European Review of Digital Administration & Law - Erdal
2021, Volume 2, Issue 2, pp. 109-120
ISSN 2724-5969 - ISBN 979-12-5994-752-9 - DOI 9791259947529 11

*Blockchain and Big Data in the Public Sector: Some Considerations About G.D.P.R. Compliance**

Giovanni Gallone
(Magistrate at T.A.R. Lecce, PhD L.U.M. University)

ABSTRACT: The application of Blockchain technology to the public sector is inevitably intertwined, also in the perspective of the automation of administrative decisions, with the theme of the management and exploitation of big data, offering enormous potential but also posing unprecedented problems of compatibility with the European and national legal framework on data protection.

1. *Blockchain technology in the modern data society*

The first studies on the application of blockchain to the public sector[1] have focused on the most obvious potential of this technology, emphasising the impact it is bound to have on organisation and administrative action.

The unmodifiability of the operations included in the blockchain, the *ab origine* sharing of data between network participants and the passage from the traditional centralised trust system to forms of distributed trust (with the consequent change in the role of the Public Administration) represent, however, only the most immediate effects of distributed computer registers.

The real potential of such a disruptive technology with its wide range of applications can, however, be fully appreciated if it is seen in the overall and varied panorama of the fourth industrial revolution[2].

From this perspective, blockchain technology is certainly a candidate for a leading role in the modern data society[3].

If, in fact, big data is the new fuel of the digital economy[4], the most delicate challenge facing not only economic operators but also public decision-makers[5] is that of making such an enormous amount of data ever more easily exploitable. Data cleansing is the main step towards a "green" transition[6] in the digital sector as well, capable of remedying, in some way, the distortions of "*surveillance capitalism*"[7]. Facilitating the clear and transparent reading of mega-data is, in fact, the first remedy to oppose an obscure management (and concentrated in the hands of a few private companies) of such an important and delicate resource[8].

---

* Article submitted to double-blind peer review.

[1] M. Atzori, *Blockchain technology and decentralised governance: is the state still necessary?*, in *Journal of governance and regulation,* vol. 6, n. 1, 2017, 45; G. Gallone, *Blockchain, procedimenti amministrativi e prevenzione della corruzione*, in *Il Diritto dell'Economia*, 3, 2019 196; M. Allena, *Blockchain technology for environmental compliance: towards a 'choral' approach*, in *Enviromental Law Review*, vol. 50, n. 4, 2020; M. Macchia, *Blockchain and public administration*, in *Federalismi*, n. 2, 2021. On the topic of the relationship between administrative contractual automation and blockchain technology, see also G. Gallone, *Public Administration and the Challenge of Contractual Automation. Notes on Smart Contracts*, in *European Review of Digital Administration and Law*, vol. 1, issue 1-2, 2020, 179.

[2] K. Schwab, *The Fourth Industrial Revolution*, World Economic Forum, Geneva, 2015.

[3] Among the first to reflect on the topic E. Karafiloski, A. Mishev, *Blockchain Solutions for Big Data Challenges. A literature review*, in IEEE EUROCON 2017 - 17th International Conference on Smart Technologies, 2017, 763-768; K. Rabah, *Convergence of AI, IoT, Big-Data and Blockchain: a review*, in *The Lake Institute journal,* vol. 1, n. 1, 2018, 1.

[4] "Data is the new oil. It's valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc. to create a valuable entity that drives profitable activity; so, must data be broken down, analysed for it to have value" according to the famous statement of the British mathematician Clive Humby.

[5] M. Maciejewski, *To do more, better, faster and more cheaply: using big data in public administration*, in *International Review of Administrative Sciences*, 2017, 120; F. Costantino, *Lampi. New frontiers of administrative decisions between open and big data,* in *Diritto Amministrativo*, 2017, 799.

[6] M. Atzori, *Blockchain technology and decentralised governance: is the state still necessary?*, 51.

[7] In the framework described by S. Zuboff, *The Age of Surveillance Capitalism. The fight for a human future and the new frontier of power*, New York, NY, Public Affairs, 2019.

[8] Data holding, like all other forms of capital coding (in the cross-cutting reading offered by K. Pistor, *The Code of Capital: How the Law Creates Wealth and Inequality*, Princeton, NJ, Princeton University Press, 2019), is a factor in wealth polarisation (J. Stiglitz, *The Price of In-*

However, the most interesting aspect is probably that linked to the virtuous combination that can arise between distributed registers, clear reading of big data and public decision-making automation. The still largely unexplored processing capacities of artificial intelligence[9], especially in the form of machine learning, are inevitably enhanced by the possibility of having a large quantity of information which, by means of the blockchain, is traceable and easier to consult. This opens up new frontiers of complexity and accuracy for public decision-makers in weighing up and assessing public interests, with obvious repercussions on the effectiveness of administrative action[10].

2. *The potential of distributed ledgers in the management and exploitation of big data*

The development of information technology has been accompanied by a progressive and unstoppable increase in the amount of data produced by users. According to the most reliable estimations[11], 59 zettabytes of data were created, collected, stored, copied and processed in 2020 alone. This amount is even more impressive and gives an idea of the exponential growth of this volume, if we consider that between 1984 and 2016 (i.e. over a total of 32 years) less than 5 zettabytes of data passed through the Internet[12].

This growth trend is, likely, also driven by the digital transition, destined to be confirmed in the coming years. The increase in the amount of data produced and potentially available is also driving the construction of ever larger data sets[13].

This requires, especially in the future, the abandonment of the traditional model of the information silo[14] in favor of more innovative and effective forms of data management, requirements that are only partly met by cloud computing[15].

In fact, the data silos model, which is still widely dominant at least in public administrations, with its flat structure due to the progressive storage of data in layers, poses problems not only of efficiency (in terms of speed in retrieving, accessing and processing data) but also of overall consistency, since when there are two or more silos for data that are wholly or partly identical, their content could be different, misaligned and, as such, a potential source of entropy and errors. Cloud computing, while reducing these drawbacks, remains deficient in terms of integrity and security[16].

equality. How Today's Divided Society Endangers Our Future, New York, WW Norton & Co, 2013). The darker implications of the phenomenon had already been predicted by V. Mayer-Schonberger and K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think*, Boston, Houghton Mifflin Harcourt, 2013; C. O'Neill, *Weapons of Math destruction*, London, Penguin Books Ltd, 2017. The "dark side of transparency" and the unknowns that lie behind the dismantling of any diaphragm with respect to the knowability of data have been investigated, also on the IT level, in a broad spectrum between "miti ed ombre" by A.G. Orofino, *La trasparenza oltre la crisi. Accesso, informatizzazione e controllo civico*, Bari, Cacucci, 2020, 285.

[9] D. Marongiu, *L'intelligenza artificiale "istituzionale": limiti (attuali) e potenzialità*, in *European Review of Digital Administration and Law*, vol. 1, issue 1-2, 2020; J. Valero Torrijos, *The Legal Guarantees of Artificial Intelligence in Administrative Activity: Reflections and Contributions from the Viewpoint of Spanish Administrative Law and Good Administration Requirements,* in *European Review of Digital Administration and Law*, vol. 1, issue 1-2, 2020, 55, C. Fraenkel-Haeberle, *Fully Digitalized Administrative Procedures in the German Legal System*, in *European Review of Digital Administration and Law*, vol. 1, issue 1-2, 2020, 105.

[10] This is the frontier represented by data *analysis automation* and *augmented analytics*, meaning the ability to use technologies such as artificial intelligence, *machine learning* and natural language to explore data and analyse how content is developed, consumed and shared. One example is the possibility of processing the huge amount of data managed through an administrative blockchain in the field of health (number of accesses, number of patients, type of pathologies, etc.) through artificial intelligence tools to establish the most efficient location of a hospital on the territory.

[11] These figures are taken from the International Data Corporation's (IDC) Global Dataphere published on 8 May 2020.

[12] Cisco Visual Networking Index, 2018.

[13] A data set is defined as a set of data structured in relational form that can, as such, be analysed. The tendency to build larger and larger data sets is explained by the almost banal observation that enlarging the database subject to contextual analysis enables the extraction of additional information compared to that which could be obtained by analysing small series with the same total amount of data.

[14] The term data silo refers to the storage of fixed data that remain under the control of a single manager and are isolated from the rest of the system.

[15] The cloud computing paradigm envisages Data as a Service (*DaaS*), i.e. making data available to the user via the web as if they were physically present on the local disk.

[16] By its very nature, *cloud computing* remains particularly exposed to *data breaches*, since data are usually stored by third parties. The decision to use the cloud also raises a number of important points for reflection on the subject of digital sovereignty and cybersecurity, as highlighted, with regard to the Italian situation and the launch of the hub of the European project Gaia X by L. Bolognini and E. Pelino, *Le tentazioni del cloud eu-*

It is evident that these solutions are destined to prove inadequate, especially if one bears in mind that, by now, the qualitative characteristics of the big data are no longer enclosed in the formula of the "3 V" (Volume, Velocity and Variety)[17] but are expressed, more fully, also by the two further attributes of "Veracity" and "Value". More specifically, the profiles of the quality of the data (in terms of reliability of the data content) and the capacity of this to be transformed into added value, through its processing, are now of primary importance.

This is the transition from big data to smart data and blockchain technology can play a crucial role in it[18].

The possibility of relying on a distributed and unmodifiable register brings a series of undoubted technical advantages in the management of large quantities of complex data. First of all, the application of blockchain technology to data storage makes it possible to significantly increase the overall level of security, almost completely eliminating the risk of tampering with the information stored (so-called data integrity). In addition, its distributed structure makes it possible to reduce the frequency and impact of malfunctioning episodes, thereby facilitating any recovery operations[19].

From another point of view, the use of blockchain technology makes it possible to implement data quality. Very often, data acquired from outside and entered into the database, coming from different sources, are unstructured and often present different formats. Inserting them into the distributed ledger allows them to be structured and standardized for the first time. This conversion into a more structured format inevitably affects the capacity of the data to generate value (the Value mentioned above) as it facilitates and accelerates data mining operations, also in the perspective of using the information in a predictive way through artificial intelligence tools[20].

In addition, the functioning of the blockchain, which is based on the sequential recording of the operations carried out by affixing a time stamp to them, opens up the possibility of real-time data analysis, through the continuous monitoring of their flow (Real Time Data Analytics). This is accompanied, with obvious organisational consequences, by a simplification of the data life cycle, which becomes more linear both in terms of its processing (fully traced) and in terms of access to it (being shared *ab origine* by all network participants)[21]. This simplification also brings obvious advantages not only in terms of efficiency, but also in terms of costs and resources needed to manage big data.

3. *The implications of applying blockchain technology to the public sector and interferences with data protection*

The huge benefits of using blockchain technology in big data management are even more evident when looking at the administrative sector.

The latter, in fact, by its nature, is predisposed to the management of a large and complex amount of data.

Moreover, public decision-makers still perceive this as a responsibility and a burden rather than a strategic resource in the development of their policies. In this attitude of mistrust, in addition to the mentality of the apparatus, the evident interference between the management (and exploitation) of big data held by the Administration and the tasks and duties of data protection play a fundamental role[22].

ropeo e nazionale: tra semplificazione politica e critica giuridica, in *Istituto italiano per la privacy e la valorizzazione dei dati*, 25 September 2020, and M. Rhao, *Gaia X: un'analisi giuspubblicistica dei cloud europei*, in *Gruppo di Pisa*, vol. 3, 2021, 283.

[17] This is the well-known model developed by D. Laney, *3D Data Management: Controlling Data Volume, Velocity and Variety*, *Gartner*, 2001.

[18] For an updated overview, see, among others, N. Deepa, Q.-V. Pham, D.C. Nguyen, S. Bhattacharya, B. Prabadevi, T.R. Gadekallu, P.K.R. Maddikunta, Fang Fang and P.N. Pathirana, *Survey on Blockchain for Big Data: Approaches, Opportunities, and Future Directions*, in *ArXiv*, 2020, 1.

[19] In this sense, the sharing of the register by the operators makes back-up operations practically superfluous, since data can be retrieved simply by drawing on what is held by each node (which holds its own copy, identical to the others, of the register).

[20] N. Deepa, Q.-V. Pham, D.C. Nguyen, S. Bhattacharya, B. Prabadevi, T.R. Gadekallu, P.K.R. Maddikunta, Fang Fang and P.N. Pathirana, *Survey on Blockchain for Big Data: Approaches, Opportunities, and Future Directions*, 4.

[21] The doctrine speaks about "Streamlining the Data Access", in N. Deepa, Q.-V. Pham, D.C. Nguyen, S. Bhattacharya, B. Prabadevi, T.R. Gadekallu, P.K.R. Maddikunta, Fang Fang and P.N. Pathirana, *Survey on Blockchain for Big Data: Approaches, Opportunities, and Future Directions*, 5.

[22] The entire framework of the European legislation on *data protection* (Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the pro-

In fact, it cannot be forgotten that the data managed by the Public Administration are normally of a personal nature pursuant to Art. 4 no. 1) of the G.D.P.R.[23].

The application of blockchain technology to the management of data held by the administration opens up a completely new scenario[24].

It is, in fact, now widely believed that blockchain technology is destined to have a significant impact on the science of administration, challenging traditional models and categories[25]. Insofar as it is of interest here, its application leads to the conclusion that the very idea of transparency is partially outdated. With the disappearance of a central authority holding the documents and information (which are instead shared by all the participants in the distributed database), accessibility and visibility from the outside is replaced, in terms of active involvement, by the sharing of the data from the source. They are born shared and from their sharing they derive, moreover, the attribute of non modifiability.

The total and immediate accessibility to the data guaranteed to participants by the distributed register seems, apparently, to be in frontal contrast with the protection of the confidentiality of the owners of the data. The possible tensions between blockchain and the G.D.P.R. have, moreover, been immediately recognised by both doctrine[26] and the European institutions[27].

The data recorded in the blockchain register, in addition to being physiologically public and potentially knowable by anyone, are unmodifiable and this may clash with the implementation of the principle of data minimisation pursuant to art. 5 G.D.P.R. and, in particular, with the definition of limited storage times of the same[28].

On an organisational level, the co-ownership of data by all participants in the blockchain also requires a rethink of the more established solutions for identifying data controllers and data processors.

The problem of G.D.P.R. compliance, which is a key aspect of blockchain

cessing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation - G.D.R.) is based on the concept of accountability of the data controller (see recitals 13, 74 and 79 and Art. 24 of the G.D.P.R. available at www.eur lex.europa.eu). For an overview of the new discipline M.G. Stanzione, *Il regolamento europeo sulla privacy: origini e ambito di applicazione*, in *Europa e Diritto Privato*, 2016, 4, 1249; L. Bolognini, E. Pelino, C. Bistolfi, *Il regolamento privacy europeo. Commentario alla nuova disciplina sulla protezione dei dati personali,* Milano, Giuffrè, 2016. The implementation and transposition of the G.D.P.R. by Public Administrations implies also particularly heavy organisational burdens, including the identification of *a Data Protection Officer* (Article 37 of the G.D.P.R.) and the preparation of a register of processing operations (Article 30 of the G.D.P.R.).

[23] And among these also "genetic", "biometric" and "concerning health" data ex art. 4 nr. 13), 14) and 15) of the G.D.P.R.. The issue of the processing of sensitive data by the administration through IT tools has been the subject of interesting reflections also before the advent of the G.D.P.R. outside the Italian borders (J. Valero Torrijos, *Las transformaciones de la innovación tecnológica en la administración pública y su proyección sobre la protección de los datos de caractér personal*, in *La revista de la Agencia de Protección de Datos de la Comunidad de Madrid,* 2012, 56; C. Sarmento e Castro, *Direito da informática, privacidade e dados pessoais*, Coimbra, Almedina, 2003, 170).

[24] These are considerations already partly made in G. Gallone, *Blockchain, procedimenti amministrativi e prevenzione della corruzione*, 196.

[25] In this regard, see the reflections of M.A. Bernal Blay, *Blockchain, Administración y contratación pública,* in *Observatorio de Contratación Pública*, 12 July 2018; B. Barraud, *Les blockchains et le droit*, in *Reveu Lamy droit de l'immatériel*, 147, 2018, 48; M. Macchia, *Blockchain and public administration.*

[26] On the topic of the relationship between this discipline and blockchain technology see M. Berberich and M. Steiner, *Practitioner's Corner. Blockchain Technology and the GDPR - How to Reconcile Privacy and Distributed Ledgers?*, in *European Data Protection Law Review*, 3, 2016; M. Finck, *Blockchains and Data Protection in the European Union*, in *Max Planck Institute for innovation & Competition Research paper*, 18-01, 30 November 2017; L. Ibanez, K. O'Hara, E. Semperl, *On blockchains and the general data protection regulation*, *EU Blockchain Forum and observatory*, 7,2018; A. M. Gambino and C. Bompressi, *Blockchain and personal data protection*, in *Information and Computer Law*, 3, 2019; A. Palladino, *The lost balance of blockchain between platform revolution and GDPR compliance*, in *Rivista di diritto dei media*, n. 2, 2019.

[27] See European Parliament Resolution of 3 October 2018 on distributed ledger and blockchain technologies: building trust through disintermediation (2017/2772(RSP)) which moves from the idea that distributed ledger technology provides for "the pseudonymisation of users, but not their anonymisation" (Recital D) to conclude in the sense of its full subjection to the G.D.P.R.. On this point, see also the *Thematic report Blockchain and the GDPR,* by the European Union Blockchain Observatory, 16 October 2018, published at www.eublockchainforum.eu.

[28] A. Palladino, L'*equilibrio perduto della blockchain tra platform revolution e GDPR compliance*, 153. This is what is set out in Article 5(1)(c) and (e) of the GDPR, according to which data must be "adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati and devono essere conservati in modo da consentire l'identificazione degli interessati per un tempo non superiore a quello necessario al conseguimento della finalità del trattamento".

governance, arises in all its complexity[29]. In relation to this also emerges the ongoing indispensability of a public actor that provides for a regulation of the network that is technically and legally compliant with domestic and European law[30].

4. *Administrative blockchain and G.D.P.R.*

The basic difficulty in reconciling the characteristics of the blockchain technology with the European discipline is represented by the circumstance that the G.D.P.R., although expressly referring also to the processing of personal data by "public authorities"[31] is modelled on the traditional forms of centralized processing through data silos.

In the absence of an ad hoc regulatory framework, in order to ease the possible tensions between administrative blockchain and G.D.P.R., it is therefore necessary to start from the structure that the distributed register can concretely assume.

The debate on the choice between the *permissonless* (or public) and the *permissioned* (or private)[32] blockchain model often takes on an almost ideological hue, so much so that purists claim that the *permissionless* blockchain is even outside the technology paradigm, as if it were a betrayal of it.

Yet, it seems that, at least in the near future, the only really viable alternative for the exploitation of blockchain technology in the public sector is the *permissioned* model[33].

The latter, in addition to having a better technical performance and not suffering from the volatility of open networks, makes it possible to better address (and overcome) the critical issues described above in terms of data protection.

In fact, only the option for the *permissioned* model makes it possible to meet the fundamental canons of privacy by design and privacy by default enshrined in Article 25 of the G.D.P.R. According to them, the architecture of the distributed registry must allow, from the moment of its design and by default, the protection of personal data in accordance with the standards of the Regulation. It follows that the problem of G.D.P.R. compliance must be addressed and resolved ex ante through the adoption of measures of a technical and organisational nature. Furthermore, the principles in question are not absolute and rigid, but, by express provision of the G.D.P.R., flexible, so that the judgement of adequacy and of exigibility of the aforesaid protection measures must be conducted in concrete terms also on the basis of a cost-benefit analysis[34]. It is clear from this that the public decision-maker, in the governance of blockchain, retains a wide

[29] In the blockchain ecosystem, the role of the Administration is that of the guarantor of the rules of the system and is therefore not just a matter of "governance by blockchain", i.e. governing by blockchain, but it is mainly about "blockchain governance", i.e. the governance of the tool itself (S. Ølnes, J. Ubacht, M. Janssen, *Blockchain in government: benefits and implications of distributed ledger technology for information sharing*, in *Government information quarterly*, 2017).

[30] Moreover, as far as Italy is concerned, it cannot be overlooked that Article 50 (1) of Legislative Decree no. 82 of 7 March 2005 (the Digital Administration Code) establishes that "I dati delle pubbliche amministrazioni sono formati, raccolti, conservati, resi disponibili e accessibili con l'uso delle tecnologie dell'informazione e della comunicazione che ne consentano la fruizione e riutilizzazione, alle condizioni fissate dall'ordinamento, da parte delle altre pubbliche amministrazioni e dai privati; restano salvi i limiti alla conoscibilità dei dati previsti dalle leggi e dai regolamenti, le norme in materia di protezione dei dati personali ed il rispetto della normativa comunitaria in materia di riutilizzo delle informazioni del settore pubblico".

[31] Numbers 7), 8), 9) and 10) of Article 3, containing the definition of "controller", of "processor", of "recipient" and of "third party" also include "the Public Authority". Furthermore, Art. 6 par. 1 letter e) contemplates as a condition of lawfulness of the processing of personal data, besides the consent of the data subject, also "the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller".

[32] In the *permissionless blockchain,* which is the basis of the Bitcoin experience, access to the network is totally free and control is distributed and entrusted to all participants. In the *permissioned blockchain* model, on the other hand, both access and validation are reserved to a restricted group of participants.

[33] G. Gallone, *Blockchain, administrative procedures and corruption prevention*, 197.

[34] In this regard, Article 25 of the G.D.P.R. establishes that the measures of protection must be defined "Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing". This is a specific declination of the theme of the management of the risks (in this case connected to the treatment of personal data) inherent in the carrying out of the administrative activity (A. Barone, *Il diritto del rischio,* Milano, Giuffrè, 2006; Id., *Amministrazione del rischio e intelligenza artificiale*, in *European Review of Digital Administration and Law,* vol. 1, issue 1-2, 2020, 63). In line with the dictates of risk management, the G.D.P.R. incorporates, moreover, a notion of risk not limited to the phase of the violation, but also to the previous one in which the violation has not yet taken place (A.M. Gambino and C. Bompressi, *Blockchain e protezione dei dati personali,* 621).

margin of appreciation, which allows him to consider different solutions capable of reconciling the protection of personal data and the potential of distributed registries, respect for the G.D.P.R. and the efficiency of administrative action.

Several technical and organisational measures have been suggested in doctrine to avoid what has been called the paradox of a *permissioned* blockchain exempt from compliance with the G.D.P.R.[35].

From an organisational point of view, if in the *permissionless* blockchain there is a clear lack of accountability in the handling of data, in the *permissioned* blockchain it is, on the contrary, possible to identify in advance, at the time of the establishment of the operating rules of the network, the node (or nodes) that are the owners or managers of the processing.

It has been, in this regard, unanimously observed[36] how it is the same Art. 26 of the G.D.P.R. to admit forms of co-ownership of the treatment. This would seem to allow, by setting the provision in question in the proper context of the blockchain (characterized by the sharing *ab origine* of the data by all the operators), the recognition of the quality of owner of the treatment to all the participants in the network. However, this would result in an unjustified multiplication of administrative requirements relating to the processing of personal data (which would fall largely on private parties), with a consequent overall burden on administrative action. Moreover, the identification of a potentially large number of co-processors risks generating dangerous interference in the performance of the functions linked to this quality.

From a purely legal point of view, it should be noted that there is co-ownership of the processing only where "two or more controllers jointly determine the purposes and means of processing"[37]. In this sense, it seems that the mere participation to the network and the recognition of the power to validate the operations carried out is not sufficient to attribute the co-ownership of the treatment, requiring a quid pluris consisting in the power to establish the purpose and method of the treatment itself.

In an administrative blockchain, this role of defining the "purposes"' and "means" of processing will reasonably fall to the "public" or "administrative" node (*id est* the public administration that has set up the network and oversees its governance). This "administrative" node will thus play a different role from the other "non-administrative" nodes which, on the other hand, although they will share all the data entered in the register from the outset, will not have the status of data controller[38].

Such a solution has the undoubted merit of reducing the subjective complexity of the personal data protection system and of ensuring its greater efficiency. On the other hand, a differentiation, from a legal point of view, of the position of the "administrative" node compared to that of the other "non-administrative" nodes, seems in any case imposed by the need to ensure the compliance of the instrument with the internal and European Union regulations on electronic identification referred to in Regulation (EU) no. 910/2014 of the European Parliament and of the Council of 23 July 2014 (the so-called "Regulation"), through the provision in its architecture of a master node that operates as a provider of the electronic signature solution, certifying the identity of the holders of the cryptographic keys admitted to the network[39].

Furthermore, it must be borne in mind that the ownership of the processing is a concept quite different from the responsibility for the same. In fact, Art. 4 no. 7) and 8) of the G.D.P.R. distinguish the already examined notion of "controller" from that of "processor", which is identified with "natural or legal person, public authority, agency or other body which processes personal data on

[35] A. Mirchandani, *The GDPR-Blockchain Paradox: Exempting Permissioned Blockchains from the GDPR*, *in Fordham Intellectual Property, Media and Entertainment Law Journal*, vol. 29, n. 4, 2019.

[36] A. Palladino, L'*equilibrio perduto della blockchain tra platform revolution e GDPR compliance*, 154.

[37] Thus, Article 26 par. 1 G.D.P.R. This notion is, after all, a reflection of the more general notion of "controller" set out in Article 4 par. 1 no. 7 of the G.D.P.R., according to which "controller [...] means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".

[38] M. Finck, *Blockchains and Data Protection in the European Union*, 26, who, however, observes, with reference to the text of art. 4 nr. 7) of the GDPR that "The use of the singular indicates that in centralised data silos there is often only one entity that qualifies as a data controller. It is to them that the GDPR is addressed".

[39] As already highlighted in G. Gallone, *Blockchain, procedimenti amministrativi e prevenzione della corruzione*, 197; F. Sarzana di S. Ippolito and M. Nicotra, *Diritto della blockchain, intelligenza artificiale e IoT*, Milano, Ipsoa Wolters Kluwer, 2018, 66.

behalf of the controller"[40].

Therefore, it cannot be excluded that in a more complex administrative blockchain, a plurality of nodes with differentiated roles could be envisaged, contemplating, alongside an "administrative" master node that would combine the role of provider of the electronic signature solution and holder of the processing of personal data, one or more further "administrative" nodes that could process the data as data controllers.

Moreover, in a scenario that sees the construction of networks of small size, with a limited number of nodes, organized in a permissioned form and subject to public regulation[41], while waiting for the creation of a single public infrastructure based on distributed registers, the question of G.D.P.R. compliance arises, always on the organizational level, with respect to the flow of data that normally takes place between public administrations[42] and, in particular, between administrative blockchains in which the role of owner of the processing of personal data belongs to different administrative subjects. This is linked to the technical issue of interoperability between the different administrative blockchain circuits and depends, essentially, on the ability of the technology to ensure that the data migrated from one distributed registry to another retain the attributes of non modifiability and temporal certainty. The reciprocal reliability of the blockchain platforms, allowing the full traceability of the life of the data, would, in fact, greatly reduce the problem of establishing the moment in which this passes from the legal availability (and responsibility) of one owner to that of the other and if and, above all, at what juncture of the treatment violations of the law have occurred.

In terms of the way in which personal data are processed, even about the administrative blockchain, there is the problem of protecting the confidentiality of those concerned. The personal data, in fact, as soon as it has been entered into the distributed register, becomes knowable by all participants in the network. This does not, however, prevent the identification of certain categories of data that are destined to remain (temporarily or perpetually) confidential[43], when establishing the rules for the functioning of the *permissioned* blockchain.

Moreover, the G.D.P.R. itself provides for the possibility of "pseudonymisation"[44] of the data. From a technical point of view, this can be achieved through various solutions such as, for example, the use of so-called "ring" signatures[45], multiple signatures affixed by

[40] What is important is that, when defining the rules for the functioning of the administrative blockchain, it is ensured, as Recital 79 to the G.D.P.R. requires, "a clear allocation of responsibilities [...] including where a controller determines the purposes and means of the processing jointly with other controllers or where a processing operation is carried out on behalf of a controller".

[41] M. Atzori, *Blockchain governance and the role of trust service providers: the trusted chain network*, in *The Journal of British Blockchain Association*, vol. 1, n. 1, 1-17, who imagines blockchain technology as embedded in a "multi-stakeholder framework" that ensures an "appropriate mix of centralization and decentralization" tailored to the concrete needs of individual processes.

[42] This phenomenon is expressly admitted, as regards the Italian legal system, by the new paragraph 2 of Article 50 of Legislative Decree n. 82 of 7 March 2005 (Digital Administration Code), as last amended by Article 264, paragraph 2, letter b), Law no. 77 of 2020, according to which "Qualunque dato trattato da una pubblica amministrazione, con le esclusioni di cui all'articolo 2, comma 6, salvi i casi previsti dall'articolo 24 della legge 7 agosto 1990, n. 241, e nel rispetto della normativa in materia di protezione dei dati personali, è reso accessibile e fruibile alle altre amministrazioni quando l'utilizzazione del dato sia necessaria per lo svolgimento dei compiti istituzionali dell'amministrazione richiedente, senza oneri a carico di quest'ultima, salvo per la prestazione di elaborazioni aggiuntive". The legal framework is completed by the subsequent par. 2 bis, which establishes that "Le pubbliche amministrazioni, nell'ambito delle proprie funzioni istituzionali, procedono all'analisi dei propri dati anche in combinazione con quelli detenuti da altri soggetti di cui all'articolo 2, comma 2, fermi restando i limiti di cui al comma 1. La predetta attività si svolge secondo le modalità individuate dall'AgID con le Linee guida".

[43] Through the affixing of a "restriction on processing", defined in Article 4 par. 1 no. 3 of the G.D.R. as a "marking of stored personal data with the aim of limiting their processing in the future". The need to maintain confidentiality as regards the content of certain blocks in the chain is particularly felt in the field of public contract award procedures with regard to the bids submitted by participants (which must remain confidential until, at least, their evaluation by the Commission).

[44] Art. 4 par. 1 no. 5 of the GDPR which defines "pseudonymization" as "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".

[45] It is a type of digital signature that constitutes the evolution of group digital signatures (category first elaborated by D. Chaum and E. van Heyst, *Group signatures*, in D.W. Davies (ed.), *Advances in Cryptology - Eurocrypt' 91*, Berlin, Springer, 1991, 257-265). It allows any member of a predefined group of users to sign mes-

more than one operator, so as not to render the data univocally traceable to one of them, or the aggregation of large quantities of data coming from different subjects in a single block with the affixing of a single digital signature[46].

It is obvious that such solutions make it possible to conceal the origin of the data, but not their content per se. However, this, by excluding the certain attributability of the data to a specific data subject in the absence of further specific information[47], seems sufficient, from a legal point of view, in the light of the notion of "pseudonymisation" used in Article 4 no. 5) of the G.D.P.R., to ensure full compliance with the European rules[48].

The same pseudonymisation techniques can also provide a solution to the different, albeit related, problems of "minimisation", "limitation of data retention" and, consequently, protection of the "right to be forgotten" of the persons concerned[49].

According to G.D.P.R. rules, once the purpose for which the processing is intended is fulfilled, the data should be deleted. This is not, however, possible in a distributed register where the operations performed are unchangeable (and cannot, therefore, be expunged from the sequence).

An alternative technically equivalent to erasure may, however, be represented by the subsequent, definitive and irreversible pseudonymisation of the data. This result can be obtained, for example, through the destruction of the cryptographic keys employed during the operations of pseudonymisation[50]. Less satisfactory is, instead, the solution of the off-chain storage of the data to be deleted, which is limited to moving the destruction operation outside the blockchain ecosystem (and, therefore, outside the guarantees of verification that it offers and with an increase in management costs).

Moreover, the issue of the limitation of data retention in the administrative blockchain appears to be considerably re-dimensioned by the provision of the same art. 5 par. 1 letter e) of the G.D.P.R. This last prevision contemplates, in fact, in its second part, a tempering to the general principle, establishing that the "personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject". Archiving in the public interest represents, in fact, a further and peculiar form of lawful processing of personal data different from the basic processing carried out under one of the conditions laid down by art. 6 of the G.D.P.R.[51]. The possibility of archiving permits, first of all, to postpone the deletion of the data with respect to the satisfaction of the purpose for which the treatment is ordered. Furthermore, it allows, pursuant to Art. 17, par. 3 lett. d of the G.D.P.R., even to exclude the right to erasure of the data subject where

sages anonymously on behalf of the group with the possibility that a designated user of the group can later revoke the anonymity and identify the author of a signature.

[46] These are the technical solutions suggested in the *Thematic report Blockchain and the GDPR,* by the European Union Blockchain Observatory, 17.

[47] In fact, by knowing the public key (which acts as an account) of a node, it becomes possible to link the address to a specific identity, thus attributing the authorship of the data.

[48] This view is shared by the Working Party established under Article 29 of Directive 95/46/EC in Opinion 05/2014 - WP 216 on anonymisation techniques adopted on 10 April 2014.

[49] Art. 17 G.D.P.R. also defined as "right to erasure" or "right to be forgotten", which identifies the hypotheses in which "The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay", among which (lett. a) are also those in which the "personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed" and those in which the "personal data have been unlawfully processed". In doctrine on the origin of the right and on its systematic placement also in a comparative perspective M. Mezzanotte, *Il diritto all'oblio*, Napoli, Edizioni Scientifiche Italiane, 2009. About the right to be forgotten in the G.D.P.R. - A. Ricci, *I diritti dell'interessato*, in G. Finocchiaro (ed.), *La protezione dei dati personali in Italia: Regolamento UE n. 2016/679 e d.lgs. 10 agosto 2018, n. 101*, Bologna, Zanichelli, 2019, 392.

[50] This solution requires, however, the identification of the node in charge of the cryptographic key destruction operations. It can, however, coincide, in the *permissioned blockchain*, with the "administrative" master node already designated as provider of the electronic signature solution.

[51] See recital 50. Therefore, a lawful processing in the form of storage of personal data already processed "for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller" pursuant to Article 6 par. 1 lett. e of the G.D.P.R. is certainly conceivable.

the exercise of the same "is likely to render impossible or seriously impair the achievement of the objectives of that processing" of archiving.

The archiving of personal data by means of a register distributed in the public interest is, moreover, a crucial aspect in the construction of a large data set, which takes on the characteristics of big data, from which to draw, also using A.I. tools, for the automated adoption of administrative decisions.

Finally, it remains to be clarified how it is possible to guarantee, also in an administrative blockchain, the right of the interested parties to the rectification or integration ex art. 16 of the G.D.P.R., of the personal data inserted in the distributed register, which by their nature are unchangeable. A part of the doctrine suggests, as a solution to a problem that is eminently technical, the use of *chameleon hashes* that make it possible to keep the sequence hash unaltered while making changes in the individual block[52]. In particular, the *chameleon hash* is nothing more than a cryptographic hash function that provides a backdoor (a method, unknown to operators, to circumvent the cryptographic system) whose key is only available to one of the nodes. It is intuitable that in the *permissioned* blockchain, the backdoor key will preferably be entrusted to the "administrative" master node, already designated, as mentioned, as the provider of the electronic signature solution.

The possibility for the "administrative" node to rectify the data by intervening on the content of the block would also seem to undermine the unchangeable nature of the distributed registry and jeopardise data integrity. However, this concern is unfounded, since every operation, including modification, is tracked and the block cannot be removed from the chain in which it is inserted. In other words, the actions of the node holding the backdoor key can be monitored, according to the logic of transparency that characterizes the blockchain, by the other participants, who will be able to ask the administrative node how it has used the tool (the so-called key management, part of the new role of governance of the infrastructure due to the Public Administration), calling it to justify (and be responsible for) any alteration of the sequence.

5. *Beyond critical issues: an evolving scenario*

If technology offers, therefore, all in all, easy solutions to overcome the critical issues that arise in the construction of an administrative blockchain respecting the European and national rules on the protection of personal data, it should not be overlooked how the technology of distributed registers is even able to raise, in some respects, the minimum standard of protection set by the G.D.P.R.[53].

This is particularly evident in terms of the involvement of the data subject in the data lifecycle. In this sense, the multi-polar structure of the blockchain can allow the single participant, upon admission to the permissioned blockchain, to establish (through an operation that is inserted in the same way as the others inserted in the distributed register) which data to provide and to modulate the consent given to their processing and, above all, to directly monitor the processing itself[54]. This device facilitates the data controller in the census of the will of the interested parties and represents one of the most interesting and innovative forms of co-management of the administrative procedure.

Moreover, it should not be forgotten that the use of blockchain technology in the public sector for the management of big data facilitates and simplifies, in various aspects, the tasks of the data controller. In particular, the transparency of the operations and the co-ownership of the data *ab origine* by all the operators, even with the precautions already seen in the point of protection of privacy, produces an effect of simplification of the procedures of management of the data, different from the treatment *stricto sensu*. The data subject's right of access, in its twofold

[52] K. Ashritha, M. Sindhu, K. V. Lakshmy, *Redactable Blockchain using Enhanced Chameleon Hash Function*, in *5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 2019, 323; M. Finck, *Blockchains and Data Protection in the European Union*, 23; A.M. Gambino and C. Bompressi, *Blockchain and Personal Data Protection,* 642.

[53] As is the aim of the G.D.P.R. itself, which in recital 6 sets out to ensure a "a high level of the protection of personal data".

[54] G. Zyskind, O. Nathan and A. Pentland, *Decentralizing privacy: using blockchain to protect personal data*, in *IEEE Security and privacy workshops*, 2015, 180, according to which through the distributed ledger "users should own and control their data" and "are always aware of the data that is being collected about them and how it is used".

dimension of right to obtain from the data controller confirmation as to whether or not personal data concerning him are being processed and to obtain a copy of such data[55], is satisfied automatically, without the activation of a specific procedure. This simplification eliminates the need for the data subject to make a request and for the data controller to receive it, examine it and decide whether to grant the request by displaying the data or not. And all this clearly translates into a considerable saving of human and economic resources for the Public Administration.

In the face of such obvious potential, the technical and regulatory scenario regarding blockchain administration is, of course, evolving.

Italy has placed itself at the forefront in Europe on a regulatory level by introducing, in paragraph 1 of Article 8b of Decree-Law No. 135 of 2018[56], a legal definition of "distributed register".

This new legislation was essentially descriptive since it was not accompanied by the provision of detailed rules[57]. However, this silence on the part of the legislator should probably be interpreted as an intentional gap, the result of a specific regulatory choice inspired by the principle of technological neutrality[58].

Moreover, if what has been said above has attempted to demonstrate that, even with unchanged legislation, there is the possibility of building an administrative blockchain that complies with the European (and, by implication, domestic) rules on the protection of personal data, it seems that significant margins remain for an intervention by the national legislator to define specific rules.

In particular, Art. 6 par. 3 of the G.D.P.R. provides, for the hypothesis in which the legal basis of the processing of personal data is the performance of a task carried out in the public interest or in connection with the exercise of official authority vested in the data controller, the possibility for the law of the Member State to which the data controller is subject (as well as for Union law) to introduce, in compliance with the principle of proportionality, "specific provisions to adapt the application of rules" of the Regulation with regard to the "general conditions governing the lawfulness of processing by the controller", to the "types of data which are subject to the processing" to the data subjects and to the "the data subjects concerned" and, above all, to "the purposes for which, the personal data may be disclosed", the "storage periods" and the "processing operations and processing procedures"[59].

[55] Art. 15 G.D.P.R. The co detention *ab origine* of the data can, at the same time, satisfy, through the adoption of technical devices allowing the clear reading of the sequence, also the right to the portability of the data ex art. 20 par. 1 G.D.P.R. according to which "The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format".

[56] The reference is to Decree-Law No. 135 of 14 December 2018, converted with amendments by Law No. 12 of 11 February 2019 of "Conversione in legge del decreto-legge 14 dicembre 2018, n. 135, recante disposizioni urgenti in materia di sostegno e semplificazione per le imprese e per la pubblica amministrazione", published in the Italian Official Gazette No. 36 of 11 February 2019, according to which distributed ledger-based technologies are defined as "protocolli informatici che usano un registro condiviso, distribuito, replicabile, accessibile simultaneamente, architetturalmente decentralizzato su basi crittografiche, tali da consentire la registrazione, la convalida, l'aggiornamento e l'archiviazione di dati sia in chiaro che ulteriormente protetti da crittografia verificabili da ciascun partecipante, non alterabili e non modificabili".

[57] Paragraph 4 of art. 8 bis delegates to the Agenzia per l'Italia Digitale the identification of the "standard tecnici che le tecnologie basate su registri distribuiti debbono possedere ai fini della produzione degli effetti di cui al comma 3" and, therefore, only those relating to the "validazione temporale elettronica di cui all'articolo 41 del regolamento (UE) n. 910/2014 del Parlamento europeo e del Consiglio, del 23 luglio 2014" (the so-called *timestamping*) of the electronic transitions, leaving out of this regulatory delegation any other disciplinary aspect of the technology.

[58] On this subject, see A.G. Orofino, *L'esternazione informatica degli atti amministrativi,* in S. Civitarese Matteucci and L. Torchia (eds.), *La tecnificazione*, Firenze, Firenze University Press, 2016, 181. The principle of technological neutrality as a guiding criterion in the approach to blockchain technology is also suggested by the European Parliament Resolution of 3 October 2018 on distributed ledger technologies and blockchain: building trust through disintermediation, recital E.

[59] The content of the provision is clarified by Recital 45 of the G.D.P.R., which states that "Where processing is carried out in accordance with a legal obligation to which the controller is subject or where processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority, the processing should have a basis in Union or Member State law. This Regulation does not require a specific law for each individual processing. A law as a basis for several processing operations based on a legal obligation to which the controller is subject or where processing is necessary for the performance of a task carried out in the public interest or in the exercise of an official authority may be sufficient. It should also be for Union or Member State law to determine the purpose of processing. Furthermore, that law could specify the general conditions of this Regulation governing the law-

Furthermore, it should be recalled that Article 23 of the G.D.P.R. allows the possibility, as a general rule, for the law of the Union or of the Member State to which the controller or processor is subject to provide for legislative measures restricting the scope of the obligations and rights referred to in Articles 12 to 22 and 34, as well as in Article 5, in so far as the provisions contained therein correspond to the rights and obligations set out in Articles 12 to 22, provided that "such a restriction respects the essence of the fundamental rights and freedoms and is a necessary and proportionate measure in a democratic society to safeguard", inter alia, in sub-paragraph e, "other important objectives of general public interest of the Union or of a Member State, in particular an important economic or financial interest of the Union or of a Member State, including monetary, budgetary and taxation matters, public health and social security".

It seems, moreover, that these margins of regulatory intervention have already been partially exploited by the Italian Legislator, going to modify, in the climate of the pandemic crisis, the text of Art. 2 ter of the Legislative Decree No. 196 of 2003 (the so-called "Codice della protezione dei dati personali"), introducing, by means of the Decree Law, 8th October 2021, No. 139[60], a new paragraph 1 bis. This last provision establishes, today, that "il trattamento dei dati personali da parte di un'amministrazione pubblica […] è sempre consentito se necessario per l'adempimento di un compito svolto nel pubblico interesse o per l'esercizio di pubblici poteri a essa attribuiti" con la precisazione che la "finalità del trattamento, se non espressamente prevista da una norma di legge o, nei casi previsti dalla legge, di regolamento è indicata dall'amministrazione […] in coerenza al compito svolto o al potere esercitato".

It clearly shows the intention of the legislator to overcome a logic of rigid *ex ante* typification of the possible legal bases of data processing by the Public Administration, thus breaking down the barriers to the use of big data in the adoption of administrative decisions.

This is a legislative intervention not expressly designed for the technology of distributed registers but invites to reflect on the theme of the modalities with which to establish, on the domestic front, the discipline of the functioning of the administrative blockchain (and not only with regard to the profile of the G.D.P.R. compliance, which is the subject of in-depth study here). It could be provided for in a unitary manner by a formal ordinary law or entrusted to secondary regulatory sources, also in order to deal with the risks of obsolescence. On the other hand, it seems less appropriate to leave the regulation of the functioning of the distributed register to the individual lex specialis of the procedure (such as the call for tenders in selective procedures) since, by fragmenting the relevant legal regime, one would be distancing oneself from the creation of a uniform disciplinary standard, which is indeed crucial also with a view to achieving the objective of interoperability between the individual platforms. Moreover, the obvious implications of both regulatory choices in terms of judicial protection cannot be overlooked.

As an alternative to the intervention of the national legislator, while waiting for an organic discipline of the technology, it could be possible to hypothesise a punctual modification of the G.D.P.R. that opens its discipline to the blockchain, in a manner not dissimilar to what is happening, with reference to the other delicate profile of legal compliance of electronic identification methods, with the eIDAS regulation[61].

This is, perhaps, the preferable solution, also in view of the ever-increasing importance of distributed register technology. In fact, the elimination, by means of a regulatory

fulness of personal data processing, establish specifications for determining the controller, the type of personal data which are subject to the processing, the data subjects concerned, the entities to which the personal data may be disclosed, the purpose limitations, the storage period and other measures to ensure lawful and fair processing. It should also be for Union or Member State law to determine whether the controller performing a task carried out in the public interest or in the exercise of official authority should be a public authority or another natural or legal person governed by public law, or, where it is in the public interest to do so, including for health purposes such as public health and social protection and the management of health care services, by private law, such as a professional association".

[60] Published in the Italian Official Gazette of 8 October 2021, no. 241.

[61] The reference is to the proposal for the amendment of the eIDAS Regulation presented by the European Commission in June 2021, which provides, in the new Articles *45 nonies* and *decies,* specific rules on the legal effects of electronic registers and the requirements for qualified electronic registers.

intervention, of any residual doubts regarding the application of the G.D.P.R. could, on the one hand, represent a push towards the concrete application of the instrument (eliminating any alibi for management, traditionally sceptical towards forms of innovation not expressly regulated by law) and, on the other, contribute to the diffusion of a homogeneous administrative language at a Union level that facilitates the increasingly frequent recourse to the model of joint administration.

In this way, a further piece would be added to the emerging and jagged European Union discipline on digital administration[62], to the definition of which the doctrine is also called upon on a theoretical level.

[62] Together with the aforementioned eIDAS Regulation. Consider also the issue of the so-called "reserve of humanity" with respect to automated administrative functions posed precisely by Article 22 of the G.D.P.R., indicated by the Italian Council of State as the basis of the well-known judgment Cons. State, Sez. VI, 8 April 2019, no. 2270 (with commentary by A.G. Orofino and G. Gallone, *L'intelligenza artificiale al servizio delle funzioni amministrative: profili problematici e spunti di riflessione*, in *Giur. It.*, 2020, 1738).