Gergely Pálmai – Szabolcs Csernyák – Zoltán Erdélyi

Authentic and Reliable Data in the Service of National Public Data Asset

Summary: The analysis focused on how efficient management of the national data asset is supported by the Hungarian regulatory framework concerning the use of public information, and whether public data constituting part of the national data asset can be deemed authentic and reliable to support the efforts for the digitalisation and artificial intelligence-based developments of the public sector. The analysis shows why the availability of authentic and reliable data in terms of the national data asset has outstanding significance. In support of this assertion, it presents the different levels of data asset use, the role of using artificial intelligence in the public sector, and the significance, risks and challenges of the authenticity and reliability of public data, from both a data protection and a public finance aspect. Inaccuracy, unreliability of input data predestines the generation of incorrect result products (conclusion, decision), even if the appropriate algorithm is used, which could lead to direct financial loss, for both the citizens and the state. The authors of the analysis therefore suggest that a paradigm shift is necessary in the strategies targeting the efficient use of the public sector's data, with the necessity to record the fundamental precondition that the national data asset must be based on reliable and authentic data.

Keywords: public sector, digitalisation, artificial intelligence, national data asset, efficient management, reliable and authentic data
JEL codes: D73, D80, D81, D89, H41, H76, H89, K23, L38, L51, O33, O38
DOI: https://doi.org/10.35551/PFQ_2021_s_1_3

The present analysis demonstrates why the availability of authentic and reliable data has outstanding significance in terms of the national data asset and in the interest of implementing the strategic target system supporting the efforts for the digitalisation of the public sector and artificial intelligence-based developments. Consequently, it presents the different levels of data asset use, the role of using artificial intelligence in the public sector, and the significance, risks and challenges of the authenticity and reliability of public data, from both a data protection and a public funds aspect. Regarding the empirical research tools, the authors chose the method of analysis to process the topic.

E-mail address: szabigvezeto@asz.hu

LEVELS AND IDENTIFIED RISKS OF DATA ASSET MANAGEMENT AND PUBLIC DATA USE

Organisations in both the private and public sectors perform the management of all data received by them, generated by them or forwarded from them, i.e. the 'data asset', pursuant to the legal framework and subject to their activity and range of duties as some kind of 'stock' (Péterfalvi, 2017, pages 263-264). Organisations performing public service are entitled and obliged to process data types of extremely large diversity in many data categories, regardless of whether they are data stored electronically or in paper-based documents. Accordingly, the subject of the data management of organisations performing public service can be classified data, personal data and public data, such as data of general interest and data accessible on general interest grounds.

Pursuant to Article VI of the Fundamental Law, everyone has the right to the protection of his or her personal data, as well as to assess and disseminate data of public interest. Pursuant to Article 39 of the Fundamental Law, data relating to public funds and national assets are data of public interest. Access to public information can be considered a certain pledge, a fundamental pillar of the rule of law, in connection with which freedom of information only exists if *'everyone has free access to the information of the public sector; this right can only be restricted by law, in a limited manner'* (Székely, 2015, page 40).

The access to data of public interest can also contradict the data security requirement. Certain institutions of the public sector must meet extremely high information security requirements pursuant to the relevant law. The requirement of meeting high-level information security conditions – with merely formal significance – is in contrast with ensuring wide-ranging access to data of public interest.

The paradox is due to the nature of 'competing' fundamental rights incorporated in the Fundamental Law – data asset security and data protection versus transparency of data of public interest. The restriction of the fundamental right incorporated in the Fundamental Law is only possible within constitutional frameworks pursuant to the conditions provisioned in Article I (3) of the Fundamental Law, i.e. by way of legal regulation, in the interest of the enforcement of other fundamental right or the protection of some constitutional interest, to the absolutely necessary extent, in proportion to the objective pursued, with the material content of the fundamental right observed. The right to assess and disseminate data of public interest is a fundamental right provisioned in Article VI of the Fundamental Law, concerning which the restriction of assessing data of public interest is only possible under expressly strict conditions within the effective legal frameworks and based on the effective legal practice. Rejection of request to assess data of public interest by a body performing public service is only possible for reasons specified by the law, in a limited range.

Pursuant to the Information Act in effect, request for data of public interest can be initiated by anyone without interest, implication; anyone can apply for assessing data of public interest. In connection with the regulation of request for assessing data of public interest, the Information Act does not require the specification of the purpose of the application; pursuant to the provisions of the law, assessing data of public interest must be ensured regardless of the purpose of the applicant. Based on the explanation of the Information Act, the purpose of the legislator is to ensure the access to data of public interest against the controller having information

monopoly. Pursuant to the relevant legal practice of the data protection authority, '*the purpose of requesting data of public interest is irrelevant in the cases of requesting data of public interest (motivation of the data request) […] according to the Authority, in the course of granting the data request, it cannot be examined if the applicant exercises his or her fundamental right as intended, and what is the purpose of the application. Data requests cannot be rejected on the basis of misuse of the law*' (Péterfalvi, 2014).

All this can result in individual data requests, occurring on a massive scale in practice, leading to misuse-type legal practice, and presenting significant additional administration challenges to the organisations fulfilling requests of data of public interest.

In its civil decision of principle No. 16/2013, the Curia of Hungary established that the exercising of the fundamental right related to requesting data of public interest must be legitimate, in the case of exercising the right with misuse, granting the data request can be legally denied (judgement No. EBH2013 P16). According to the interpretation of the data protection authority – related to the referenced judgement –, *'in the case of every single data request, the existence of the misuse of law must be examined and the priority of public interest related to rejection (elimination of unlawful legal practice against the principle of the Civil Code) must be weighed independently. Otherwise, the organisation performing public service could deprive a particular applicant of exercising his or her fundamental right if it is assumed from the beginning that regardless of its content and subject, the application for data effects misuse*' (Péterfalvi, 2020).

The above illustrate the narrow boundaries within which the access to data of public interest can also contradict the requirement of data security. In addition to the access to data of public interest, the constitutional interest related to the protection of national data asset can also be deduced from the Fundamental Law. Data – including the totality of data of public interest, personal data and data accessible on public interest grounds – processed by the institutions of the public sector, i.e. the bodies performing public service enjoy outstanding protection by the law; they constitute part of the so-called national data asset. Pursuant to the law, the national data asset is deemed a national asset, an asset element belonging to the national assets, and as such, enjoys outstanding constitutional protection pursuant to the Fundamental Law, including the requirements of the protection of national assets and responsible management of the national assets [Article 38 of the Fundamental Law, Section *i)* of Article 2 of Act CXCIV of 2011 on National Assets]. Provisions ensuring protection of the data do not apply to persons holding data removed from this protected circle by application for assessing data. All this has an adverse effect on the information security efforts of the public sector's organisations in terms of data confidentiality and integrity.

Several levels and dimensions of data asset management are known, including internal use of data within the organisation, forwarding data to and sharing data with external organisations, as well as national and international use of the organisation in broader dimension. Data asset management has an economic aspect, namely the question of efficient management of data asset as stock. Regarding laws concerning information, '*the key to appropriate data asset management is lawful, planned and secure data processing"* (Péterfalvi, 2017, page 264).

State data asset use has several dimensions, however, in the absence of long-term data asset strategy document system, state data asset is considered 'uncut diamond' because despite the fact that the state can be deemed the biggest data owner, *'significant part of the*

state data asset is still not exploited today; it lies unused' (Schopp, 2020). Part of this is linked with the deficiency of the related regulatory environment that 'data asset' is not defined by the effective laws.[1]

Data constituting elementary part of the data asset are typically approached by the practice via the protection relation of data or the concept relation of personal data (whether data include personal data or not). Some try to define the concept of data asset with the tools of copyright, and attempt to deduce it from the concept of business secret. It is more likely, however, that the future legal concept of data asset will be regulated as a completely new, independent legal institution, after which the concept system of the data asset can also be established.

Further problems are posed by the disintegration of data asset-related tasks and powers, and the fact that the precise number of administrative records is not known, the degree to which the data sets therein have been processed is low.

The transparency and use of state data asset are still hindered by the fact that there is still no publicly available record of the state data asset in Hungary. This is despite that, the Government decision on the Digital Welfare Programme of Hungary prescribed the obligation of the overall survey of the public data asset and the preparation of the public data cadastre with a deadline of 31 March 2017 [pursuant to Article 7/*g)* of Government decision No. 2012/2015. (XII. 29)]. The strategic necessity of establishing a national cadastre was already specified in the Digital Welfare Programme (DJP) published in 2015 (DJP, 2015).

The absence of publicly available records of the state data asset significantly hinders the re-use of the public sector's data, and public actors often collect and generate data in parallel. In addition to the businesses potentially interested in re-use, without a state data asset cadastre, the bodies of the public sector cannot see in depth the range of public data they could use from data generated or collected by other bodies (Börcsök et al., 2019, page 67).

Another problem is the undervaluation of data. There is no precise calculation method concerning the calculation of the data asset's value and its settlement in the course of its use (Schopp, 2020),

The intention of the government to establish a detailed legal background for the data asset – by creating the Data Asset Framework Act[2] – in the near future is considered a positive step. Further positive development is that the Artificial Intelligence Strategy approved by government decision in September 2020 also includes strategic goals concerning data asset (AI Strategy, 2020).

In connection with the broader dimension of managing public data, the PSI directive on further use of public sector information and the act on the re-use of public data adapting it to the Hungarian legal environment (Public Data Act) laying down the foundations of further use of data of public interest have significant relevance.

USE OF PUBLIC DATA, ITS 'RE-USE' AS OPEN DATA

The European Union also recognised the potential hidden in the data of public administration, as one of the goals of its open data strategy announced in the interest of enforcing efficiency viewpoints is to facilitate the secondary – market-based – use of public data of public administration bodies unexploited to date. Accordingly, the data economy building strategy and the artificial intelligence strategy of the European Union also facilitate, among others, wider access to and efficient use of public data or open data.

Considering its objective system, the PSI directive and the Public Data Act, which adapted the former to the Hungarian legal system in 2012, do not target the transparency of the public sector or the strengthening of free assessment of information generated by the public sector. Instead, they wish to ensure further use of public data – primarily for market or business purposes – based on uniform EU regulatory frameworks, and through it, the wider and more efficient use of the public data asset.

The so-called open data strategy of the EU issued in 2011 is worth mentioning as a historic antecedent prior to the adaptation of the PSI directive to the Hungarian legal system. The open data strategy of the EU served the implementation of the central objective of the Europe 2020 strategy, the goal of which was to put the European economy on a strong and sustainable growth path. According to the strategy, *'increase of the European innovation potential and more efficient exploitation of the available sources are necessary to achieve this objective'*, which, from among the group of sources requiring exploitation, primarily named public data. Public data consist of information generated, collected or purchased by public bodies on the territory of the European Union. Pursuant to the open data strategy of the EU, *'making these sources publicly available – in the interest of increasing the efficiency of new products, services or the efficiency of public administration bodies – could result in an annual EUR 40 billion economic profit in the European Union'* (EU Open Data Strategy, 2011). How much from this planned economic growth was actually realised by 2020 is outside the framework of the topics examined by the present analysis, however, the quantitative plan figures shed light on the weight and significance of potential economic advantages to be gained by the exploitation of open data. The European Commissioner of the European Commission responsible for the uniform digital market issued an approximately annual EUR 12 billion worth of economic forecast for the coming decade related to the use of open data, according to which *'the total direct economic value of the public sector's information and data originated from the public undertakings is expected to grow from annual EUR 52 billion in 2018 to EUR 194 billion by 2030'* (Ansip, 2019).

The data economy building strategy of the European Union announced in 2017 identified the so-called data localisation requirements prescribed by the Member States in connection with the public administration bodies, which may restrict the free flow of data within the EU, as a factor hindering the growth of the data economy of EU. As an example of such data localisation requirements, the document mentions the rules generally prescribing the local storage of archived data generated in the public sector (EU Data Strategy, 2017). In this respect, it should be noted that the regulation concerning the Hungarian electronic information systems of the Hungarian public sector also prescribes such data localisation requirements, which, for the purpose of enforcing increased data security of the data, restricts the storage and operation of national data asset elements and electronic information systems outside the territory of Hungary [pursuant to Article 3 (1) of Act L of 2013 on Electronic Information Security of State and Local Government Bodies, hereinafter referred to as: Information Security Act]. Despite the potential obstacles, the 2017 data economy building strategy forecasted significant growth in the value of the EU data economy, the estimated value of which was 1.85% of the GDP of EU in 2014, as opposed to its estimated value of 3.17% of the EU GDP by 2020 (EU Data Strategy, 2017).

In the interest of the better use of public and

private data, the new European data strategy of the EU published in 2020 also envisages similar data economy development potential. As an ambitious target, the new European data strategy specified that, through the appropriate policies and investments of the Member States and businesses, the Commission invests a total of EUR 4-6 billion in common European data areas and the European integration of cloud-based infrastructures and services (EU Data Strategy, 2020).

In addition to the re-use of public data as open data, the application of new technologies supporting digital transformation, such as the use of artificial intelligence (AI), also promises significant economic and social use. With regard to the economic significance of AI technology, the European Union budgeted an EUR 20 billion annual investment target in connection with the artificial intelligence technology for the coming decade (AI Coordinated Plan, 2018).

AI IN THE WORLD OF PUBLIC DATA

The difficulty of defining artificial intelligence – as a concept – is also shown by the number of literatures offering solution for this problem. Within the framework of the present analysis, we consider the following definition in the Artificial Intelligence Strategy of Hungary issued in 2020 as standard: *'Artificial intelligence is a piece of software capable of mapping parts of human intelligence and supporting or autonomously performing processes of perception, interpretation, decision making and action'* (AI Strategy, 2020, page 9).

Digital data revolution and the new technologies supporting it – among others, the use of AI – have an undoubtedly increasing influence on our everyday lives, from which the organisations of the public sector and the public services are no exceptions.

In our data-driven digital economy, the question in connection with the use of AI 'conquering' the world of public data all at once is not whether the use of AI in the public sector is justified, but how the use of data also concerning the public sector can be supported more efficiently with AI technologies.

In the interest of exploiting the opportunities given by AI and managing the resulting challenges, the European Union announced the necessity of a specific 'European approach' in its 2018 AI strategy (AI Strategy, 2018) and in its White Book on artificial intelligence. Among others, the 'anthropocentrism' of AI, the necessity of building 'trust' in the technology and building the European AI sector on values and fundamental rights, such as the protection of human dignity and privacy, are incorporated in this European approach (AI White Book, 2020).

Notably, the legal and technological regulatory frameworks of the use of AI are not mature yet either in Hungary or in the European Union. In the interest of establishing the ethical and legal regulatory frameworks of AI, in October 2020, the European Parliament published a recommendation on the 'ethical and legal challenges' of AI development, without binding legal force for the Member States, in which it is emphasised that technology could not develop at the expense of the humankind and the protection of intellectual property rights and patents, private persons and business with civil liability (AI Recommendation, 2020).

No legal framework system for the use of AI has yet been established in Hungary, however, the related strategic document system is available. A significant progress in the field of using AI in Hungary was the Artificial Intelligence Strategy of Hungary being approved by government decision in September 2020. The Hungarian AI strategy was approved with the objective that *'the citizens,*

businesses and sectors of public administration could prepare for the changes caused by artificial intelligence and could exploit its advantages' [as regulated in Government decision No. 1573/2020 (IX. 9.)]. Digital transformation concerning data of the national data asset in the Hungarian public sector reached another milestone by incorporating the use of AI in the national strategic framework system.

The government decision approving the AI strategy of Hungary established the institution system supporting the implementation of the strategic target system, provisioning, among others, the establishment of an Artificial Intelligence Innovation Centre, Artificial Intelligence National Laboratory and Hungarian Data Asset Agency.

The question of what conditions are necessary in using AI efficiently in the public sector arises. According to the professional head of the Digital Welfare Programme, *'artificial intelligence will be used well in Hungary if data economy is established and the subjective and material frameworks of data asset are specified for it'* (Gál, 2020). In keeping with this, the building of data economy is essential, for which the relevant legal frameworks, such as the concept of data asset, must be rethought.

Data are the foundation of fulfilling the goals of AI and the necessary computer learning . According to the professional head of the Digital Welfare Programme, *'artificial intelligence cannot learn without data, therefore it needs clean information for its operation'* (Gál, 2020).

In our opinion, good quality, reliable data, in addition to clean information, are also necessary to achieve the wished operation, as the use of an unsuitable data set in the process of computer learning may lead to undesired goals and my cause serious damage. This is also specified by the so-called GiGo (*'Garbage in, garbage out'*) principle, one of the fundamental laws of informatics, according to which bad data will give bad results.

This is supported by the fact that several EU Member States deem the existence of reliable data of strategic importance for the wide ranging and secure use of AI. Danish AI strategy specifies the following: *'The progress of artificial intelligence is subject to the quality and quantity of data'* (Danish AI Strategy, 2019). The strategic document of Germany drawn up in relation to AI states the following: *'Regarding AI and the methods of computer learning, availability and quality of data are central conditions and determining factors of the results' quality'* (German AI Strategy, 2018).

The question of when we can deem data reliable arises. In order to answer this question, approaches from both data security and public funds viewpoints are worth taking.

RELIABLE DATA REGARDING DATA SECURITY

'Data' constituting the basis of digitalisation in the public sector enjoy outstanding legal protection compared to private sector's data, because, in addition to the data protection legal regulations enforced in the public sector, controllers must comply with extremely strict data security legal requirements.

Concerning the public sector, information security is provisioned by the Information Security Act and its acts of implementation. The basis of the outstandingly high data protection and data security requirements of the public sector's organisations is a social requirement according to which the security and protection of data in the relation between the citizen and the state – including personal data of the citizens (!) – must have outstanding priority with regard to the institution protection obligation of the state.

In this context, the strict legal requirement of the public sector's organisations is the provision of closed, complete, continuous

and risk-proportional protection of the information related to information systems and data stored therein. The protection of personal data therefore cannot be implemented without meeting the information security requirements.

Concerning data security, the 'reliability' of data is ensured by the enforcement of the information security requirements. With regard to electronic information security, in the case of the information technology systems and applications of the state and local government bodies, the enforcement of the threefold principle incorporated in the Information Security Act has outstanding importance.

The Act adapted the spirituality of Standard No. ISO/IEC 2700.1 constituting the foundation of information security in the competitive sector. Accordingly, the measures guaranteeing data protection are built around the following three fundamental categories *(see Figure 1):*

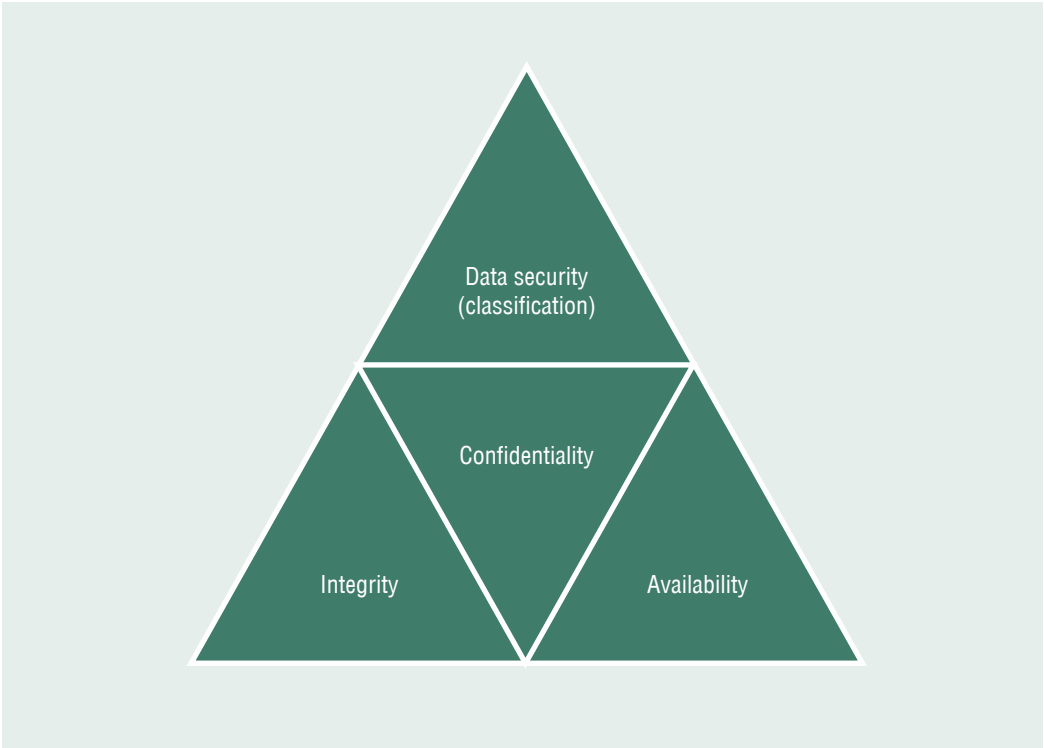
- confidentiality,
- integrity,
- availability.

Confidentiality is the property of the data ensuring that it can only be known, used by authorised parties, only they can make decisions about is, , and only to the extent of their authorisation.

Integrity means that the content and properties of the data are identical with the expected content and properties, including the certainty that it originates from

Figure 1

DATA SECURITY – VIEWPOINTS OF DATA CLASSIFICATION



Source: own edition

expected source (authenticity), as well as the possibility of verifying its origin, certainty (nonrepudiation).

Availability means that authorised parties can access the wished data at a specified time and for a specified period.

In the case of the protection of data – as well as their integrity –, in harmony with the international standards and good practices, the Information Security Act subjects the protection measures to be set up to the classification of data per security class and to the classification of organisations per security level. Both in terms of the security level and security class, the Act requires the use of a scale from 1 to 5, the classification in which must be performed based on risk analysis. The set-up of protection measures proportional with the risks fulfilling the goal specified in the law can thus be ensured. In the absence of this fundamental act, the given organisation only 'shoots blindly', as it cannot identify the requirements deemed necessary by the Act and therefore it cannot take action to establish them appropriately either.

On that basis, compliance with the Information Security Act is essential in terms of assessing the reliability of data generated and processed in the public sector, the responsibility of which lies with the data owners of the data. Clear definition of data owner roles within the given organisation, therefore, has outstanding significance, which – as we have mentioned it before – is also an important precondition of using AI in the public sector.

Concerning data security, data are 'reliable' if the integrity of data (confidentiality, integrity, availability) is ensured within an organisation, therefore, due to the closed, complete, continuous and risk-proportional protection of information, the risk of data vulnerability is negligible.

One of the most significant challenges of providing data necessary for the operation of AI and the reliability of data is how to ensure the reliability of data once they are removed from the responsibility of the data owner.

Audit experience regarding the practical operation of the Hungarian framework system of data protection of the State Audit Office of Hungary is also available. In the course of the 2017 audit of the Hungarian data protection framework system and certain priority data bases, the SAO found that the fact that the audited organisations did not always perform the classification of their electronic systems used for data processing and the classification of the entire organisation per security class and security level pursuant to the provisions was a vulnerability risk from a data protection viewpoint (SAO report (2017). In 2020, within the framework of a subsequent audit, the SAO audited the implementation of the tasks specified in the action plan of the audited organisations related to the findings made in its referenced report. Concerning the audited organisations, the subsequent audit found an overall improvement in the field of data processing security, however, at the same time, part of the risks related to performing data protection and data security supervisory tasks still exists (SAO report, 2020).

All of this suggests that there is still room for improvement in the interest of increasing data integrity of the state information systems and state data asset before we can speak of the availability of 'reliable' data in connection with the data asset elements.

In conclusion, concerning the information security of state and local government electronic information systems, it can be established that in the absence of compliance with the referenced principles, the integrity of data, i.e. their reliability and authenticity, processed and stored in the given information technology system is not ensured. This questions whether fundamental requirements of the system are fulfilled.

ON THE PUBLIC FUNDS ASPECTS OF AUTHENTIC AND RELIABLE DATA

In its Article 39, the Fundamental Law specifies that every organisation managing public funds is obliged to publicly account for its management of public funds. Public funds and national assets must be managed according to the principles of transparency and the purity of public life. Under the extraordinary circumstances experienced during the Covid–19 pandemic, the role of increased enforceability of the principles of the accountability and transparency of public funds, the necessity of their enforceability and guaranteed provision regarding freedom of information have gained more weight.

In the case of certain public sector data, the law quasi presumes the 'reliability' and 'authenticity' of the data. Such are the so-called 'official public registers', the purpose of keeping which is to offer authentic proof of confirmation of data contained by them [pursuant to Section *b)* of Article 97 (1) Act CL of 2016 on General Public Administration Procedures (Ákr.)]. In connection with these records, the authenticity of data contained therein can be refuted in retrospective based on public or court proceedings, i.e. refutable legal protection exists, that the included data are authentic. Based on the relevant legal provision pending proof to the contrary, data contained in the official registers must be presumed to exist, and data deleted from the official registers must be presumed not to exist [Pursuant to Article 97 (2) of Ákr.].

The authenticity of data contained by these official registers is recognised by the force of the law. As a general rule, it must be presumed that a party acquiring certain rights relying upon the data obtained from an official register was acting in good faith [Article 97 (2) of Ákr.].

The fundamental principles for the registration of legal persons related to the authenticity of registers are provisioned by Article 03:13 of the Civil Code. According to Section (1), all entries made to the register of rights, facts or data must be evidenced by a document, court or administrative decision specified by law. The definition of the authenticity of the register is included in Section (2), pursuant to which the rights, facts and data it contains must be presumed to exist and to be authentic.

The issue of the 'authenticity' and 'reliability' of data content of registers without authenticity relating to the asset elements of state data asset needs special attention. In the course of its audits, the SAO experienced the absence of authenticity and reliability of the audited data on several occasions. Several findings of the SAO audits include the fact that the audited organisations do not perform their obligation related to the publishing of the annual report prescribed by the law, or do not perform it appropriately. In several cases, the SAO also found that the reports of the audited organisation were published without the legal action of the body authorised to approve the report. It is a deficiency regularly revealed in the course of the SAO audits that the audited organisation attempts to verify its reporting obligation not with the authentic report signed by the organisation but with the report without signature recorded in the company information service, the authenticity of which is not ensured in the present regulatory environment.

The relevant regulations, such as the Accounting Act and the separate regulations on the preparation of accounting report, prescribe that the accounting report must be signed by the person authorised to sign by the law and must be approved by the body authorised to approve it.

Regardless of whether the register is authentic or not, pursuant to the legal regulations in effect, the organisations

performing tasks related to the publishing of the reports do not have the expressed task and authority to confirm the authenticity of the reports' data content ex officio. The examination of the reports submitted for publishing is not ensured in the sense of whether data included therein are reliable.

There is demand for the establishment of a regulatory environment for the principle of transparency and accountability of public funds incorporated in the Fundamental Law, the enforcement of the conditions of fair economic competition specified in the law, the security of economic turnover and the protection of the creditors' interest to examine if the accounting reports of the organisations submitted for publishing are reports signed and approved by the authorised persons, which guarantees data included in them. This is particularly relevant in cases where the register is deemed an authentic database pursuant to the law.

The SAO experiences related to the publishing of reports referred to above shed light on the outstanding significance of the reliability and authenticity of data asset elements in connection with the registration of state data asset.

Concerning the recording of national data asset, the absence of availability of authentic and reliable data may have the consequence of including unreliable and non-authentic data in certain official public registers. In this case, the fundamental authenticity of official public registers can be questioned, contradicting the legal protection, according to which official public registers authentically prove by the force of the law that data contained therein, the registered rights and facts, and their modifications exist pending proof to the contrary.

If we use data in such database – giving the impression of being a database containing authentic data – with AI technology, it will certainly have adverse effects, because – as we already mentioned it – *'artificial intelligence cannot learn without data, therefore it needs clean information for its operation'* (Gál, 2020).

A further challenge regarding the availability of authentic and reliable data is posed by the fact that the organisations operating in the central system of the general government and in the subsystem of the local governments must meet extremely strict compliance requirements resulting from dual sector (data security, data protection and accounting, professional and general government management) regulations.

It was experienced on several occasions in the course of the SAO audits that the information technology system of the audited organisation was not suitable to ensure the generation of annual budgetary reports giving a reliable and authentic picture in compliance with the professional accounting and general government management-related sector requirements.

In this respect, the effective regulation, pursuant to which the certification of the compliance of document management software tools usable by organisations performing public service is carried out by so-called certifying bodies not holding official authority, can be identified as a good practice. The certifying body issues a certificate on the compliance check, in which it certifies that a given data management software tool meets the conditions of use prescribed in a separate regulation.

In connection with information technology systems supporting the availability of authentic and reliable data, the issue arises that similarly to the certification of the utilisation conditions of data management software tools usable at organisations performing public service, the compliance of the state information technology system with public funds compliance requirements should also be certified with a certificate.

According to Article 38 (1) of the Fundamental Law, the legal prescription and implementation of the certification procedure concerning the public funds compliance requirements of state and local government information technology systems would significantly support the practical enforceability of the requirement of the responsible management of national assets.

Based on the presented risks and challenges, a paradigm shift is necessary regarding the processing of the public sector's data, and the preparation of strategies aimed at it, with the fundamental condition provisioned that the national data asset must be based on data which can be deemed reliable and authentic. A more efficient use of public data or open data and its support with artificial intelligence developments can exclusively be built on authentic and reliable data asset elements.

Until it is fully ensured that the records contain reliable and authentic data in connection with the asset elements of state data asset registers – due to the deficiencies of the regulatory environment and necessary controls built in the process –, we cannot speak about efficient management of the data asset as 'stock'. Therefore, concerning AI, such type of data are an increased risk, which may produce the GiGo effect repeatedly as a result.

SUMMARY, CONCLUSION

The more efficient use of the national data asset's data is in the focus of supporting the use of data processed in the public sector with artificial intelligence technology. The potential hidden in the data of the public sector was also recognised by the European Union, since one of the goals of its open data strategy announced in the interest of enforcing efficiency viewpoints is to facilitate the secondary – market-based – use of public data of public administration bodies unexploited to date. Accordingly, the data economy building strategy and the artificial intelligence strategy of the European Union also facilitate wider access to and efficient use of data essential for the use of artificial intelligence, among others.

Despite all this, it is emphasised less that in order to achieve strategic goals concerning the use of artificial intelligence, there is need for primarily reliable and authentic data preceding any efficiency issue. Regarding the integrity of data processed in the public sector, the development of digital economy and the artificial intelligence developments, it is essential that the national data asset is built on secure foundations, reliable and authentic data.

In the course of its audits, the SAO experienced the absence of authenticity and reliability of the audited data on several occasions. In the course of the audit of the Hungarian data protection framework system and certain priority data bases, the State Audit Office of Hungary found that the fact that the audited organisations did not perform the classification of their electronic systems used for data management and the classification of the entire organisation per security class and security level appropriately was a vulnerability risk regarding data protection.

With regard to electronic information security in the case of the information technology systems and applications of the state and local government bodies, the enforcement of the threefold principle – integrity, confidentiality and availability – incorporated in the act on information security has outstanding importance. In the absence of this the integrity, i.e. the reliability and authenticity of data processed and stored in the given information technology system, is not ensured. This questions whether the fundamental requirements of the system are fulfilled.

Additionally, in the course of its audits, the SAO also found that certain electronic databases – such as databases processing the reports of companies and other organisations – contained unreliable, non-authentic data. The reason for this is that the databases are organised without regard to the accounting-professional legal provisions, which is a material risk regarding reliability and authenticity.

Concerning the recording of national data asset, the absence of availability of authentic and reliable data may have the consequence of including not valid and non-authentic data in certain official public registers. In this case, the fundamental authenticity of official public registers can be questioned, contradicting the legal protection according to which official public registers authentically prove by the force of the law that data contained therein, the registered rights and facts, and their modifications exist pending proof to the contrary.

In conclusion, a paradigm shift is necessary regarding the strategies targeting the efficient use of the public sector's data, with the necessity to record the fundamental precondition that the national data asset must be based on reliable and authentic data.

A more efficient use of public data or open data and its support with artificial intelligence developments can exclusively be built on authentic and reliable data asset elements. By projecting the principle applied to the world of informatics and mathematics to artificial intelligence, according to which bad result is obtained from bad data, the fact that the inaccuracy, unreliability of input data predestines the generation of incorrect result products (conclusion, decision), even if the appropriate algorithm is used, can be demonstrated well. This – in the case of assessing an application, for example –- may result in direct financial losses both for the citizens and the state.

Notes

[1, 2] At the time of the closing of the manuscript, the bill T/14949 on the national data assets was under discussion in Parliament.

References

Ansip, A. (2019). Press release – Digital single market: EU negotiators agree on new rules for sharing of public sector data In: EC Europa.eu honlap [EC Europa.eu website], 22 January 2019, https://ec.europa.eu/commission/presscorner/detail/en/IP_19_525

Börcsök, S. (2019). Adatpolitikai stratégiai javaslat az MI-alapú innováció beindítására Magyarországon. [Data policy strategy proposal for the launch of AI-based innovation in Hungary.] 15 June 2019, page 67, https://www.magyary.hu/wp-content/uploads/2019/07/AdatpolitikaiStrate%CC%81giaiJavaslat.20190627.Magyary.pdf

Gál, A. L. (2020). Elindult a Moór Gyula Digitális Jog- és Államtudományi Szakkollégium előadássorozata: fókuszban az adatvagyon és annak

szabályozása. [Series of lectures were launched by the Gyula Moór Digital Law and Political Science College for Advanced Studies.] In: *DJP honlap [DJP website]*, 19 November 2020, https://digitalisjoletprogram.hu/hu/hirek/elindult-a-moor-gyula-digitalis-jog-es-allamtudomanyi-szakkollegium-eloadassorozata-fokuszban-az-adatvagyon-es-annak-szabalyozasa

Péterfalvi, A. (2014). A Nemzeti Adatvédelmi és Információszabadság Hatóság (NAIH) jelentése Mohács Város Önkormányzatának információszabadsággal kapcsolatos jogsértése tárgyában. [Report of the National Authority for Data Protection and Freedom of Information (Hungarian acronym: NAIH) in the subject of freedom of information related infringement of the Municipality of Mohács Town.] In: *NAIH honlap [NAIH website],* 16 April 2014 https://naih.hu/files/Infoszab-NAIH-2309-11_2013_V_jelentes.pdf

Péterfalvi, A. (2020). A Nemzeti Adatvédelmi és Információszabadság Hatóság (NAIH) NAIH/2020/3433/2 ügyszámú állásfoglalása [Standpoint of the National Authority for Data Protection and Freedom of Information in Case Number NAIH/2020/3433/2], In: *NAIH honlap [NAIH website],* April 2020 https://www.naih.hu/files/infoszab_allasfoglalas_NAIH-2020-3433-2.pdf

Péterfalvi, A., Sziklay, J. (2017). Gazdálkodás az adatvagyonnal [Management of data asset.] In: Bábosik, M. (edit.): *Vezetés a közjó szolgálatában – Közpénzügyi gazdálkodás és menedzsment. [Management in the service of common good - Public funds management]* State Audit Office of Hungary – Typotex Kiadó [Typotex Publisher], Budapest, pages 263–279

Schopp, A. (2020). Állami adatvagyon: csiszolatlan gyémánt. [State data asset: uncut diamond.] In: *ITbusiness.hu honlap [ITbusiness.hu website],* 9 April 2020, https://itbusiness.hu/technology/aktualis_lapszam/center/allami-adatvagyon-csiszolatlan-gyemant

Székely, I. (2015). Közadatok és nyilvános adatbázisok: a hozzáférés kérdései. [Public data and public databases: issues of access.] *Educatio,* Year 24 Edition 3, pages 40-50, https://folyoiratok.oh.gov.hu/educatio/kozadatok-es-nyilvanos-adatbazisok-a-hozzaferes-kerdesei

ÁSZ-jelentés (2017). [SAO report (2017).] Az Állami Számvevőszék „Az adatvédelem ellenőrzése – az adatvédelem hazai keretrendszerének és egyes kiemelt adatnyilvántartások ellenőrzése nemzetközi együttműködés keretében 2017." című 17061. azonosító számú jelentése, 2017. március 14. [Report of the State Audit Office of Hungary titled 'Audit of data protection - audit of the national framework system of data protection and some special data records 2017' of ID No. 17061, 14 March 2017.], https://www.asz.hu/storage/files/files/jelentes/2017/17061.pdf?ctid=1125

ÁSZ-jelentés (2020). [SAO report (2020).] Az Állami Számvevőszék „Az adatvédelem ellenőrzése – az adatvédelem hazai keretrendszerének és egyes kiemelt adatnyilvántartások ellenőrzése nemzetközi együttműködés keretében 2020." című 20077. azonosító számú jelentése, 2017. március 14. [Report of the State Audit Office of Hungary titled 'Audit of data protection - audit of the national framework system of data protection and some special data records 2020' of ID No. 17061, 22 May 2020.], https://www.asz.hu/storage/files/files/jelentes/2020/20077.pdf?download=true

Danish AI Strategy (2019). National Strategy for Artificial Intelligence, In: https://en.digst.dk/policy-and-strategy/denmark-s-national-strategy-for-artificial-intelligence, March 2019, page 33

DJP (2017). A Digitális Jólét Program 2.0. [Digital Welfare Program 2.0.], In: *DJP honlap*

[DJP website], July 2017, page 15, https://digitalisjoletprogram.hu/files/57/1c/571c60381c274901733f8a2fc8a1cca5.pdf

EBH (2013). P16. számú ítélet [Judgement No. EBH2013 P16]: Kúria Pfv. IV. 20.137/2013. számú ítélete, 16/2013. számú polgári elvi határozata [Judgement No. Pfv. IV.137/2013, civil decision in principle No. 16/2013 of the Curia of Hungary], https://kuria-birosag.hu/hu/elvhat/162013-szamu-polgari-elvi-hatarozat

EU Nyíltadat-stratégia (2011). [EU Open data strategy (2011).] Communication No. COM 2011/882 of the European Commission - Open data - an engine for innovation, growth and transparent governance, In: *Eur-lex honlap [Eur-lex website],* 12 December 2011, page 2, https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF

EU Data Strategy (2017). Communication No. 2017/9 from the European Commission to the European Parliament, the Council, the European Economic and Social Committee – Building a European data economy, In: EC Europa.eu honlap [EC Europa.eu website], 10 January 2017, pages 2-6, https://ec.europa.eu/transparency/regdoc/rep/1/2017/EN/COM-2017-9-F1-EN-MAIN-PART-1.PDF

EU Data Strategy (2020). The European Data Strategy - Shaping Europe's digital future *EC Europa.hu honlap [EC Europa.hu website],* 19. February 2020 https://ec.europa.eu/commission/presscorner/detail/en/fs_20_283

Fundamental Law of Hungary (Fundamental Law)

Act C of 2000 on Accounting (Accounting Act)

Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information (PSI directive)

Act V of 2006 on Public Company Information, Company Registration and Winding-Up Proceeding (Company Act)

Act CXII of 2011 on the Right of Informational Self-Determination and on Freedom of Information (Information Act)

Act LXIII of 2012 on the Re-Use of Public Sector Information

Act V of 2013 on the Civil Code (Hungarian acronym: Ptk.)

Act L of 2013 on Electronic Information Security of State and Government Bodies (Information Security Act)

Act CL of 2016 on the Code of General Administrative Procedure (Hungarian acronym: Ákr.)

AI Recommendation (2020). Sajtóközlemény – Az EP a mesterséges intelligencia fejlesztésének etikai és jogi vetületeiről fogadott el ajánlást. [Press release - EU approved a recommendation on the ethical and legal challenges posed by the development of artificial intelligence] In: *Európai Parlament honlap [European Parliament website]*, 21 October 2020. https://www.europarl.europa.eu/news/hu/press-room/20200925IPR87932/making-artificial-intelligence-ethical-safe-and-innovative

AI White Book (2020). European Commission COM/2020/65, White Paper On Artificial Intelligence - A European approach to excellence and trust, In *Op.Europa.eu honlap [OpEuropa.hu website],* 19 February 2020, page 2, https://op.europa.eu/hu/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1/language-en

AI Coordinated Plan (2018). European Commission COM/2018/795, Coordinated

Plan on Artificial Intelligence, In *EC Europa.eu honlap [EC Europa.eu website],* 7 February 2018, page 4, https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence

AI Strategy (2018). Communication from the Commission COM(2018) 237 - Artificial Intelligence for Europe, In: *Eur-lex honlap [Eur-lex website],* 26 June 2018, page 2, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN

AI Strategy (2020). Magyarország Mesterséges Intelligencia Stratégiája 2020–2030 [Hungary's Artificial Intelligence Strategy (2020-2030)], In: *DJP honlap [DJP website]*, May 2020, https://digitalisjoletprogram.hu/files/6f/3b/6f3b96c7604fd36e436a96a3a01e0b05.pdf

German AI Strategy (2018). Artificial Intelligence Strategy, In: https://knowledge4policy.ec.europa.eu/publication/germany-artificial-intelligence-strategy_en, November 2018, page 32