figure no. 2. Percentage of Employment for ICT Specialists from Total Workforce



Source: (European Union – Eurostat)

Datafication is increasing due to digital enhancement and usage, which is also increasing with the internet access, with more than 93% in 2023, as shown in the next figure.

Figure no. 2. Level of Internet Access



Source: (European Union – Eurostat)

Online learning is a specific item of Datafication, it generates data related to the courses, students and lots of related items. It can be identified an important increase starting 2020 (as a consequence of the COVID-19 pandemic protection rules), with a peak in 2021.

An important conclusion of the paper is that actually we have advanced with 2 steps: the first regards generating and saving data and corresponds to the Fourth Industrial Revolution, as described by Schwab (2017), the second one is linked to usage of the available data and can be part of a new industrial revolution, related to how data are used and processed. The automation and enhancement of data processing gets to a totally different level through big data and AI.

Datafication creates challenges also related to data governance, especially regarding privacy. Actually, it creates the need to shift from thinking about definitions of privacy (characteristics of privacy) to models of privacy (how privacy works), as a change of paradigm (Mai, 2016).

*Another conclusion is that ‘Disconnection is futile’ (Karppi, 2020). The world paradigm has changed, our life ways have changed, nobody and nothing can disconnect totally, whatever men do, they have for sure a digital mark, through data.*

Datafication is a phenomenon in development, and it is the foundation of the world as we know it. This is because the need to govern data, but also datafication constitutes the nouns of the grammars used by humans and machines.

## 5. Conclusions

The study of datafication in the EU context shows multiple important perspectives:

1. Transforming through Datafication: an essential part of the Fourth Industrial Revolution, datafication revolutionizes the ways we generate, save and use data. This transformation has progressed in two stages: first stage concentrated on data generation and saving, and the current phase focused on data usage and processing through advanced technologies, like big data and artificial intelligence.
2. Regulations impact: The EU regulation framework, through instruments like GDPR and the NIS Directive (Network Information Systems Directive) has put order into the datafication process. The EU regulations draw and ensure strict rules regarding data management, with a focus on consent, transparency, accountability and cybersecurity.
3. Confidentiality and Data Governance: As datafication phenomenon is in continuous evolution, we need to change the paradigm regarding the confidentiality. Instead of defining only confidentiality, it is necessary to model the way confidentiality works in a data-based world. This change is essential for developing data effective frameworks able to govern the data usage complexity and to protect individual rights;
4. Inevitable Digital Footprint: the nature of datafication means that total disconnection from full digital footprint is practically impossible. As data enters into each life aspects, inevitably each person leaves digital marks, showing the inevitable place of data in the modern society.
5. Competency and educational development: The COVID-19 Pandemic accelerated the datafication processes, especially in education and personal development. The transition to online learning and digital platforms generated lots of data, showing the necessity to improve the digital competencies.
6. Future directions: as datafication continues its evolution, the implications for a lot of areas, like education, business, personal data administration will become more and more prominent. Future researches and public policies development should address the challenges and opportunities generated by datafication, ensuring that data administration practices are in line with ethical standards and with the society needs.

Generally, datafication represents a fundamental change in the way data comingle with the day to day life and with the economic activities. Continuous developments regarding data governance, confidentiality and the technological advances will continue to shape of datafication in Europe and everywhere.



## 6. References

- Bellanova, R. and Glouftsiou, G., 2022. Controlling the Schengen Information System (SIS II): The Infrastructural Politics of Fragility and Maintenance. *Geopolitics*. DOI: <<https://doi.org/10.1080/14650045.2022.1952356>> [Accessed 10 Jun 2024]
- Biltgen, P. & Ryan, S., 2016. *Activity-Based Intelligence: Principles and Applications*. 1st ed. Norwood, MA: Artech House
- Breiter, A., 2016. Datafication in education: a multi-level challenge for IT in educational management. In *Stakeholders and Information Technology in Education: IFIP TC 3 International Conference, SaITE 2016, Guimarães, Portugal, July 5-8, 2016, Revised Selected Papers 1* (pp. 95-103). Springer International Publishing. [https://doi.org/10.1007/978-3-319-54687-2\\_9](https://doi.org/10.1007/978-3-319-54687-2_9) [Accessed 10 March 2024].
- EUR-Lex, 2002. Document 02002L0058-20091219 - Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), [online], Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02002L0058-20091219&qid=1718225692463> [Accessed 10 Jun 2024]
- EUR-Lex, 2016. Document 02016R0679-20160504 - Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), [online], Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504> [Accessed 10 Jun 2024]
- EUR-Lex, 2022. Document 02022L2555-20221227 - Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive), [online], Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02022L2555-20221227> [Accessed 10 Jun 2024]
- Eurostat (2024) *Employed ICT specialists - total [isoc\_sks\_itspt\_custom\_11567955]*. Eurostat. Available at: [https://ec.europa.eu/eurostat/databrowser/view/isoc\\_sks\\_itspt\\_custom\\_11567955/default/line?lang=en](https://ec.europa.eu/eurostat/databrowser/view/isoc_sks_itspt_custom_11567955/default/line?lang=en) [Accessed: 18 May 2024].
- Eurostat (2024) *Level of internet access - households [tin00134\_custom\_11569871]*. Eurostat. Available at: [https://ec.europa.eu/eurostat/databrowser/view/tin00134\\_custom\\_11569871/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/tin00134_custom_11569871/default/table?lang=en) [Accessed: 18 May 2024].
- Gulson, N.G., Sellar, S. and Taylor Webb, P., 2022. *Algorithms of Education. How Datafication and Artificial Intelligence Shape Policy*. Minneapolis, University of Minnesota Press
- Hayes, A., Cheng, J., 2020. Datafication of epistemic equality: advancing understandings of teaching excellence beyond benchmarked performativity. *Teaching in Higher Education. Critical Perspectives*. [online]. Available at: <https://www.tandfonline.com/doi/full/10.1080/13562517.2019.1689387> [Accessed 25 May 2024]. <https://doi.org/10.1080/13562517.2019.1689387>
- Karppi, T., Couldry, N., Mejias, U., 2020. Disconnection is futile – Theorizing resistance and human flourishing in an age of datafication. *European Journal of Communication, Vol.35 Issue 3*. [online] Available at: <https://journals.sagepub.com/doi/10.1177/0267323120922094> [Accessed 25 May 2024]. <https://doi.org/10.1177/0267323120922094>
- Kennedy, H., Steedman, R., Jones, R., 2020. Approaching public perceptions of datafication through the lens of inequality: a case study in public service media. *Information, Communication & Society*, [online]. Available at: <https://www.tandfonline.com/doi/full/10.1080/1369118X.2020.1736122> [Accessed 20 May 2024]. <https://doi.org/10.1080/1369118X.2020.1736122>
- Knox, J., Williamson, B. and Bayne, S., 2019. Machine behaviourism: future visions of ‘learnification’ and ‘datafication’ across humans and digital technologies, *Learning, Media and Technology*, <https://www.tandfonline.com/doi/full/10.1080/17439884.2019.1623251> [Accessed on 25 May 2024]. <https://doi.org/10.1080/17439884.2019.1623251>
- Kuntsman, A., Miyake, E. and Martin, S., 2019. Re-thinking Digital Health: Data, Appisation and the (im)possibility of 'Opting out'. *Digital Health*. <https://doi.org/10.1177/2055207619863148>
- Mai, J.-E., 2016. Big data privacy: The datafication of personal information. *The Information Society. An International Journal* [online]. <http://dx.doi.org/10.1080/01972243.2016.1153010>



- Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, 7(2). [online]. Available at: <https://journals.sagepub.com/doi/10.1177/2053951720948087> [Accessed 25 May 2024] <https://doi.org/10.1177/2053951720948087>
- Nguyen, D., 2020. Mediatisation and datafication in the global COVID-19 pandemic: on the urgency of data literacy. *Media International Australia. Extraordinary Issue: Coronavirus, Crisis and Communication*, [online] Available at: <https://journals.sagepub.com/doi/full/10.1177/1329878X20947563> [Accessed 20 May 2024] <https://doi.org/10.1177/1329878X20947563>
- Micheli, M., Ponti, M., Craglia, M. & Suman, A.B., 2020. Emerging models of data governance in the age of datafication. *Big Data & Society*, 7(2), p.2053951720948087. <https://doi.org/10.1177/2053951720948087>
- Ruckenstein, M. and Schüll, N.D., 2017. The datafication of health. *Annual Review of Anthropology*, 46, pp.261-278. <https://doi.org/10.1146/annurev-anthro-102116-041244>
- Sadowschi, J., 2019. When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, January-June 2019, 1-12. Available at: <https://journals.sagepub.com/doi/10.1177/2053951718820549> [Accessed 25 May 2024] <https://doi.org/10.1177/2053951718820549>
- Schwab, K., 2017. *The Fourth Industrial Revolution*. New York: Crown Business. <https://dl.acm.org/doi/10.5555/3137529>
- Van Eck, N.J. and Waltman, L., 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), pp.523-538. <https://doi.org/10.1007/s11192-009-0146-3>