

# Digital transformation and data ecosystem: implications for policy actions and competency frameworks

Monika Rozkrut<sup>a</sup>

**Abstract.** The article discusses EU policy for digital transformation and the associated development potential. The article aims to critically analyse the current progress of the operationalisation and implementation of the relevant policies. It is followed by the recognition of the challenges that can contribute negatively to the necessary strategic objectives and obstacles that may hinder reaching the policy goals. In particular, a significant obstacle may be a major deficiency of adequately prepared experts ready to work in new roles in a dynamically developing data ecosystem. A remarkable example is the role of the Data Steward. This role is essential for fostering the rapid development of the data ecosystem in the EU. We propose creating a universal competence framework for Data Stewards to streamline human resource allocation. The article proposes a basic outline of the necessary skills and competencies ensuring effective data stewardship.

**Keywords:** digital transformation, data ecosystem, Data Steward, data stewardship, data sharing, competence framework

**JEL:** C80, O30, J24, O15

## Transformacja cyfrowa i ekosystem danych – implikacje dla tworzenia polityk i wymagań kompetencyjnych

**Streszczenie.** W artykule omówiono politykę Unii Europejskiej dotyczącą transformacji cyfrowej i związany z nią potencjał rozwojowy. Celem pracy jest krytyczna analiza postępu, jaki dokonuje się obecnie w zakresie operacjonalizacji i wdrażania odpowiednich polityk. Ponadto zidentyfikowano wyzwania, które mogą mieć negatywny wpływ na osiągnięcie koniecznych celów strategicznych, oraz przeszkody mogące utrudniać realizację przyjętych polityk. Jako szczególne utrudnienie postrzega się znaczny niedobór ekspertów przygotowanych do pełnienia nowych funkcji w dynamicznie rozwijającym się ekosystemie danych. Znaczącym przykładem jest rola data stewarda, kluczowa dla wsparcia szybkiego rozwoju ekosystemu danych w UE. Zaproponowano stworzenie uniwersalnych ram kompetencji dla data stewardów w celu usprawnienia zarządzania zasobami ludzkimi. W artykule przedstawiono podstawowy zarys niezbędnych umiejętności i kompetencji zapewniających efektywne zarządzanie danymi.

---

<sup>a</sup> Uniwersytet Szczeciński, Wydział Ekonomii, Finansów i Zarządzania, Instytut Ekonomii i Finansów, Polska / University of Szczecin, Faculty of Economics, Finance and Management, Institute of Economics and Finance, Poland. ORCID: <https://orcid.org/0000-0002-9701-6388>. E-mail: [monika.rozkrut@usz.edu.pl](mailto:monika.rozkrut@usz.edu.pl).

**Słowa kluczowe:** transformacja cyfrowa, ekosystem danych, data steward, zarządzanie danymi, wymiana danych, ramy kompetencji

## 1. Introduction

Over the last few years, modern digital technologies have significantly influenced changes in economies and societies, affecting individual enterprises, entire sectors of activity, public administration, and citizens' everyday lives. Digital transformation is changing the way people live, work and communicate. Data are at the heart of this transformation. The amount of information generated by citizens, businesses and public authorities is constantly growing. Whole new data ecosystems are created, with enterprises, public institutions, and even households collecting data. These can be treated as a new resource and a new factor of production. The growing amounts of industrial and public data, combined with technological changes in how they are stored and processed, represent a potential source of growth and innovation that should be harnessed. The wide availability of data and their safe and open exchange offer the possibility to respond to many vital economic and societal challenges. The proper use will enable enterprises, the public sector and citizens to make more informed and beneficial decisions positively influencing their development. Fast and secure data exchange will ensure the further growth of a knowledge-based economy and enhance the quality of life. Therefore, there is a growing need to tackle data governance, better organise and improve the functioning of the data ecosystem, and make it cohesive and coherent, at least in the basic scope, to facilitate data exchange and application between all interested parties. New regulations are necessary to increase trust in sharing data and ensure the security of their processing; guidance over data quality is likewise crucial. All these needs, in turn, translate into a growing demand for qualified staff with competencies in specific areas related to the broadly understood digital transformation. The article aims to critically analyse the strategic activities of the European Union aimed at stimulating the progress of the digital transformation and creating a data space. The analysis points out that a significant obstacle to reaching the policy goals is the shortage of adequately trained specialists to perform the role of Data Stewards. The emergence and rapidly increased interest in jobs related to data stewardship seem particularly important. It seems that fostering this newly reinvented role determines the development of properly set up and functional data ecosystems adequate to its potential and needs.

## 2. Digital transformation

As general-purpose technologies, information and communication technologies (ICT) saturate all aspects of socio-economic life. On the one hand, they change the ways of social interactions, and on the other hand, they more and more often enable direct communication between devices and objects that create what is called the Internet of Things. Companies are changing their internal structures to use ICT effectively. Employees acquire new skills to use them, process information and learn. Public administration adjusts the ways of interacting with citizens and businesses accordingly. New means of communication lead to the creation of new behaviour models and, as a result, consumption patterns are changing as well. Undoubtedly, new technologies are currently one of the fastest-growing sectors of the economy in developed countries, characterised by an increasing share in the creation of GDP.

The progress in modern computer techniques and technologies has led to a revolution in collecting, processing, storing and transmitting information (Bell, 1973). Modern digital technologies influence and support business activity. The digitalisation process concerns communication and business management and often affects production or service processes. This results in IT systems in enterprises that collect, send, store and process larger and larger data sets.

These technologies fundamentally change how various entities operate: producers, service providers, customers and eventually the whole economy (Horton, 2007; MacKay & Vogh, 2012). Hence, we discuss how the private and public sectors are transforming, including the transformation of government and local government authorities and administration. The use of modern technologies in providing services by public administration bodies is described as e-services or e-government, synonymous with electronic administration, e-administration, and e-governance. Contacts between a given organisation and its service providers and recipients occur electronically, thanks to ICT.

The digitalisation of the economy produces more and more data. Data can, therefore, be determined as a new resource, a new production factor. Data collected by companies, especially on what are called platforms (platform economy), allow them to assess the demand, profile the advertising content and personalise the product. Digitalisation technology, i.e. everything that allows data to be collected, processed and analysed leads to new business models that affect user behaviour and business response, distorting companies' established operations in numerous sectors of the economy. The emergence of Airbnb led to a shock in the hotel market; the appearance of Uber revolutionised public transport. The platform economy expands to other sectors of the economy in the first place, where information is at the heart of the business model.

The transformation processes permeate the entire economy, gradually changing it into a digital economy founded on electronic data exchange. It is the crucial axis of a modern economy that also develops in virtual space. A digital economy results from technological progress, new means of communication, the accumulation of knowledge and data processing methods. It is a consequence of the development and convergence of data processing techniques (dynamic growth of computing power, data storage capacity), telecommunications (data transfer speed, advancement of protocols, infrastructure investments) and knowledge accumulation (algorithms, data science).

### **3. Data ecosystems – a strategic perspective**

Technological changes in storing and processing the growing amount of data are a major source of innovation. New digital data ecosystems are created that can be understood as systems for generating processes, collecting and storing data resources, and continuous exchange between individual entities of this ecosystem participating in social and economic life. They are the effect of the possibilities that digital technologies jointly offer. Data enable ecosystem entities to make better decisions, resulting in higher performance, competitiveness and more efficient management. Here, the issue of entities entering the data ecosystem and their potential roles appears. We deal with an outburst of possibilities in this case as well. With advancing digital transformation processes, the generation and acquisition of data cease to be the domain of only a narrow group of specialised entities. These are not only IT enterprises anymore, because digital technologies are increasingly often used across almost all industries and beyond that, not only in businesses but also in the public sector, NGOs or simply among citizens who can passively share their data using modern devices and services, or actively participate in citizen-generated data projects.

Data is at the core of the digital economy. The increasing volume of the generated data and the number of actors involved lead to a growing need to undertake necessary governance-related policies to improve the functioning of data ecosystems. Firstly, it is important to facilitate the exchange of the collected data between individual entities of socio-economic life. The gathered data should be available to everyone regardless of the nature of the entities (public, NGOs or private), size or maturity. Thanks to this proceeding, the benefits that data bring will be maximised.

Decomposing the EU strategic perspective may be a helpful tool to contextualise further considerations in this article. The European data strategy aims to enable the EU to make better decisions and solve current political issues, such as resource and climate problems, leading to a data-agile economy driving overall innovation. In

February 2020, the European Commission published two documents key to the data-based economy: 'Shaping Europe's Digital Future' (European Commission [EC], 2020b) and 'A European Strategy for Data' (EC, 2020a). In the first one, three essential elements were indicated for the correct shaping of the digital future of Europe: technology benefits people, a fair and competitive economy, and an open, democratic and sustainable society. The document on the digital future of Europe contains essential elements of the EU's strategy for a data-based economy. It is an introduction to the content laid out in the second regulation. In the data strategy, the European Commission described the vision of European Data Spaces, a common, single data market on which they could be used regardless of the country of origin. It provides a list of the legal activities and investments that will be made over the next few years. The goal is to exploit data and demand for goods and services based on data. The strategy for data undertaken by the European Commission activities is based on four pillars. Firstly, there is a plan to create a legal framework for data management, tackling the availability of public sector data and actions used to exchange data efficiently within sectors should also be intensified. The second strategy pillar applies to the provision of support for technology development and digital infrastructure. The third pillar of the strategy involves investment in competencies and general skills to use data. The European Commission's activities aim to reduce the gap in extensive data acquisitions and data analysis capacities. The fourth pillar of the strategy for building a data-based economy concerns the development of a common European data space in strategic sectors and fields of public interest (e.g. relating to mobility, green governance, industrial data, energy, agriculture and health).

As indicated in the document, the possibilities of using data for innovative purposes are essential in the data-based economy. Data availability is a crucial problem. Even in the cases where one may expect an abundance of data, a lack of sharing mechanisms may severely hamper the innovation potential. Therefore, the data value is best assessed through the prism of its reuse. Hence, the need to unblock the flow channels is one of the critical elements of the European strategy. A typical data ecosystem for the development of collaboration in the field of data sharing may be decomposed into four main spheres – B2B (Business-to-Business), B2G (Business-to-Government), G2B (Government-to-Business) and finally, G2G (Government-to-Government).

In the case of B2B data sharing, the consensus is that it is still not dynamic enough and needs enhancement in the EU. In the coming years, it should be made more viable economically and, thus, more attractive thanks to the new policy initiatives. Fears against the loss of competitive advantage, the lack of mutual trust, concerns about data appropriation and exploitation by third parties are among the factors that

affect the current state. There are plans to run new policies and programmes to financially support this type of data sharing while offering guidelines to facilitate agreements guaranteeing equal negotiation positions of the parties and the security of the information provided.

Next, B2G data sharing is associated with the use of private data by public authorities. The European Commission's documents state that using private-sector data is insufficient. At the same time, the public sector can significantly improve the policy-shaping process and delivery of public services based on the evidence available through data. One prominent example of the above is the potential to considerably increase official statistics' scope, granularity and timeliness. Strengthening data sharing in this regard would accelerate the development of a data-based society and improve decision-making processes in the public sector. The European Commission recommends developing appropriate incentives to build data sharing. It proposes to analyse the legitimacy of introducing an EU framework regulating the reuse of private data in the public interest. There is a chance that the individual member states will initiate national programmes facilitating access to privately-held data by the public sector.

The primary assumption in the G2B model is to provide data to businesses by public administration. Data generated thanks to public funds should benefit the society to the broadest possible extent. It applies particularly to high-value data sets, including different forms of protected data, which are only sometimes made available for research due to the need for mechanisms consistent with the provisions on personal data protection.

A crucial role in building a data-based economy is G2G data sharing, i.e. data exchange between public authorities. In this respect, the preparation and implementation of the relevant regulations remain a competency of member states' administration at the national level. Cooperation on this platform will significantly improve the policy and provision of public services and reduce administrative burdens for market-operating enterprises.

The EU should become a place where data enable better decisions. However, this goal must be founded on a solid legal framework dealing with fundamental rights, data protection and security. Suppose the EU is to play a leading role in a data-based economy. In that case, it must take action now and in a coordinated manner to deal with access and data storage, computational and cyber security. In addition, the EU will have to improve its data processing structures and increase the pool of high-quality data available for reuse. Issues related to cyber security tackle many fields – public administration, financial institutions, international corporations, small and medium-sized companies, and individual users. A balance must be kept between

a vast flow and use of data and a high level of privacy, security and ethical standards to unleash the potential of data to the fullest.

#### **4. EU policy actions**

Undoubtedly, effective policies need the support of appropriate regulatory frameworks and governance. The primary goal is to create regulations that raise confidence in the sharing process and ensure the security of processed data. In 2014, the European Commission published the ‘Towards a thriving data-driven economy’ (EC, 2014b) communication, which relied on a coordinated action plan involving the member states and the EU. The European Commission proposed policy initiatives to address bottlenecks in the data ‘reuse potential’ by creating a common European data space. Consequently, the adopted ‘Towards a common European data space’ (EC, 2014a) European Commission communication proposed a package of measures pointing to the four primary modes of data sharing referred to above, i.e. B2B, B2G, G2B, and G2G. As one of the results, Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, also known as the ‘Open Data Directive’, entered into force on 16 July 2019, providing a common legal framework for a European market for government-held data (public sector information). The Directive was to be transposed into national legislation by June 2021.

In its European Data Strategy put forward in February 2020 (EC, 2020a), the Commission declared subsequent initiatives: an implementing act on high-value data sets and two major legislative proposals: a governance framework for common European data spaces (Data Governance Act; EC, 2020d) and a Data Act. Implementing the regulation on high-value data sets will include a list of data with high commercial potential, speeding up the emergence of value-added EU-wide information products.

The new regulations are to become the basis for the European Data Policy. Many regulatory initiatives are in preparation (see below), but two are fundamental. The main policy implications of the EU data strategy are prepared as the Data Governance Act and the Data Act. The Data Governance Act provides an overarching governance framework for establishing and functioning common European data spaces, constituting the core part of the European Commission’s broader data and digital strategies. It is designed to encourage investment in new infrastructure for sharing data, increasing data availability, strengthening the mechanisms for their sharing and ensuring the safety of processed data, helping to build a single digital market for data across EU member states. Legal solutions apply mainly to ensuring a high level of privacy and data security, which are subject to

sharing based on a new data management framework. The task of the Data Governance Act is primarily to increase market trust in the sharing process. The proposed provisions build the confidence of enterprises, government authorities and citizens to share information on the European data market.

The Data Governance Act has four pillars:

- granting access to public sector data for reuse in situations where these data may be copyrighted;
- sharing data between enterprises in exchange for remuneration in any form;
- enabling the use of personal data with the aid of an 'intermediary' helping individuals to follow 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 [GDPR]);
- enabling the use of data from altruistic motives.

The document contains information on how public and private sector data (usually unavailable due to intellectual property rights, trade secrets or privacy rights) could be made available using mechanisms guaranteeing anonymity and confidentiality under the applicable legal acts. These shall also include personal data protected by the General Regulation on Data Protection and within its legal standards. The Data Governance Act outlines how the data in possession of the public sector, charged with third-party rights (e.g. trade secrets and intellectual property rights) may be used. It contains provisions enabling data intermediaries (data brokers), defined as trustworthy actors, and sets the registration rules of entities that collect and provide data for charitable purposes. The new law aims to improve information protection and create a safe environment for obtaining, using and controlling data. The entire process of transferring and managing data is based on the neutrality and transparency of 'data intermediaries' responsible for data collection. These are to be trustworthy entities to which data are made available. Data intermediaries will be required to maintain neutrality and observance of the strict requirements, including the ban on using data for their interest. A certification or labelling framework has been proposed together with a notification obligation and then the monitoring of the compliance with the requirements for designated competent authorities in the member states. These proposals do not apply to data-sharing initiatives in closed groups.

In conclusion, the regulation is primarily to help build trust and raise process security, thus strengthening the practice of using data and creating innovative solutions. Government agencies should consider what data can be made available to guarantee their quality and what policy (and perhaps a price list) of their sharing should be adopted. The analysis of which data may be helpful and how to ensure



their safety and compliance with the requirements of other regulations will be necessary on the part of companies. In turn, potential data intermediaries should consider the relevant ideas for business and ways to facilitate data access and sharing. In each of the above, one has to be able to 'reign over data' to know what data one has and how they are processed and used. In practice, applying the proposed regulation will involve the inventory of the data held, classifying them according to the regulations they are subject to and ensuring their safety. The Data Governance Act also sets out a framework for the voluntary registration of entities that collect and process data made available for charitable or altruistic purposes.

The Act aims to enable data-driven innovation by setting a governance framework to promote confidence in data sharing and incentivise the expansion of EU data spaces while ensuring that natural persons and legal entities are in control of the data they generate. We are currently facing many forms of barriers to sharing data, like restrictive intellectual property rights, concerns about compliance with GDPR, fears of breaches of confidentiality, or fears of others deriving value from shared data when those who actually shared them were unable to do so. Data governance aims to promote data assets, going beyond some simple set of protection rules, but aiming more broadly at breaking down barriers to sharing. Thus, it has the potential to benefit businesses and the society more generally. The Data Governance Act establishes a framework to promote confidence in data exchange between organisations. It creates the basis for managing data spaces that comply with the values and laws of the EU, such as personal data protection, consumer protection and competition rules.

The last of the main proposed acts was planned to be adopted at the end of 2023. The Data Act will explicitly support B2B data sharing and B2G data sharing for public interest purposes, fostering access to data held by private sector entities when these data are of public interest. The European Commission hints that the right to data portability could be enhanced to give individuals more control over who can access and use their data; changes in the EU's intellectual property rights framework may be introduced, particularly in database rights and trade secrets.

## **5. The increasing complexity of data ecosystems**

Europe's digital strategy also highlights many other legislative initiatives the European Commission plans to introduce. These include: laws on crypto assets and digital operational and cyber resilience in the financial sector; online platforms (Digital Services Act; EC, 2020c), data centres, cybersecurity, a review of the Network and Information Security Directive, EU's existing eIDAS Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on

electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC (digital identities), action plans on '5G' and '6G'; new digital sector inquiry; or new strategies for payments, the industry, blockchains and quantum computing. Add the Open Data Directive, supporting the use of public sector data collections, the Database Directive, and the well-known General Data Protection Regulation, and we will receive an increasingly complex EU legal system related to data. Naturally, this brings in substantial issues, some still unclear. For example, the Digital Governance Act provides an oversight of data sharing by competent authorities. However, there is a potential for uncertainty of responsibilities if providers are subject to regulatory oversight by several different authorities in the countries in which they are located.

The next issue to mention relates to cross-border transfers of data. According to the same Governance Act, non-personal data that is subject to the rights of others may be transferred from an EU country to a third country only if proper safeguards are in place. The Act should ensure the protection of the fundamental rights of data holders. At the same time, third countries providing some equivalent (to the EU) level of protection should be allowed to transfer data across borders. How the European Commission is going to operationalise it needs to be made clear. One solution might be offering model contract clauses to gain reassurance that the non-personal data transferred outside the EU is protected. Next, the Digital Governance Act does not require data-sharing service providers to have an EU establishment, though the provider must appoint a legal representative in the EU.

Prospective data-sharing service providers with multiple establishments in the EU will be deemed to have their main establishment where their central unit is located. The GDPR, however, allows choosing the main establishment where the most important decisions about personal data processing are taken, not necessarily the central administration unit. Some other risk emerges in data altruism, specifically concerning forms proposed to gather consent from individuals to use their data, which is very broadly described in the regulation as consent to specific purposes or data processing in certain areas of research or parts of research projects when it is initially difficult to precisely identify the purpose at the time of data collection.

The above-mentioned issues naturally impact all modes of possible data sharing and may become extensively problematic in more complex cases like data collaboration and collaboratives. In these cases, the distributed nature of the data supply is matched with the distributed nature of demand for data. These can bring in a value-added insight that can potentially be generated only thanks to such initiatives. Naturally, the policy's intention is to instigate such collaboration in a more agile and instantaneous manner. The pandemic is clear evidence of such a need; it also clearly exposed a lack of preparedness for deep and swift data

collaboration. Data innovations may happen, yet many concerns still need to be addressed regarding incentives, limitations, obstacles, the lack of regulatory solutions or governance framework. New regulations are expected to fill this gap, but not without problems. However, the purpose of this article is not a detailed analysis of the potential problems, difficulties or issues requiring solutions. The intention is to demonstrate that the entities or persons who want to take advantage of the digital transformation and exponentially growing data ecosystems will need help with this increasingly complicated matter. Inevitably, a high degree of complexity can result in a significant burden, cost and lack of trained staff.

## **6. The role of the Data Steward**

Although permeated with significant social and economic goals and undoubtedly necessary for the digital development of the EU, numerous legislative proposals create an increasingly complex system of relations, connections, rules and limitations. Undoubtedly, as often in such cases, attempts to solve particular problems create new ones. In such a complicated system of mutual dependencies and numerous regulations, high competencies are necessary to enable efficient navigation between individual elements of the system according to the prescribed set of recommendations and rules. Although the perspective of broader use of data by analysts, resulting from the adopted measures seems very attractive, the analysts themselves will need to constantly acquire new competencies and extensive contextual knowledge in such a complex arrangement.

The solution is already available in the form of Data Stewards, a concept that has long been developed in the business and scientific community (Peng, 2018). Undoubtedly, the role of specialised Data Stewards will become more significant in an increasingly complex and exponentially growing data ecosystem. At this point, however, we want to propose the following side theses: changes caused by the dynamic progress of digital transformation mean that the current group of specialists in this field is insufficient for today's needs and that the growing strategic, political and regulatory interest causes that the usual extent of knowledge, competence and skills of an exemplary Data Steward must be significantly developed. Let us agree that these theses are valid. It subsequently leads to the following conclusions: the need for a notable increase of Data Stewards supply while ensuring an appropriate profile of knowledge, skills and competencies. Thus, we put forward the central thesis that the most appropriate tool to respond to the formulated conclusions is to create a proper competency framework for Data Stewards tailored to a rapidly changing environment, conditioned by the ongoing digital transformation processes and extensive new legal regulations. It is the fundamental conclusion of this paper.

The growing volume of data and effective data management processes require a proper blend of technologies and people. Demand for digital competencies goes far beyond the extent and quality needed before, even just a few years ago (Organisation for Economic Co-operation and Development [OECD], 2016a, 2016b, 2018, 2020). The skills necessary to manage and supervise a whole data life-cycle, assess data value, keep adequate data quality, their efficient sharing and reuse are highly important. The expectations towards adequately trained staff are growing. New roles in the data ecosystems emerge, further extending the scope of training. One particular example supported by the regulations is that of an intermediary. Their task will ensure the availability, quality and reuse of the existing data. By organising pooling and sharing data, they will not process it on their account. However, experts (European Commission & Directorate-General for Research and Innovation, 2016) point out substantial deficits in people and skills. Good Data Stewards are exceedingly rare. Some claim that one Data Steward is needed for every twenty data analysts, so a quick estimate suggests that many Data Stewards will be needed over the coming years to fill in the gap (Versweyveld, 2016).

Data stewardship concerns those who work with, protect and use data. A Data Steward is at the heart of any data governance programme. In short, data stewardship aims to design, build, implement and manage data, enabling users to make consistent decisions based on information. Typically, the tasks of Data Stewards include:

- helping define and implement data definitions, shared metadata, standardised, controlled dictionaries and standards;
- setting data quality guidelines that face requirements for what is considered to be good data quality;
- consistent management of data resources throughout the entire life-cycle in order to preserve their quality, integrity and consistency, and avoid redundancies and integrity-related failures;
- facilitating the reuse of and providing access to data resources (for internal or external purposes);
- maintaining high-quality metadata;
- cooperating with others engaged in creating, collecting, accessing, using, sharing and maintaining data;
- collecting information on the needs and feedback on the quality of data resources for which they are responsible, ensuring that data is fit for the purpose;
- being aware of the data protection policies, intellectual property and information security; ensuring data is protected and security procedures are enforced;
- organising and contributing to communication and promotional activities to increase awareness and use of data resources for which they are responsible;

- building, supporting and sharing knowledge; helping users understand the data better and recommending improvements.

Undoubtedly, the above list does not fully exhaust the scope of activities of Data Stewards; depending on the context, not all of these tasks need to be performed or can only be performed to a certain extent. Nevertheless, such a synthetic list of essential items constitutes a complex set of competencies, requiring years of training and gathering experience. While carrying out his or her activities, a Data Steward must not only demonstrate knowledge and deep understanding of the issues that remain in a strong relationship with his or her area of competence, but also of those going beyond it. For example, when assisting data analysts, a Data Steward should have some experience in carrying out such research, which can help to efficiently prepare an appropriate quantitative analysis development environment, allowing analysts to concentrate on the task at hand. The evolution of the role of the Data Steward requires an appropriate skill set. Data Steward duties cover all aspects of data governance; they must understand all levels of the business and are expected to demonstrate cooperation skills and the ability to collaborate and effectively communicate with other internal and external stakeholders, both in the language of business processes and technology. Data Stewards must be able to build relationships with others, promote best practices and demonstrate the proper use of data, following appropriate guidelines, rules and regulations, whether internal or external (Sen, 2018). Data stewardship aims to use information as an asset by defining strategies, standards, policies, models, processes, tools and methodologies to identify opportunities. Undoubtedly, the proper set of competencies is challenging to achieve.

## **7. Competency framework for Data Stewards**

Technological progress inevitably raises the issue of digital competencies. The discussion is not easy; the first obstacle is the very concept of competencies. They are interdisciplinary and broad, referring primarily to the practical ability to use modern technologies' wealth of tools and methods. Discussions on digital competencies can be found mainly in national and international strategic documents and scientific literature. The variety of definitions and approaches is undoubtedly associated with the dynamic evolution of technology and its proliferation in the economy and society. The concept of digital competencies is related to technology-oriented skills. Initially, digital skills were associated predominantly with access to equipment and the Internet and extended to more complex usage skills to achieve various life goals. 'Digital competencies' may be defined as a harmonious set of knowledge, skills and attitudes that allow the effective use of digital technologies in various areas of life

(Erstad, 2010). The concept of digital competencies covers an extensive set of skills that determine the efficient and conscious use of new technologies and the active participation in the life of the information society.

In this article, we argue for the need to eliminate the deep deficit of Data Stewards. A natural proposition to solve this problem is to invest in training appropriate personnel. A quick review of the existing literature in this area and observing what is happening in the education and labour markets indicate two crucial aspects to consider when planning further activities. Firstly, the role of Data Stewards has long been present in business, and there are also career paths that prepare for this role. Secondly, the analysis of the existing education programmes indicates a fundamental need for more opportunities at the university level.

The first of the indicated problems may mean that in the dynamic progress of the digital transformation, accompanied by an active response in the legislative sphere, we will face a deficit of specialists in the government and non-governmental sectors. The second problem means that there is still no consensus as to what competencies should characterise a Data Steward, at least one with a general profile of competencies, ready to assume the duties, regardless of the specificity and context of the tasks performed (after all, it is already possible to indicate a few potential specialisations for Data Stewards).

Therefore, we propose to develop an appropriate framework for the competencies of Data Stewards. A basic set of ready-made requirements would undoubtedly increase the awareness of the concept of Data Stewards, facilitating the preparation of appropriate fields of study, vocational training and training materials. There is no doubt that the expectations of the labour market in this respect will increase. Employers wish to know the potential employees' knowledge, skills and personal and social competencies. In turn, the candidates want to be aware of the content and level of learning outcomes, within and outside higher education, in line with lifelong learning. Competency frameworks also balance the different paths to achieve the intended qualifications (Punie et al., 2013). When developing the desired competencies of Data Stewards, it is possible to use the existing schemas based on the standard defined in the European Qualifications Framework (EQF) – the structure of qualifications levels adopted in the EU, constituting a reference system of the national qualifications framework, enabling the comparison of qualifications obtained in different countries. The qualifications framework defines the qualifications obtained in the education system and their mutual relations, defining e.g. learning pathways, making it easier to compare qualifications acquired at different times, places and forms. The qualifications framework aims to adapt competencies to the needs of the labour market and increase employee mobility, promoting and facilitating lifelong learning.

The primary element of the framework is the learning outcomes, defining what the learner knows, understands and can do after completing the learning process. In other words, it is a language for the description of competencies to be understood by all stakeholders in the learning process. The scheme for describing the levels of qualifications is based on identifying three categories of learning outcomes: knowledge, skills and competencies, according to which specific characteristics of learning outcomes are composed. Competencies in this tripartite are understood as the proven ability to apply knowledge, skills, personal, social or methodological abilities demonstrated at work or study and professional and personal careers. In the EQF, competencies are defined in terms of responsibility and autonomy (in the Polish competence framework, the term ‘social competencies’ is used). The first category describes the expected range of theoretical and factual knowledge. Skills are described as cognitive (involving logical, intuitive and creative thinking) and practical (relating to manual dexterity and using methods, materials, tools and instruments). Responsibility and autonomy are described as the learner’s ability to apply knowledge and skills independently and responsibly (Europass, n.d.).

A detailed elaboration of the possible variants of the competency framework is quite a significant endeavour. Here, we limit ourselves to indicating essential areas and issues to consider when developing such a framework. A Data Steward, in a nutshell, is someone who understands the value of data as a resource, who knows about the problems of information systems in the organisation, who wants to question the status quo and introduce changes, who has the necessary communication skills, and is a person with high credibility, above all thanks to the unique competencies in the area of data governance.

The scope of the expected competencies can be divided into three main groups: technical competencies, non-technical (business) competencies and soft competencies. The scope of technical competencies should primarily cover the following areas: knowledge management, data processing, data archiving, infrastructure and network, and services. This list can be further elaborated. For example, these should include a good understanding of data and information concepts, data modelling skills (conceptual, logical and physical), knowledge of the organisation’s data management technologies and systems, and the related tools. Non-technical skills refer to strategic planning and change management, project management, compliance and legal environment, and a deep understanding of the organisation’s foundations, functions, goals and environment. While often underestimated, soft traits and skills are particularly important in the case of Data Steward. Here, one can indicate creativity, independence, teamwork, listening skills, openness and diplomacy, objectivity, communication and persuasion skills, networking skills,

patience, calmness, and composure. To sum up, the competence of a Data Steward is about the careful and responsible management of the data ecosystem.

## 8. Conclusions

Currently, businesses are confronted with a fairly complex interplay between the existing and forthcoming EU data laws. From a business perspective, practical guidance is needed to help enterprises navigate through this complexity and meet their legal obligations while harnessing the power of data. A recognised Data Steward function should be promoted in both the public and private sectors, along with developing digital skills and capacity building to ultimately create a data-sharing culture based on the principle of reciprocity (EC, 2020e). Trained Data Stewards are needed to increase capacity, both on the supply and demand side of data, to democratise access to the data and facilitate the growth of data-driven solutions within functional data ecosystems. Data Stewards are at the heart of data governance efforts.

Technology and regulation can only solve some problems of the lack of adequate human resources, expertise or poorly defined functions. Data Stewards must take on diverse roles, tasks and responsibilities, aligning data processes and applications in developing, and enforcing data governance in compliance with regulations and data ethics. Implemented and planned legislation requires support through designing education programmes in the field of new roles necessary for the efficient functioning of data ecosystems, particularly the role of a Data Steward. Unfortunately, the current state of data stewardship education and training, especially in the public sector, is in its infancy. Future education in data stewardship should be built on a solid collaboration between the private and public sectors and the academia to ensure relevant competencies in documenting, curating and structuring data across public and private organisations, facilitating data sharing and use in an increasingly regulated environment.

We argue that democratic countries with highly effective state authorities should conduct an active policy on creating the state's information order and provide citizens and business entities with adequate information as a public good, regardless of whether the information is needed. Thanks to this, information also reaches these citizens or entities who do not realise what information they need. In this way, it effectively minimises a social information gap and increases social knowledge resources (Oleński, 2000). The right policy should focus on developing infrastructure and the knowledge, skills and attitudes necessary for the smooth functioning of data ecosystems or (more generally) public information governance. Skills and habits for using information and stock of derived knowledge which



a citizen, an enterprise or the public administration can use constitute an essential aspect of the country's socio-economic development. With the exponential increase in the potentialities of generating, collecting and processing information in all areas of socio-economic life, this puzzle's critical element becomes the Data Steward's role.

## References

- Bell, D. (1973). *The Coming of Post-Industrial Society*. Basic Books.
- Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (OJ EU L 172, 26.6.2019). <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024>.
- Erstad, O. (2010). Educating the Digital Generation. *Nordic Journal of Digital Literacy*, 5(1), 56–71. <https://doi.org/10.18261/ISSN1891-943X-2010-01-05>.
- Europass. (n.d.). *Description of the eight EQF levels*. Retrieved April 3, 2021, from <https://europa.eu/europass/en/description-eight-efq-levels>.
- European Commission. (2014a). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Towards a common European data space*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0232>.
- European Commission. (2014b). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Towards a thriving data-driven economy*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014DC0442>.
- European Commission. (2020a). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A European strategy for data*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066>.
- European Commission. (2020b). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Shaping Europe's digital future*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0067>.
- European Commission. (2020c). *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825>.
- European Commission. (2020d). *Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act)*. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=COM:2020:767:FIN>.
- European Commission. (2020e). *Towards a European strategy on business-to-government data sharing for the public interest. Final report prepared by the High Level Expert Group on Business to Government Data Sharing*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2759/731415>.

- European Commission & Directorate-General for Research and Innovation. (2016). *Management Plan 2016*. [https://commission.europa.eu/system/files/2016-05/management-plan-2016-dg-rtd-march2016\\_en.pdf](https://commission.europa.eu/system/files/2016-05/management-plan-2016-dg-rtd-march2016_en.pdf).
- Horton, F. W. (2007). *Understanding information literacy: a primer*. United Nations Educational, Scientific and Cultural Organization. <https://www.ifla.org/publications/understanding-information-literacy-a-primer/>.
- MacKay, K., & Vogt, Ch. (2012). Information Technology in Everyday and Vacation Contexts. *Annals of Tourism Research*, 39(3), 1380–1401. <https://doi.org/10.1016/j.annals.2012.02.001>.
- Oleński, J. (2000). *Elementy ekonomiki informacji. Podstawy ekonomiczne informatyki gospodarczej*. Uniwersytet Warszawski.
- Organisation for Economic Co-operation and Development. (2016a). *Skills Matter. Further Results from the Survey of Adult Skills*. OECD Publishing. <http://dx.doi.org/10.1787/9789264258051-en>.
- Organisation for Economic Co-operation and Development. (2016b). *G20/OECD INFE Core Competencies Framework on Financial Literacy for Adults*. <https://www.oecd.org/finance/Core-Competencies-Framework-Adults.pdf>.
- Organisation for Economic Co-operation and Development. (2018). *Financial Markets, Insurance and Pensions. Digitalisation and Finance*. <https://www.oecd.org/finance/Financial-markets-insurance-pensions-digitalisation-and-finance.pdf>.
- Organisation for Economic Co-operation and Development. (2020). *OECD Digital Economy Outlook 2020*. OECD Publishing. <https://doi.org/10.1787/bb167041-en>.
- Peng, G. (2018). The State of Assessing Data Stewardship Maturity – An Overview. *Data Science Journal*, 17, 1–12. <http://doi.org/10.5334/dsj-2018-007>.
- Punie, Y., Brečko, B., & Ferrari, A. (2013). *DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe*. Publications Office of the European Union. <https://publications.jrc.ec.europa.eu/repository/handle/JRC83167>.
- Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC (OJ EU L 257, 28.8.2014).
- 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) (OJ EU L 119, 4.5.2016).
- Sen, H. (2018). *Data Governance. Perspectives and Practices* (1st edition). Technics Publications.
- Versweyveld, L. (2016). *We need 500,000 respected data stewards to operate the European Open Science Cloud*. <http://e-irg.eu/news-blog/-/blogs/we-need-500-000-respected-data-stewards-to-operate-the-european-open-science-cloud>.