

THE ETHICAL AND LEGAL CHALLENGES OF DATA ALTRUISM FOR THE SCIENTIFIC RESEARCH SECTOR

Ludovica Paseri

Law Department, University of Turin (Italy)

ludovica.paseri@unito.it

EXTENDED ABSTRACT

Scientific research nowadays is increasingly data-driven and therefore requires a growing amount of data, which need to be accessible and of high quality. The data altruism mechanism, that results as a means to meet this demand, is regulated in the Data Governance Act (DGA, hereinafter). The DGA is a Regulation of the European Union, which is applicable from 23 September 2023⁸, that aims to “foster the availability of data for use by increasing trust in data intermediaries and by strengthening data-sharing mechanisms across the EU”, as described in the explanatory memorandum accompanying the proposal for a Regulation⁹. The DGA is a crucial part of the so-called “politics of data” (Pagallo, 2022) developed by the European Commission in 2020¹⁰ and can also be considered as complementary to the Open Data Directive (ODD, hereinafter)¹¹, integrating the European framework on data sharing and reuse (Ruohonen & Mickelsson 2023). The Article 3 of the DGA, which identifies the scope of application, expresses the complementarity between the DGA and the ODD by stating that the DGA provides for the reuse of certain categories of data, such as data held by the public sector that are protected on the basis of commercial confidentiality, statistical confidentiality, protection of third parties’ intellectual property rights and protection of personal data. Therefore, the DGA concerns the reuse of those public sector data excluded from the scope of the ODD (Van Eechoud, 2021, p. 376).

Data altruism is defined by the DGA, in the Article 2(16):

‘data altruism’ means the consent by data subjects to process personal data pertaining to them, or permissions of other data holders to allow the use of their non-personal data without seeking a reward, for purposes of general interest, such as scientific research purposes or improving public services.

This contribution aims to investigate data altruism mechanism for the scientific research sector. This mechanism, based on the voluntary release of data, raises several ethical and legal

⁸ Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), ELI: <http://data.europa.eu/eli/reg/2022/868/oj>.

⁹ Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act), COM/2020/767 final, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52020PC0767>.

¹⁰ European Commission Communication, A European strategy for data, COM/2020/66 final (2020), ELI: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0066>.

¹¹ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast), ELI: <http://data.europa.eu/eli/dir/2019/1024/oj>.

challenges. From the legal viewpoint, the data altruism in the research sector raises the following challenges: (1) the risk of fragmentation; (2) the security concerns; and (3) the duty of control on data altruism organisations.

From the ethical perspective, there is (1) a problem of terminological uncertainty; (2) a need for decoding the concept of altruism underlying the release of data; and (3) a concern related to the autonomy of subjects in giving consent.

The mechanism of data altruism identifies the activities of several actors: (i) the data subject of personal data; (ii) the data holder of non-personal data; (iii) the data altruism organisations; (iv) the data users; and (v) the competent authority for registration. Underlying the operations of this multiplicity of actors are two conditions, expressed in the Article 2(16) of the DGA. The first condition is that reuse by data subjects and data holders must be granted to the data altruism organisation, free of charge, not in return for a reward. This condition may be interpreted as a measure to avoid the establishment of a buying and selling of personal data.

The second condition is that the reuse must pursue general interest purposes. These purposes are specified in the Recital 45, which states: “Such purposes would include healthcare, combating climate change, improving mobility, facilitating the establishment of official statistics or improving the provision of public services. Support to scientific research, including for example technological development and demonstration, fundamental research, applied research and privately funded research, should be considered as well purposes of general interest”.

The mechanism of the data altruism, based on data “voluntarily made available by individuals or companies for the common good” (Proposal DGA, Explanatory Memorandum, 2020, p. 8), generates a considerable impact on the data management in the scientific research sector.

According to the European institutions, the data altruism mechanism involves an articulated process with several phases. First, any entity intending to be recognised as data altruism organisation has to undergo a registration process, and among other information, it is required to declare that “the purposes of general interest it intends to promote when collecting data” (Article 19(4)h, DGA). The general interest is primarily identified by scientific research by the DGA, when presenting the notion of data altruism, the Article 2(16) states that this mechanism shall be realised “for purposes of general interest, such as scientific research purposes or improving public services”. However, it is often difficult to be precise about the aims pursued in a specific scientific research project (Pagallo & Bassi, 2013, p. 183). After that, if the requesting entity meets all the requirements laid down by the DGA, it will be included in the national register of data altruism organisations, by the competent national authority or authorities, within 12 weeks from the date of application, pursuant to the Article 19(5) of the DGA.

The voluntary release of personal data by data subjects, or non-personal data by data holders, to data altruism organisations is based on the provision of the consent. The consent needs to be given in compliance with the two conditions described above, i.e., no reward and public interest purposes. The registered data altruism organisations provide to several natural and legal persons the possibility to process the data they hold, for purposes of general interest, eventually on the basis of a fee. Each data altruism organisation is required to keep accurate records – very similar to the processing register set out in the Article 30 of the General Data Protection

Regulation (GDPR, hereinafter)¹² – concerning a set of accurate information about the specific data processing activities, based on the data altruism consent. In addition, each data altruism organisation, pursuant to the Articles 20 and 21 of the DGA, has several reporting obligations towards data holders. In particular, entities are required to communicate the purposes for which further processing of data is permitted to third parties. In this regard, it is significant that the corresponding article in the proposal of the Regulation (Article 19 of the DGA Proposal) made explicit reference to the duty to communicate “any processing outside the Union”, expression that no longer appears in the wording of the Regulation. In addition, it is relevant (and problematic) that any organisation of data altruism also has a function of control over the entire lifecycle of the data that is given to third parties to process. The Article 21 of the DGA states, in fact, that the “entity shall also ensure that the data is not be used for other purposes than those of general interest for which it permits the processing”. However, the problematic aspect emerging in this phase is represented by the fact that the purpose and regulation of reuse of public sector data is intrinsically generic and open to any possible use of the data and appears from the very definition of reuse (Bassi, 2011, p. 67).

In light of the analysis of the data altruism mechanism, the contribution illustrates the legal challenges, which are: (1) the risk of fragmentation; (2) the security concerns; (3) the duty of control on data altruism organisations.

(1) The risk of fragmentation is generated by the envisaged difficulties in implementing data altruism. Even though the DGA is a Regulation, in the implementation of data altruism the Member States play a decisive role. While waiting to understand how they will implement the data altruism mechanism, it is worth analysing the possible Member States’ strategies and approaches.

(2) Concerning the security, the data altruism organisation must also ensure a solid infrastructure system. The goal is to create pools of data and this data must be stored, transferred, and managed, which makes the infrastructure absolutely central. The centralisation of data always brings with it several challenges from a security viewpoint, making those holding the data both very powerful, and at the same time very vulnerable. Very powerful, because it generates “the emergence of pools of data made available on the basis of data altruism that have a sufficient size in order to enable data analytics and machine learning, including across borders in the Union” (Recital 45, DGA). Highly weak because they are more easily targeted by cyber-attacks and data breaches.

(3) Then, the Article 21 of the DGA establishes a duty to control in charge of any data altruism organisations, over the third parties that are allowed to process the data. Although this requirement is understandable on principle, it does not seem easy achievable in practice. This mechanism of mutual controls seems to strongly refer to the mechanisms of accountability of the GDPR which establish, for the data controller, a set of duties, also with regard to the data processors, those who actually process the data, in the name and on behalf of the data controller (Durante, 2021, p. 134). Admittedly, data altruism organisations must only ensure that these users conduct processing for purposes of general interest. However, the DGA itself considers such purposes in a very broad and general way, referring to scientific research or the

¹² Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) ELI: <http://data.europa.eu/eli/reg/2016/679/oj>.

improvement of public services. These categories are very broad, and without the identification of further boundaries, much can fall within the all-encompassing terms set by the European Regulation.

From the ethical perspective, the contribution intends to shed light on three aspects: (1) the terminological uncertainty; (2) the decoding the concept of altruism; and (3) the individual autonomy of the consent.

(1) As emerged from the description of the different phases of the data altruism mechanism according to the DGA, the concept of 'general interest' is crucial. However, there is terminological uncertainty insofar as the concept of 'public interest' emerges in the proposal of the Regulation. The 'public interest' diverges from the concept of 'general interest' and entails considerable debate if conceived according to the GDPR. In addition, the explanatory memorandum also mentions the concept of 'common good', amplifying the terminological uncertainty that impacts the ethics underlying the data altruism mechanism.

(2) Furthermore, the paper intends to develop the analysis regarding the application of the notion of altruism to the release of data, referring to the studies conducted on the so-called 'data philanthropy' (Taddeo, 2016; Taddeo 2017; Giannopoulou, 2019) that pave the way for the assessment regarding the practical feasibility of this mechanism (Veil, 2022). In light of a well-established trend, it is not difficult to envisage that there might be a fair amount of success on the side of data subjects and data holders who release their personal and non-personal data for the pursuit of general interest purposes (Ienca, 2023, p. 2). Several experiences show a general inclination to release more easily personal data for scientific research purposes (Pagallo, 2022, p. 74). Similarly, it is not difficult to envisage data users, which may include universities, research centres, but also private companies, foundations, etc. It is more difficult to identify entities undergoing the registration process to become data altruism organisations. For this reason, the role of potential private actors performing the function of data altruism organisations and the impact, from an ethical perspective, on the integrity of scientific research deserve further investigation.

(3) Finally, the data altruism is a consent-based mechanism. The goals of the introduction of the GDPR, compared to the previous discipline of the Directive 95/46/EC, was precisely to overcome the model of personal data processing primarily based on consent. The Article 6 of the GDPR provides for a set of mandatory legal bases for the processing of personal data: consent thus becomes one of the possible bases. The *ratio* for this choice made by the European lawmaker in 2016 was precisely to replace a consent-based approach that had proved to be ineffective (Solove 2012; Schermer, *et al.* 2014). It will therefore be necessary to further investigate the impact of the use of consent at the basis of the data altruism mechanism in relation to the choice of the legal bases required for processing personal data for scientific research purposes, according to the GDPR. The analysis of the role of consent in the altruism of data is relevant to the extent that it generates an impact on the personal autonomy of the individual giving such consent.

Given the set of challenges, both legal and ethical, and the multiplicity of actors involved in the data altruism mechanism, it is worth investigating how and whether to develop governance mechanisms that are able to hold together all the aspects at stake.

KEYWORDS: Data governance act, DGA, data governance, data altruism, scientific research, public interest.

REFERENCES

- Bassi, E. (2011). PSI, protezione dei dati personali, anonimizzazione. *Informatica e diritto* 37.1-2, 65-83.
- Durante, M. (2021) *Computational power: the impact of ICT on law, society and knowledge*. New York: Routledge.
- Giannopoulou, A. (2019). Access and Reuse of Machine-Generated Data for Scientific Research. *Erasmus Law Review* 2, 155-165.
- Ruohonen, J., Mickelsson, S. (2023). Reflections on the Data Governance Act. *Digital Society* 2.1, 1-10.
- Ienca, M. (2023) "Medical data sharing and privacy: a false dichotomy?", *Swiss Medical Weekly* 153.1, 1-3.
- Pagallo, U., Bassi, E. (2013). Open Data Protection: Challenges, Perspectives, and Tools for the Reuse of PSI, in Hildebrandt, M., et al. (eds), *Digital Enlightenment Yearbook 2013*. Amsterdam: IOS Press, 179-189.
- Pagallo, U. (2020). The Politics of Data in EU Law: Will It Succeed? *Digital Society* 1.3, 1-20.
- Pagallo, U. (2022). *Il dovere alla salute. Sul rischio di sottoutilizzo dell'intelligenza artificiale in ambito sanitario*, Milano-Udine: Mimesis.
- Schermer, B. W., et al. (2014) The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology* 16.2, 171-182.
- Solove, D. J. (2012) Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review* 126, 1880-1903.
- Taddeo, M. (2017). Data philanthropy and individual rights. *Minds and Machines* 27.1, 1-5.
- Taddeo, M. (2016). Data philanthropy and the design of the infraethics for information societies. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2083, 1-12.
- Van Eechoud, M. (2021). A Serpent Eating Its Tail: The Database Directive Meets the Open Data Directive. *IIC - International Review of Intellectual Property and Competition Law*, 52, 375-378.
- Veil, W. (2022). Data Altruism: How the EU is Screwing up a Good Idea. Discussion paper. *AlgorithmWatch*, 1-8.