Open Knowledge. Access and Re-Use of Research Data in the European Union Open Data Directive and the Implementation in Italy

Marta Arisi*

Abstract

This paper provides initial observations on the inclusion of scientific research data in the scope of the EU Public Sector Information Directive of 2019, Directive (EU) 2019/1024, also known as the Open Data Directive, related rules for the re-use of such data enshrined in Art 10, and the implementation in Italy with the decreto legislativo 8 November 2021 no 200. The work seeks to examine how the EU Public Sector Information rules on research data – and, to a lesser extent, data from cultural establishments – may contribute to the objectives of Open Knowledge, elected as an umbrella term with primary reference to Open Access, Open Science, and Open Data, given the difficulties of identifying exhaustive conceptual contours for them. In order to do so, the paper critically examines the exemptions and safeguards related to Intellectual Property and Personal Data protection and identifies the circumstances under which these may obstruct the re-use of research data.

I. Introduction

The present paper analyzes the inclusion of scientific research data in the scope of the Directive (EU) 2019/1024 on open data and the re-use of public sector information (PSI), also known as the Open Data Directive,[1] and related rules for the re-use of research data. The paper is informed by the concept of open knowledge and critically examines the mentioned rules from such perspective. This is to be understood as an umbrella term with primary reference to open access, open science, and open data, given the difficulties of identifying exhaustive conceptual contours for them, and since terms are often used interchangeably. Access and re-use of research data is the focus of the work, while data from cultural establishments is also briefly considered, due the latter are vital part of the open knowledge narrative. The analysis will especially consider the numerous intersections of the EU PSI subject matter with intellectual property and data protection laws and explore how related exemptions and

* Holder of a research grant at the University of Trento, Italy, Faculty of Law. I would like to thank my colleagues at the University of Trento, as well as my anonymous reviewers, for their helpful comments. All errors remain my own.

[1] 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 [2019] OJ L172/56.

safeguards may to some extent represent obstacles to the re-use of research data. The ultimate objective is to shed light on the rules recently introduced in Italy with the decreto legislativo 8 November 2021 no 200, transposing the Open Data Directive into national law, and potential discrepancies in relation to the objectives of open knowledge – that, to put it simply, calls for a more open re-use of research data and data from cultural establishments.

The work is structured as follows. Para II begins by tracing the development of EU public sector information rules, from the PSI Directive of 2003, Directive 2003/98/EC,[2] later amended in 2013 with Directive 2013/37/EU,[3] until the most recent Directive of 2019, and examining the debate that led to the introduction of the rules on research in Art 10. Para III focuses on the provisions that detail the scope of application of rules on scientific research, and relevant exemptions.

Para IV attempts to give a more detailed account of the rules on research data set out in the Open Data Directive and it is organized in three different sub-paras. After illustrating the core rules to be applied in sub-para 1, sub-para 2 and 3 critically examine the exemptions and safeguards related to copyright law and personal data protection. In addition, para V offers a brief overview of the PSI rules on data from cultural establishments as it seems useful to compare the status of research data and cultural data in the Open Data Directive, being reputed equally fundamental elements of open knowledge.

Finally, building on the previous paragraphs, the paper proceeds with a detailed analysis of the Italian transposition of the Open Data Directive in para VI. Brief conclusive remarks follow.

II. Public Sector Information Rules in the European Union

The present paragraph briefly describes the development of the PSI rules in the European Union, focusing on the lively debate on research data and the path that led to including it into its scope, while offering insights into the broader policy and legislative context of such amendment.

1. The Public Sector Information Directives in the European Union: Main Characteristics and Rationale

The acknowledgment of the potential of PSI in the EU should be primarily traced back to the Green Paper of the Commission in 1999,[4] but the first legislative

[2] Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information [2003] OJ L345/90.

[3] Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information [2013] OJ L175/1.

[4] European Commission Communication, 'Public Sector Information: A Key Resource in Europe, Green Paper on Public Sector Information in the Information Society' [1998] COM(1998) 585 final. R. Sanna, *Dalla trasparenza amministrativa ai dati aperti, Opportunità e rischi delle*

action taken by the EU is the Directive of 2003. The Directive called on Member States to adopt a set of minimum harmonized rules (eg including redress mechanism, time limit for answering requests, fees, and transparent conditions thereof) governing the re-use of certain documents held by public sector bodies – despite relevant exclusions. At the same time, member States were also free to enact more permissive rules.

In the opinion of many, the subsequent reform of 2013 introduced an obligation for member States to make certain documents re-usable.[5] Such a mandate would emerge from the conjunct reading of Art 3(1) of the Directive,[6] as amended, and recital 8 of the PSI Directive of 2013.[7] However, on closer inspection, such an obligation for re-use would be rather limited: in particular, it would only apply to the documents that are not excluded by the scope of the Directive, which essentially referred to provisions to be detailed by Member States and was further circumscribed by several safeguards.

This still seems true after the latest overhaul of 2019, despite the material and subjective scope of the PSI rules having expanded. The Directive (EU) 2019/1024 on open data and the re-use of public sector information is a recast that brings together the amendments made to the previous acts and represents the output of a revision process started between 2017 and 2018.[8] The new essential elements of the Open Data Directive are the introduction of research data in its scope and the introduction of the principle of 'open by design and default' in Art 5(2) of the new Directive.[9] Most notably, the new Directive also has a different title, which includes – next to the re-use of public sector information – open data, although its open vocation remains to some extent unclear. This is more thoroughly discussed in relation to the topic of research data in para IV.

In the new Directive, the member States' obligation to allow re-use of public sector data remains substantially limited by a detailed scope of application, with

autostrade informatiche (Torino: Giappichelli, 2018), 1.

[5] ibid 253, 257; M. Van Eechoud, 'Making Access to Government Data Work' 9(2) *Masaryk University Journal of Law and Technology*, 61, 64 (2015).

[6] Art 3(1) of Directive 2003/98/EC, as amended, recites: 'Subject to paragraph 2 Member States shall ensure that documents to which this Directive applies in accordance with Article 1 *shall be re-usable* for commercial or non-commercial purposes in accordance with the conditions set out in Chapters III and IV'.

[7] Recital no 8 of Directive 2013/37/EU recites: 'Directive 2003/98/EC should therefore be amended to lay down *a clear obligation* for Member States to make all documents re-usable unless access is restricted or excluded under national rules on access to documents and subject to the other exceptions laid down in this Directive. The amendments made by this Directive do not seek to define or to change access regimes in Member States, which remain their responsibility'.

[8] See Procedure 2018/0111/COD, available at https://tinyurl.com/2p8ctbwv (last visited 30 June 2022).

[9] This recalls 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 [2016] OJ L119/1 (hereinafter GDPR), Art 25 titled 'Data protection by design and by default.'

several exemptions and safeguards provided in Art 1. However, member States are specifically encouraged to go beyond the minimum requirements and apply the related rules to documents held by public bodies as well as private undertakings providing services of public interest,[10] while being exhorted to establish policies that would permit a more extensive re-use of data.[11] Ultimately, the new PSI rules also signal the intention to fit into the emerging technological context, since significant progress has been made from the first Directive of 2003, as for instance considering artificial intelligence applications, distributed ledgers, the Internet of Things and smart cities.[12] Provisions on dynamic data, subject to frequent updates, have been introduced.[13]

Even after the most recent evolutions, it remains true that the rationale of the EU PSI rules is strengthening the internal market as regards information services.[14] The underlying assumption is that if information retained by public sector bodies is free for re-use, it can generate positive and essential contributions to the EU Internal market.[15] The private sector could therefore benefit from re-use of public data not only because this would allow government oversight and democracy, but because it would enable data users to create innovation.

Authors underline the need to distinguish between what is usually regarded as an economic right (the re-use) versus a civic right (the access),[16] and suggest that the main goal of the PSI rules differs from the so-called Freedom of Information legislation (also FOI), aimed at enhancing transparency and participation of citizens in the *res publica*.[17] Although their different rationale may be evident, it

[10] Directive (EU) 2019/1024, recital 19.

[11] ibid recital 20.

[12] ibid recitals 3, 9, 13; European Commission Staff Working Document, Impact assessment Accompanying the document Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information SWD(2018) 127 final [2018], 7.

[13] Directive (EU) 2019/1024, Art 2(2)(e).

[14] Directive 2003/98/EC, as amended, recitals 3, 5, 9; Directive (EU) 2019/1024, recitals 7-9.

[15] COM(1998) 585 final n 4 above, 1; C. Sappa, 'Selected intellectual property issues and PSI re-use' 6(3) *Masaryk University Journal of Law and Technology*, 445, 447 (2012); K. Janssen, 'The influence of the PSI directive on open government data: An overview of recent developments' 28 *Government Information Quarterly*, 446, 447 (2011). See also T. Streinz, 'The Evolution of European Data Law', in P. Craig and G. de Búrca eds, *The Evolution of EU Law* (Oxfod: OUP, 3rd ed, 2021), 27: the author cites the European Commission Guidelines for improving the synergy between the public and private sectors in the information market (1989).

[16] P. Keller et al, 'Re-use of public sector information in cultural heritage institutions' 6(1) *International Free and Open Source Software Law Review*, 1, 2 (2014).

[17] In the EU, a right of access to documents of the Union's institutions, bodies, offices and agencies, is currently enshrined in Art 42 of the Charter of Fundamental Rights of the European Union [2012] OJ C326/391 and Art 15 Consolidated version of the Treaty on the Functioning of the European Union [2012] OJ C326/1. The first EU Regulation on the matter appeared in 2001, two years before the first PSI Directive of 2003: Regulation (EC) 1049/2001 regarding public access to European Parliament, Council and Commission documents [2001] OJ L145/43. Absent the EU competence to ensure access to documents held by public bodies at a national level, the matter of access to information from national public sector bodies has been primarily regulated at the national level. M. Salvadori, 'Right of Access to Documents: The Implementation of Article 42 of the Charter of

is not always easy to trace a strict line of separation between the FOI and PSI laws because of relevant overlaps.[18] However, one conspicuous observation is that PSI rules do not grant access to information, but only address re-use thereof. More specifically, as expressively reiterated in the Directive of 2003, later amended in 2013,[19] as well as in the new Open Data Directive,[20] the PSI rules build on national access regimes and are without prejudice to them, so that which public sector information can be accessed and ultimately re-used still remains determined by member States at the national level.[21] It seems plausible that the confusion between the two subject matters is currently exacerbated, since both are increasingly informed by open knowledge,[22] where the notion of open government data is becoming the subject of scholarly attention.[23] As an example, the relevant sets of rules for FOI and PSI may both refer to 'open' definitions, as in the case of Italy, described in para VI.

2. The Inclusion of Research Data and the Evolutions of the Public Sector Information Rules in the European Union

Documents held by educational and research establishments, such as schools, universities, archives, libraries, as well as by research institutes were excluded by the scope of the first PSI Directive of 2003.[24] The possibility to extend the scope of the Directive to both the educational and research sectors was supported by respondents to the public consultation opened in 2010.[25] Following a lively

Fundamentals Rights', in M. Biasiotti and S. Faro eds, *From Information to Knowledge - Online Access to Legal Information: Methodologies, Trends and Perspectives* (Amsterdam: IOS Press, 2011), 2-3.

[18] K. Janssens, n 15 above, 447 describes the possible origins of this confusion, to be also linked to the first years of the transposition by member States, and related risks for freedom of information rights. Proposing a conceptual distinction between access, dissemination and re-use of public sector information A. Cerrillo-i-Martinez, 'Fundamental interests and open data for re-use' 20(3) *International Journal of Law and Information Technology*, 203, 205-214 (2012).

[19] Directive 2003/98/EC, as amended, recital 9, Art 1(3).

[20] Directive (EU) 2019/1024, recitals 18, 23.

[21] J. Andrasko and M. Mesarcik, 'Quo Vadis Open Data' 12(2) *Masaryk University Journal of Law and Technology*, 179, 187 (2018).

[22] In particular, the FOI legislation in European Union seems to be evolving towards open models, according to Open Government, Open Government Data and also E-Government trends. F. Faini, *Data Society* (Milano: Giuffrè, 2019), 12-22. International Conventions on the subject matter have also appeared, most notably Council of Europe Convention on Access to Official Documents [2009], Council of Europe Treaty Series - No. 205.

[23] M. Dulong de Rosnay and K. Janssen, 'Legal and Institutional Challenges for Opening Data across Public Sectors: Towards Common Policy Solutions' 9(3) *Journal of Theoretical and Applied Electronic Commerce Research*, 1, 3, (2014); D. Arcidiacono and G. Reale. 'Open Data as a Commons? The Disclosure of Public Sector Information from a Comparative Perspective' *Rassegna Italiana Di Sociologia*, 235, 237-239 (2018).

[24] Directive 2003/98/EC, Art 1(2)(e).

[25] H. Richter, 'Open Science and Public Sector Information – Reconsidering the exemption for educational and research establishments under the Directive on re-use of public sector information' 9 *JIPITEC,* 51, 55 (2018); European Commission Staff Working Paper, Impact Assessment

debate, the rules were only partially amended in 2013 to cover data from cultural establishments.

The Staff Working Paper that preceded the reform of 2013 contains a few helpful insights in this regard. While the potential value of sharing research data and making it publicly available was not denied,[26] one initial argument presented to disallow research data from the scope of the Directive was that this material would be covered by intellectual property or other third-party rights.[27] This argument seems unconvincing because data should in principle be excluded by copyright, in line with the well-established idea/expression dichotomy, enshrined in the major international codifications.[28] The principle has been eroded in time by a controversial and well-discussed trend of closure in the most recent copyright reforms.[29] However, the dichotomy remains paramount to safeguarding public interests when discussing copyright, data and emerging applications, as emerges from the scholarly debate on copyright, text and data mining and algorithms.[30] Nevertheless, while the Working Paper acknowledged that Intellectual Property Rights (IPR) protection 'does not extend as far as pure research data', it added there are often unclear boundaries between types of data and the status of third-party rights, as well as differences in 'researchers' attitudes, patterns of behavior and needs or in the existence and robustness of available infrastructure'. Overall, this would imply that the burden to clarify the status of research data could exceed the related benefits.

Another main argument for excluding research data from the material scope of the Directive was the approach that the Open Access (hereinafter OA) debate was a separate, although parallel, discussion channel for disseminating and exploiting research findings and results.[31] Considering the initiatives on open knowledge at the time, the most important were identified in non-binding documents. The European Commission Communication 'Towards access to better

accompanying the document Proposal for a Directive of the European Parliament and the Council amending European Parliament and Council Directive 2003/98/EC on the re-use of public sector information, SEC(2011) 1152 final [2011] 67-69.

[26] ibid 33.

[27] ibid 17, 33.

[28] Most notably, Art 2 of the World Intellectual Property Copyright Treaty (1996) reports: 'Copyright protection extends to expressions and not to ideas, procedures, methods of operation or mathematical concepts as such'.

[29] J.P. Barlow, 'Selling Wine Without Bottles: The Economy of Mind on the Global Net' 18 *Duke Law and Technology Review*, 24 (2019); J. Boyle, *The Public Domain: Enclosing the Commons of the Mind* (Yale University Press, 2008). The most important evidence thereof being the creation of *sui generis* database rights. The topic is linked to the emerging debate on data ownership in the EU: M.L. Montagnani and A. Von Appen, 'IP and Data (Ownership) in the New European Strategy on Data' 43 (3) *European Intellectual Property review*, 156 (2021).

[30] Discussing freedom of expression and Text and Data Mining: R. Ducato and A. Strowel, *Ensuring Text and Data Mining: Remaining Issues With the EU Copyright Exceptions and Possible Ways Out,* CRIDES Working Paper Series no 1/2021, 8-9.

[31] SEC(2011) 1152 final, n 25 above, 17, 34.

scientific information' of 2012[32] and the 'Recommendation on access to and preservation of scientific information' of 2012[33] promoted measures to ensure that the results of Europe's publicly funded research, including both publications and data, are accessible. Moreover, relevant steps were being taken as regards EU-funded projects (FP7 - Seventh framework program from 2007 to 2013 and most notably its successor Horizon 2020). Against this backdrop, the Working Document implied that only such initiatives could take into account the specificities and limitations of the research sector, while the 'generic' PSI debate, despite very close objectives, could not tackle the issue.[34] One last remark referred to the difficulties in establishing a clear terminology to limit the application of the PSI Directive – ie, with regard to research institutions.[35] Defining research institutes at EU level was considered an 'impossible endeavor', since member States' traditions differ, but also appeared disproportionate to the issue, failing the subsidiarity scrutiny.

A possible explanation for the recent changes may be primarily framed within the fostering of EU regulatory efforts to enhance open scientific research, to the point that the argument about OA being the separate channel to promote the wider availability and reuse of research data seems to have been superseded. In fact, commenting on the new proposal of the Directive, influential doctrine suggested the potential re-union of two worlds that were conceived as separate: the scientific OA world and the general PSI world.[36] First, the initial Recommendation on access to and preservation of scientific information of 2012 was replaced by the Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information,[37] calling on member States to adopt measures for the dissemination of, and open access to, both scientific publications and research data resulting from publicly funded research activities. The Recommendation's objectives and goals resemble the new rules on research data set out in the Open Data Directive,[38] but only the latter is provided with binding force concerning the objectives. Second, the premise of the impact assessment conducted on 2018 and accompanying the proposal for a reformed

[32] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Towards better access to scientific information: Boosting the benefits of public investments in research [2012] COM(2012) 401 final.

[33] European Commission Recommendation 2012/417/EU of 17 July 2012 on access to and preservation of scientific information [2012] OJ L 194/39.

[34] SEC(2011) 1152 final, n 25 above, 27.

[35] ibid 34.

[36] H. Richter, n 25 above, 52.

[37] European Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information [2018] OJ L134/12.

[38] In particular, the latest Recommendation calls for the adoption of clear policies, to be detailed in national plans, for the management of research data resulting from publicly funded research, including open access, in Point 3 of the Recommendation. Point 4 declares that member States should ensure the implementation of policies and national plans by research funding institutions responsible for managing public research funding and academic institutions receiving public funding.

Directive[39] explicitly linked the reform to the EU international commitments for opening research data,[40] including the OECD Council Recommendation of 2010[41] and the G8 Open Data Charter in 2013.[42] The impact assessment criticized the insufficient availability of research data for re-use,[43] indicating different factors: the fact that policies are fragmented, not fit for purpose and partially outdated, scarce focus on re-use compared to access and incentives, and a complex reality of different data sharing cultures in the scientific community.[44] In addition, the Consultation on output between June 2017 and late January 2018 was in favor of reviewing the scope of the PSI Directive to include research establishments.[45] As a result, different policy options were presented in the impact assessment, including adding top-down European legislative open access mandate for both publication and research data in the PSI or, as a second option, covering only research data that would have been made available as a result of open access mandate; in any case, the assessment affirmed the need to update the recommendations on access to and preservation of scientific information.[46] The second, low intensity option was eventually chosen.[47]

In addition to this, the introduction of rules on research data in the PSI Directive of 2019 should also be examined considering how the EU policy and legislative initiatives have converged towards data driven innovation, while increasingly urgent discourses on data ownership are emerging.[48] From this perspective, the dispositions on research data in the new PSI Directive 2019 may enhance the role of research data in the data economy, an objective presented in the so-called EU Open Data Policy.[49] The Digital Single Market

[39] European Commission, Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast) [2018] COM (2018) 234.

[40] SWD(2018) 127 final, n 12 above, 3.

[41] Organisation for Economic Co-Operation and Development (OECD) Recommendation Of The Council For Enhanced Access And More Effective Use Of Public Sector Information [2008] C(2008)36.

[42] G8 Open Data Charter and Technical Annex (2013), available at https://tinyurl.com/bddw3k46 (last visited 30 June 2022).

[43] SWD(2018) 127 final, n 12 above, 15.

[44] ibid 16.

[45] ibid 64-65.

[46] ibid 30-32.

[47] ibid 49.

[48] M.L. Montagnani, 'Dati e proprietà intellettuale in Europa: dalla "proprietà" all'"accesso" ' *Il diritto dell'economia*, 539 (2020); A. Wiebe, 'Protection of Industrial Data – a New Property Right for the Digital Economy?' 12(1) *Journal of Intellectual Property Law & Practice*, 62 (2017); H. Zech, 'A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data' 11(6) *Journal of Intellectual Property Law & Practice*, 460 (2016); V. Zeno-Zencovich, 'Do "Data Markets" Exist?' *MediaLaws.eu*, 23 July 2019, 17-18, available at https://tinyurl.com/2d6awywk (last visited 30 June 2022).

[49] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Open data - An engine for innovation, growth and transparent governance Communication [2011] COM(2011) 882 (also referred to as the EU Open Data Policy). This promoted the creation of an EU Open Data Portal; see

Strategy in Europe in 2015 also promoted a strong link with research and open science, envisioned in the launch of the European Open Science Cloud (EOSC).[50] Besides, it is noteworthy that the proposal for the new Open Data Directive was published the same day that the EU Commission also proposed the Communication Towards a Common European Data Space, together with a Guidance on Sharing Private Sector Data in the European Data Economy.[51]

Beyond the Open Data Directive, the cornerstone of such current developments should be identified in the Data Strategy of 2020.[52] This describes the data driven innovation potential as pervasive, also for the realization of the EU Green Deal,[53] and emphasizes the availability of data for the public good,[54] providing examples of both data generated by the public sector and data from the private sector. Most relevantly, considering public sector information, the proposal for a Data Governance Act[55] was presented in November 2020. Art 3 of the Proposal details measures that facilitate the use of some categories of data held by public sector bodies. Moreover, the proposal for the so-called Data Act[56] was published very recently in February 2022. This allows for public sector bodies to access and use data held by the private sector when this is necessary due to exceptional circumstances – ie, in case of a public emergency – or to implement a legal mandate if data are not otherwise available. On this point, initial reactions have outlined that the proposal introduces an exception to the general prohibition to re-use the obtained data, for the use of scientific research and in a public interest context.[57] These acts, once final and implemented, will therefore prove decisive in

European Union, Open Data Portal webpage, available at https://data.europa.eu.

[50] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A Digital Single Market Strategy for Europe Communication Digital Single Market Strategy [2015] COM(2015) 192 final. This acknowledges the role of research in the data economy, linking this to Open Science and announcing the European Cloud initiative including the Open Science Cloud (EOSC). The latter was promoted with the European Commission Communication Building a competitive data and knowledge economy in Europe [2016] (COM(2016) 178 final.

[51] B. Gonzalez Otero, 'Evaluating the EC Private Data Sharing Principles Setting a Mantra for Artificial Intelligence Nirvana?' 10 *JIPITEC*, 65, 66 (2019).

[52] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A European strategy for data [2020] COM(2020) 66 final.

[53] ibid 1.

[54] ibid 6-8. More specifically, four key-cases are identified: 1) data of the public sector is used by the business; 2) data is used and shared from business-to-business; 3) data of the business is shared with the public sector; 4) different public authorities share the data.

[55] Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) COM/2020/767 final.

[56] Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act) [2022] COM(2022) 68 final.

[57] F. Vogelezang and A. Takowski, 'Data Act: Business to Government Data Sharing' *Open future*, 23 February 2022, available at https://tinyurl.com/f8udecfz (last visited 30 June 2022). More specifically, Art 21 of the aforesaid proposed Regulation would permit that public bodies make data available to individuals and organizations that conduct scientific research, or statistics institutions, at

applying the provisions of the Open Data Directive.

III. Research Data and the Directive (EU) 2019/1024: Scope of Application and Relevant Exemptions

The scope of application of the Directive is primarily detailed in Art 1, while Art 2 contains definitions.[58] According to Art 1(1) the Directive applies to three main groups of documents: a) existing documents held by public sector bodies of the member States, b) existing documents held by certain public undertakings and, as recently introduced by the Directive of 2019, c) research data, pursuant to the conditions established under Art 10.

On the other hand, Art 1(2) details the documents to which the Directive does not apply. While Arts 1(2)(a) and (b) exclude certain documents held by public bodies or public undertakings, the following letters (c) to (d) contain more specific exemptions that essentially refer to the existence of rights and interests. Only a few of these exemptions are covered by the present paragraph. More specifically, this tries to outline which research data are covered by the scope of application of the Directive, what are the limitations deriving from intellectual property and data protection laws and, finally, whether there are other relevant limitations to re-use.

1. Research Data and Its Subjects

Art 1(1)(c) affirms that research data are amongst the documents to which the Directive applies, pursuant to the conditions set out in Art 10. Research data in Art 9 no 6 of the Directive is defined as 'documents in a digital form, other from scientific publication' that can either be collected or produced in the course of scientific research activities and used as evidence in the research process or, alternatively, be commonly accepted in the research community as necessary to validate research findings and results. The difference between research data and scientific articles is also found in recital 27, that provides a few examples: research data would include 'statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images', but also 'meta-data, specifications and other digital objects'.

Art 10 is the provision which defines not only conditions for access and re-use of research data but the material scope of application of related rules. As a premise, Art 10(1) calls on member States to adopt policies for making research data available addressed 'to research performing organizations and research funding organizations'; Art 10(2) on the other hand states that research data

least when these are no-profit or operate in the context of a public-interest mission.

[58] For instance, document (Directive (EU) 2019/1024 Art 2(1) no 6, research data (Directive (EU) 2019/1024 Art 2(1) no 9) or re-use (Directive (EU) 2019/1024 Art 2(1) no 11).

shall be re-usable for commercial and non-commercial purposes in accordance with Chapters III and IV. More precisely, Art 10(2) establishes two ground and cumulative conditions for the rules to apply: first, research data should be 'publicly funded'. What is deemed public funding (eg considering potential complementation by other sources of funding) is, however, not defined by the Directive nor otherwise easy to establish. Existing rules and criteria are difficult to identify and apply across member States, as well as at the national level, when they are present, for the subject matter may be regulated differently across different scientific fields or legal areas. Examples thereof are the so-called secondary publishing rights in copyright law.[59] Recital 28 seems of some relevance in this regard: building on the fact that open access policies would always be limited and not absolute, as for intellectual property reasons or national security reasons, recital no 28 affirms that certain obligations stemming from this Directive

> 'should be extended to research data resulting from scientific research activities subsidized by public funding or co-funded by public and private-sector entities'.

The recital could thus be interpreted that Member States should apply open policies when funding is even partly public, suggesting the introduction of flexible rules for the definition of what constitutes publicly funded research.

Second, for the rules to apply, researchers, research performing organizations or research funding organizations must have *already*[60] made the research data publicly available through an institutional or subject-based repository. According to recital 28, Member States could also extend the application to other data infrastructures, through open access publications, as an attached file to an article, a data paper or a paper in a data journal. The most striking aspect of this provision is that it refers to the behaviors of researchers, research performing organizations or research funding organizations. Commentators on the proposal observe how such a rule could impact the personal incentives and the informal norms of research communities, which traditionally represent the main drivers for disseminating scientific information and knowledge.[61]

One initial question to be answered is whether research data should be considered only the data produced by research organizations or include other types of organizations as well. The hereby described rules seem not to refer only to research organizations. The requirement that data is produced only by research

[59] See for instance ReCreating Europe - Rethinking digital copyright law for a culturally diverse, accessible, creative Europe, Horizon 2020 funded project, grant agreement n. 870626, Webinar: Secondary Publishing Right: Exploring Opportunities and Limitations. Video available at https://tinyurl.com/yc5h2trw (last visited 30 June 2022).

[60] This is further explained by recital 28, which links the reason for the requirement to the opportunity to avoid administrative burdens, but also not impose extra costs for the retrieval of the datasets, or require additional curation of data.

[61] H. Richter, n 25 above, 74.

organizations does not emerge in Art 1(1)(c), Art 9 nor Art 10. Moreover, considering exclusions, Art 1(2)(l) basically affirms that the Directive does not apply to the documents held by research performing organizations and research funding organizations (including organizations established for the transfer of research results), unless they are research data as defined by Art 1(1)(c), pursuant to the conditions further explained in Art 10. In addition to this, Art 1(2)(k) merely excludes that the Directive would apply to documents held by educational establishments of secondary level and below, and, in the case of all other educational establishments, documents other than those referred to in Art 1(1)(c). Therefore, a comprehensive reading of these provisions reasonably leads to the conclusion that when research is publicly funded, regardless of the type of organization, the related rules would apply.

Ultimately, it does not emerge clearly who the subjects are to which the obligations on re-use should apply. As mentioned above, Art 10(2) states that research data shall be re-usable for commercial and non-commercial purposes in accordance with Chapters III and IV. These Chapters include rules addressed to public sector bodies or public undertakings (ie Art 5 and following). What is more, recital 28 seems to confirm the research organizations targeted by the rules on research data are not public sector bodies or public undertakings only. The recital affirms that 'research performing organizations and research funding organizations could also be organized as public sector bodies or public undertakings'; in this case, the Directive should apply to such 'hybrid' organizations 'only in their capacity' as research performing organizations and to their research data.[62]

Overall, opting for a comprehensive reading of Art 10(1), Art 10(2), and related recitals 27 and 28, it seems realistic that a more precise definition of such subjects will to some extent be referred to member States, since they will address the open access policies to research performing organizations and research funding organizations for making publicly funded research more available. In addition, referring to recital 28, a positive element for enhancing re-use of research data is the interpretation that, on the one hand, member States may be required ('it is appropriate to set an obligation') to adopt and implement policies on publicly funded research data to be applied by all research performing organizations and research funding organizations.[63] On the other hand, Member States may possibly ('certain obligations stemming from this Directive should') extend the related obligations to scientific research activities

[62] S. Gobbato, 'Open Science and the reuse of publicly funded research data in the new Directive (EU) 2019/1024' 2(2) *Journal of Ethics and Legal Technologies*, 145, 153-154 (2020).

[63] Directive (EU) 2019/1024, recital 28: 'For the reasons explained above, it is appropriate to set an obligation on Member States to adopt open access policies with respect to publicly funded research data and ensure that such policies are implemented by all research performing organisations and research funding organisations […]'.

subsidized by public funding or co-funded by public and private-sector entities.[64] While this formulation of Art 10 allows for greater flexibility at the national level, the result can be criticized in terms of legal certainty and harmonization.

2. Relevant Exemptions for Research Data

As previously mentioned, only a few exemptions are covered in further detail by the present work, due to the importance of IPR in research data, and the delicate relationship between research and personal data protection. The present paragraph additionally explores what may be the other relevant exemptions prescribed by the Open Data Directive that would affect the application of the Directive to research data.

Considering intellectual property first, the related exemptions in the Directive are found in Art 1(2)(c), which excludes documents for which third parties hold intellectual property rights. Logos, crests, and insignia are excluded by virtue of Art 1(2)(g). Recital 54 confirms that documents covered by industrial property rights are excluded, such as patents, registered designs, and trademarks. Such limit of third-party intellectual property rights requires further attention for it seems partially unclear. The provision is first supported by recital 54, affirming that property rights of third parties are to be understood as being different from the intellectual rights held on materials by the public bodies themselves. It is also held that third parties' rights shall not be affected by the Directive; similarly, rights of public sector bodies or public undertakings shall not be affected by the Directive, and the exercise of the same rights shall not be limited by the Directive. What may generate some confusion, however, is the example provided by recital 55, explaining the case whereby a document is 'held' by cultural establishments,[65] 'if a third party was the initial owner'. In this case, the recital affirms, the document should be reputed a document for which third parties hold intellectual property rights for the purpose of the Directive. Therefore, this may be the case for all licensing agreements, even though these may allow for certain uses of the works, and more generally all cases in which intellectual property rights have not expired or have always been attributed to the subject concerned. Problematically, this excludes a great number of cases from the application.

A major exemption also regards data protection and privacy, as detailed in Art 1(2)(h). This provision essentially refers to national laws to define the extent

[64] Directive (EU) 2019/1024, recital 28: 'For that reason, certain obligations stemming from this Directive should be extended to research data resulting from scientific research activities subsidised by public funding or co-funded by public and private-sector entities […]'.

[65] Directive (EU) 2019/1024, recital 55 reports: 'If a third party was the initial owner of the intellectual property rights for a document held by libraries, including university libraries, museums and archives and the term of protection of those rights has not expired, that document should, for the purpose of this Directive, be considered to be a document for which third parties hold intellectual property rights […]'.

to which documents containing personal data could be included in the scope of the Directive. This excludes the documents – or parts thereof – where access is limited by national access regimes on grounds of personal data protection or otherwise deemed adverse for personal data protection and privacy concerns by national laws. More specifically, the Directive would not apply to documents to which access is excluded or simply restricted by virtue of those access regimes on grounds of protection of personal data, which may diverge across member States. Moreover, the Directive would also not apply to parts of documents that would be accessible by virtue of those national regimes and that contain personal data, when their re-use is defined by the law, alternatively, as 'incompatible with the law concerning the protection of individuals with regard to the processing of personal data',[66] or – as of 2019 – also 'undermining the protection of privacy and the integrity of the individual'. This should, however, be in accordance with Union or national law regarding the protection of personal data.

Focusing on research data, other exemptions which deserve to be mentioned are the following. Art 1(2)(d) excludes documents 'such as sensitive data'. The Directive would not apply when access is excluded by national access regimes on grounds of national security, but also statistical confidentiality and commercial confidentiality. On this point, it should be noted that it is not easy to grasp how such concepts would apply to research data as defined in the Directive. It is not immediately clear whether commercial secrecy could be perfectly identified within the EU subject matter of trade secrets, which are regulated by Directive (EU) 2016/943 on trade secrets.[67] Indeed, commercial confidentiality in the PSI Directive is defined as including business, professional or company secrets, while the Trade Secrets Directive refers to information that is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question; second, such information has commercial value because it is secret and has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret.[68]

Other relevant exemptions are presented in Art 1(2)(e) referring to the Directive on critical infrastructures[69] and Art 1(2)(f). These provisions reiterate that access to administrative documents remains governed at the national level:

[66] Art 29 Working Party, Opinion no 6/2013 on open data and public sector information ('PSI') reuse (2013), 10- 11.

[67] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1.

[68] Directive (EU) 2016/943 Art 2 no 1.

[69] Directive 2008/114/EC of 8 December 2008 on the identification and designation of European critical infrastructures and the assessment of the need to improve their protection [2008] OJ L345/75.

those documents which can be accessed upon proof of particular interest should be excluded from the scope of application. Finally, it can be added that the documents subject to the so-called INSPIRE Directive, Directive 2007/2/EC,[70] and thus including spatial data, are expressively included in the scope of application of the Directive when they are held by public sector bodies and public undertakings, by virtue of Art 1(7).

IV. Research Data: Analysis of Art 10 of the Directive (EU) 2019/1024

The present paragraph attempts to give a more detailed account of rules for research data set out in Art 10 of the Open Data Directive. After illustrating the core principles and rules to be applied (sub-para 1), the objective is to critically examine safeguards and limits provided with reference to copyright law (sub-para 2) and data protection law (sub-para 3). The analysis tries to identify the circumstances under which these provisions may obstruct the re-use of research data.

1. Principles and Rules for the Re-Use of Research Data

The rules on research data in the Open Data Directive are accompanied by a set of principles in Art 10(1) and related recitals, including open access policies, open by default principle, FAIR principles, and the principle of 'as open as possible, as closed as necessary' (see also figure 1 below). A brief conceptual reordering of the complex interplay of different open concepts, primarily including open access, open science, open data, and open knowledge, shall help to understand which open practices the Directive effectively promotes.

The link between the new PSI rules on research data, open access and Open Science (OS) already emerged in examining the debate on their introduction. Both OA and OS are to be considered consistent with the freedom of scientific literature and research.[71] The first part of Art 10 calls on member States to

[70] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) [2007] OJ L108/1.

[71] T. Margoni et al, 'Open Access, Open Science, Open Society', *Trento Law and Technology Research Group Research Paper no 27*, 1, 6-9 (2016). There is extensive literature on this point. For a very influential literature review on Open Science, B. Fecher and S. Friesike, *Open Science: One term, Five schools of thought*, RatSWD Working Paper Series, 2013. The main elaborations of the movement could be considered the so-called BBB Declarations - having been proclaimed, respectively, in Budapest, Berlin, Bethseda, which are all dated by the first years of the 21st century and refer to the Net as the emergent tool to access and share knowledge: Open Society Institute (OSI), Budapest Open Access Initiative in 2001; Max Planck Institute, Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities [2003]; Bethesda Statement on Open Access Publishing [2003]. Originally shaped by spontaneous initiatives from civil society and the academic community, Open Access and Open Science have also been subject to regulatory initiatives of non-binding nature. One prominent example is the Organisation for Economic Co-operation and Development (OECD) Council Recommendation concerning Access to Research Data from Public Funding [2006]

support the availability of research data by adopting national policies, as well as relevant actions, with the objective of making publicly funded research available: these are defined as 'open access policies'. These policies shall be addressed to research performing organizations and research funding organizations.

Art 10(1) affirms that these policies shall follow the 'open by default' principle. The principle can also be linked to Art 5 of the Directive on available formats, that calls on member States to encourage public sector bodies and public undertakings to produce and make available documents in accordance with the broader principle of 'open by design and by default'. Openness by default can be especially understood in relation to data and the movement for open data, after which the Directive is entitled. For instance, the International Open Data Charter calls on adherent governments and organizations to respect six main principles tantamount to data being open by default (1), timely and comprehensive (2), accessible and usable (3), comparable and interoperable (4), for improved governance and citizens engagement (5) and for inclusive development and innovation (6).[72] More generally, open data can be comprised under the OS and OA movements, but a definition proves elusive since it varies in the literature and open data embodies a multitude of concepts in the data-centric society – being also a buzzword – including the access, use and re-use of data in the digital domain.[73]

According to Art 10(1), policies shall also be compatible with the FAIR principles. While OA and OS address different scientific materials beyond publications, and possibly including research data, the FAIR Data principles – proclaiming that data should be Findable, Accessible, Interoperable and Re-usable – were originally elaborated by the Force1 group between 2014 and 2016[74] and they should be understood as specifically referred to scholarly data.

Art 10(1) also affirms that the policies would take into account the principle of 'as open as possible, as closed as necessary'. The principle should be linked to

C(2006)184. The latter was recently revised in 2021 in the course of the Covid-19 pandemic: Organisation for Economic Co-operation and Development (OECD) Council Recommendation concerning Access to Research Data from Public Funding [2021] OECD/LEGAL/0347.

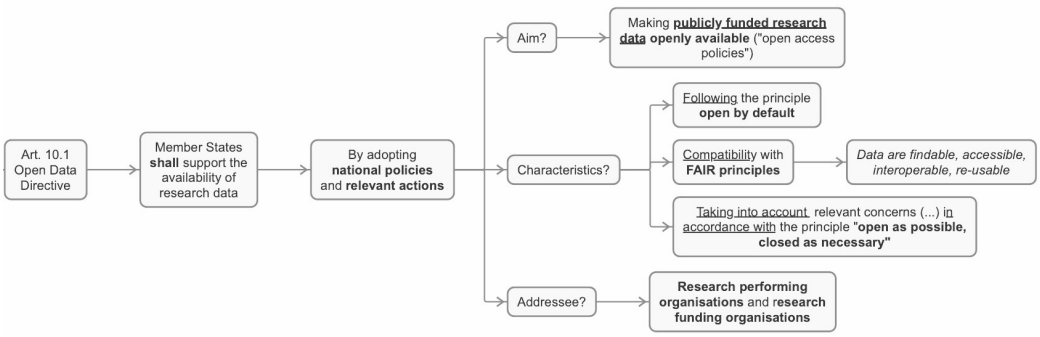
[72] International Open Data Charter [2015] available at https://opendatacharter.net/principles/. The Charter builds on the G8 Open Data Charter of 2013, n 42 above.

[73] The numerous definitions proposed, both in the regulations or by stakeholders, may further specify whether the adjective 'open' refers to a data format, the possibility to use data freely or subject to costs and for certain purposes (ie commercial purposes or not) at certain conditions (eg defined by a licenses), and the types of datasets that are targeted (eg data from the public sector, data shared by private parties, scientific research data, etc). As an additional example, next to the already mentioned Internal Open Data Charter, the Open Knowledge Foundation, a non-profit organization launched in 2004, defines Open data as 'the building block of open knowledge' – knowledge that is free to access, use, modify and share, while preserving provenance and openness. Cultural, science, finance, statistics, weather, environment are mentioned as open data categories. See Open Knowledge Foundation webpage, available at https://tinyurl.com/23ay6s2d (last visited 30 June 2022).

[74] M.D. Wilkinson et al, 'The FAIR Guiding Principles for scientific data management and stewardship' 12 *Scientific Data,* 1 (2016), available at https://tinyurl.com/8sahhsee (last visited 30 June 2022). See Force11 webpage, available at https://tinyurl.com/57yf9nfb (last visited 30 June 2022).

the EU Commission elaborations on open access to research data in the Guidelines for Horizon 2020; in particular, the Open Data Research Pilot acknowledges the possibility to opt out from research data sharing based on some incompatibility grounds.[75] In the text of the Directive, closure namely refers to the protection of rights and interest of others, the protection of personal data and confidentiality, security and legitimate commercial interests, and intellectual property rights.[76]

Figure 1. Graphic representation of Art 10(1) of the Directive (EU) 2019/1024.



For a more precise understanding of the duties and obligations regarding the re-use of research data in the Directive, briefly summarized as follows, the main reference is Art 10(2). This affirms that research data – when publicly funded and already made publicly available, as explained – shall be re-usable for commercial or non-commercial purposes in accordance with chapter III (describing conditions for re-use) and chapter IV (entitled to non-discrimination and fair trading). The article calls for mandatory action to be taken by member States ('research data *shall* be'). The mentioned rules are therefore applicable, notwithstanding the fact that they primarily address obligations directed at public bodies or public undertakings, with the uncertainties previously discussed in para III(1) as to subjects. Relevantly, Art 10(2) adds there should be no prejudice to Art 1(2)(c) (third intellectual property rights) and, as mentioned above, concludes that in this context legitimate commercial interests, knowledge transfer activities and pre-existing intellectual property rights 'shall be taken into account'.

[75] European Commission Guidelines on FAIR Data Management in Horizon 2020, III [2016], 3-4, available at https://tinyurl.com/2p8ay5b8 (last visited 30 June 2022); European Commission H2020 Online Manual, Chapter: Cross-cutting issues - Open access & Data management, available at https://tinyurl.com/3kc3sjcd (last visited 30 June 2022). See also A. Landi et al, 'The "A" of FAIR – As open as possible, as closed as necessary' 2 *Data Intelligence*, 47, 50 (2020).

[76] Directive (EU) 2019/1024, recital 27 introduces the principle 'as open as possible, as closed as necessary' in relation to the issue of rights and interests of others, and it urges that despite the certain obligations established by the Directive for member States towards the opening of publicly funded research, concerns related to the existence of rights on the data, rights of others or different interests, should be taken into account.

For chapter III, this means applying the rules as regarding formats, charging, transparency, licensing, arrangements for the search of documents. According to Art 5, member States shall first encourage the principle of 'open by design and by default' (Art 5(2)), which is one of the most relevant elements of innovation introduced by the Directive. There is also an obligation for public sector bodies and public undertakings that data should be made available in any pre-existing format or language and, where possible and appropriate, by electronic means, in formats that are open,[77] machine-readable, accessible, findable and re-usable (Art 5(1)). This is to the extent to which the creation of documents, adaptation of documents or provision of extracts does not involve disproportionate effort, going beyond a simple operation (Art 5(3)). It bears emphasis that Art 5 affirms the data should be made available together with their metadata. Finally, Art 5(1) adds that both the format and the metadata shall comply with formal open standards,[78] when possible, and namely standards laid down in written form that detail specifications for the requirements on software interoperability (Art 2 point 15) when possible. Nevertheless, regrettably, metadata is not defined in the Directive. More specific rules apply to dynamic data and high-value datasets,[79] but these are not detailed in the present work.

Re-use of documents is in principle free of charge according to Art 6, although the recovery of marginal costs is allowed. Such costs include not only those for the reproduction, provision, and dissemination of documents, but also – which seems crucial considering research data – the ones for anonymization of personal data and for the measures taken to protect commercially confidential information. This rule includes a few exceptions, as for cultural establishments (Art 6(2)), but more importantly Art 6(6)(b) explicitly states that the re-use of research data shall always be free of charge for the user.[80]

Different requirements for the conditions of re-use are detailed in Art 8: there shall be no conditions, unless they are objective, proportionate, non-discriminatory, justified on grounds of a public interest objective, and they shall

[77] Directive (EU) 2019/1024 Art 2 no 14 defines an open format as 1) platform-independent and 2) made available to the public without any restriction that impedes the re-use of documents.

[78] Directive (EU) 2019/1024 Art 2 no 15 defines open format standards as laid down in written form that detail specifications for the requirements on software interoperability.

[79] It should be questioned whether research data may fall under the category of high-value datasets under Directive (EU) 2019/1024 Art 14. This assessment is essentially based on their potential for generate significant socioeconomic or environmental benefits and innovative services, benefit a high number of users, and in particular SMEs, assist in generating revenues, and finally the potential to be combined with other datasets. Thematic categories are detailed in Directive (EU) 2019/1024 Annex I and correspond to 1) Geospatial, 2) Earth observation and environment, 3) Meteorological, 4) Statistics, 5) Companies and company ownership, 6) Mobility. Whether research data would fall under these categories, the principles detailed in Art 14 (namely: availability free of charge with a few exceptions, machine-readability, the provision via API and as bulk download) would apply, plus their re-use would be regulated by specific implementing acts of the Commission.

[80] This excludes the application of Directive (EU) 2019/1024 Art 7, that regards transparency of charging conditions.

not unnecessarily restrict possibilities for re-use. Conditions shall also not be used to restrict competition. The use of standard licenses is also encouraged.

Finally, Art 9 outlines, on the one hand, practical arrangements that Member States shall make to facilitate the search of documents and calls on member States to encourage public sector bodies to make practical arrangement for measures facilitating the preservation of documents made available for re-use. On the other hand, Art 9(2) mentions that the member States shall pursue cooperation efforts with the EU Commission to simplify access to datasets. Such efforts would include in particular the provision of a single point of access and the making available of suitable datasets (for the documents held by public bodies to which the Directive applies, as well as for the data held by the Union institutions) in formats that are accessible, readily findable and re-usable by electronic means.

Chapter IV contains rules on non-discrimination (Art 11) and exclusive agreements (Art 12). Non-discrimination means that applicable conditions for the re-use should not differentiate between comparable categories of re-use, including for cross-border re-use, while establishing a rule that the same charges plus other conditions applying to the re-use by a public sector body for commercial purposes should apply to other users for the supply of those documents for those activities. Exclusive arrangements – ie contracts or related arrangements that grant exclusive rights – are excluded unless an exclusive right is necessary for the provision of a service in the public interest, but these, together with periods of exclusivity exceeding 10 years, are subject to review.[81]

2. Re-Use of Research Data and Intellectual Property

Considering the re-use of research data and limits descending from intellectual property laws, the safeguards provided in Art 1(5) are particularly important. The provision affirms that the obligations imposed in accordance with the Directive shall apply only when compatible with the provisions of international agreements on the protection of intellectual property rights – the Berne Convention, the TRIPS Agreement and the WIPO Copyright Treaty being mentioned. Since the documents in which third parties hold IPR are outside the scope of the Directive, this article suggests that further limitations to the re-use of documents may derive from intellectual property laws nevertheless. It should be remembered, as recital 54 clarifies, that intellectual property rights comprise related rights, including *sui generis* forms of protection. On this point, Art 1(6) states that the *sui generis* right for the maker of a database – provided for in Art

[81] According to Directive (EU) 2019/1024 Art 12 specific rules prescribing transparency and review also applies if there are legal or practical arrangements that, although they not expressly grant an exclusive right, seek or could reasonably be expected to lead to, a restricted availability for the re-use of documents.

7(1) of Directive 96/9/EC[82] – shall not be exercised by public sector bodies so they can prevent the re-use of documents or restrict re-use. Crucially, the final sentence of recital 54 also affirms that public sector bodies should exercise their copyright in a way that facilitates re-use. Above all, it should be remembered that the possibility to apply the *sui generis* right to databases created by public entities is argued in the doctrine.[83]

Art 1 combines with additional limits for the re-use of research data and IPR that emerge in different parts of the text. Besides recital 28 (whose contents were analyzed in para IV(1)), Art 10 recalls concerns of intellectual property rights and, in addition to expressively recalling the IP exemption of 1(2)(c), urges to take into account, *inter alia*, knowledge transfer activities and pre-existing intellectual property rights. The reference seems partially obscure as knowledge transfer is a typical dynamic of licensing IP considering, for instance, Universities' partnerships with private companies or public bodies, while pre-existing intellectual property rights seem to refer to a situation that pre-exists any contractual arrangement. What is more, how such circumstances should ultimately be taken into account is not specified.

Taken together, these provisions considerably restrict the extent to which scientific research data can be subject to re-use. In doing so, the complexities characterizing the context of IPR and research data are scarcely addressed,[84] despite the topic being acknowledged as a challenge in the preparatory works, and the fragmentation of policies and inconsistency of related sharing practices for research data (deeply affected by IPR and especially copyright) were pointed out as one reason for promoting legal change with the Open Data Directive.

As anticipated in para III(1), one main underlying issue regards the idea/ expression dichotomy. The definition of research data in the Directive regards documents other than scientific publications that are collected, produced, and used across different phases of scientific research, as well as accepted in the scientific community. While publications – ultimate target of copyright – are excluded, the definition includes documents in a digital form and this is a broad formula that points to a variety of materials potentially protected by copyright. This would include different media, including images (possibly also 3D digital models), videos or other types of texts that cannot be framed as scientific

[82] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L 77/20.

[83] Considering Italy, F. Faini, n 22 above, 123-124. For a thorough analysis whether public entities could be the subjects of database sui generis rights, including the case decided by the Court of Justice of the European Union Case -138/11, *Compass-Datenbank GmbH* v *Republik Oesterreich,* Judgment of 12 July 2012, available at at www.eur-lex.europa.eu, P. Guarda, *Il regime giuridico dei dati della ricerca* (Trento: Università degli Studi di Trento, 2020), 124-125.

[84] J.H. Reichman and R. Okediji, 'When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale' 96 *Minnesota Law Review*, 1362 (2012). More recently, in relation to the pandemic context, K. Walsh et al, 'Intellectual Property Rights and Access in Crisis' 52 *International Review of Intellectual Property and Competition Law*, 379 (2021).

publications. Specific attention should be attributed to code, eg considering computer programs or algorithms, whose copyrightability, together with patentability, is discussed. Indeed, despite recital 30 mentioning that the definition of document is not intended to cover computer programs, member States remain free to extend the application to them. Considering, more to the point, datasets, while in line with the idea/expression dichotomy principle their content should not be protected by copyright, they may still be protected if, by reason of the selection or arrangement of their contents, they are original (Art 1(2) of the Directive 96/9/EC). Even more importantly, *sui generis* rights can protect datasets in presence of investment (Art 7(1) Directive 96/9/EC).

A second underlying issue is that IPR in research are often characterized by shared, fragmented, and sometimes uncertain, authorship; this descends from the essentially cumulative nature of scientific knowledge and the free circulation of ideas, as well as the resort to contractual agreements for IPR management, eg in knowledge transfer. As a consequence, the limits imposed by the described IP safeguards in the Open Data Directive – and consequent activities required for compliance, such as rights clearance – seem rather severe, for the obligations for re-use on research data could be even more difficult to attribute. For instance, it could be difficult to establish whether and how Art 1(6) of the Directive – that encourages not exercising the *sui generis* rights to prevent or restrict re-use – would be applicable in the context of research data. As noted by distinctive authors, the proposal for a Data Act provides for an identical rule in Art 5(7):[85] although the proposal was eagerly awaited to amend the subject of *sui generis* rights on databases, in its current version it does not introduce other relevant provisions on this utterly controversial set of rights.

3. Re-Use of Research Data and Personal Data Protection

Safeguards for the respect of personal data protection laws are found in Art 1(4) of the Directive. This states that the Directive is without prejudice to Union and national law on the protection of personal data, in particular the Regulation (EU) 2016/679 (GDPR),[86] the ePrivacy Directive[87] and corresponding national law. Recital 154 of the GDPR mirrors this provision, as it affirms that the EU legislation on the re-use of public sector information does not affect the EU data protection provisions. Overall, this means that, given that some documents containing personal data would be excluded by the scope of application of the Directive *a priori*, in light of Art 1(2)(h), the Directive may still apply to documents

[85] P. Keller, 'A vanishing right? The Sui Generis Database Right and the proposed Data Act' *Kluwer Copyright Blog*, 4 March 2022, available at https://tinyurl.com/2p8dkab6 (last visited 30 June 2022).

[86] Regulation (EU) 2016/679, n 9 above.

[87] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L201/37.

that contain personal data and, whenever this is the case, access and re-use of the documents should comply with data protection principles and rules.

A necessary premise is that the subject of Open Data and Data Protection can be considered to suffer a contrast at the conceptual level. Put more bluntly, it is difficult to see how opening to non-discriminatory re-use of data for any purpose (ie commercial and non-commercial) could be compatible with the principles of purpose limitation, data minimization, accuracy and possibly accountability, principles now enshrined in Art 5 of the GDPR.[88] Useful information about the interplay of PSI and Data Protection rules was set out by the Art 29 Working Party (hereinafter WP29, now European Data Protection Board, also EDPB), in 2003[89] and 2013.[90] During the preparation of the EU Commission Guidelines on the amended Directive of 2013 and the related consultation, the European Data Protection Supervisor (hereinafter, EDPS) also strengthened the WP29 considerations on PSI rules and data protection.[91]

As for the considerations advanced by the WP29, this first addressed the idea that because the re-use is a 'non-obligation' in the PSI Directive, related public bodies may decide to make the data available or not; it also underlines how such a decision is impacted by personal data, as data protection principles and rules should be subject to a dedicated assessment.[92] The option of making available data after anonymization is a crucial one according to WP29,[93] but it recalls that this comes with the critical need to assess and test risks of re-identification.[94] It is indeed a well-worn argument that the advance of technology, ie cryptography, has increasingly rendered complete anonymization impossible.[95]

[88] This issue has been described providing a fresh perspective on the Open Data Directive and the GDPR in the recent work of P. Guarda, n 83 above, 206; on Directive 2013/37/EU and the proposed GDPR M. Van Eechoud, n 5 above, 75-76. See also R. Ducato, 'Data Protection, Scientific Research, and the Role of Information' 37 *Computer Law & Security Review*, 36 (2020); F. Zuiderveen Borgesius et al, 'Open Data, Privacy, and Fair Information Principles: Towards a Balancing Framework' 30(3) *Berkeley Technology Law Journal*, 2073 (2015); I. Graef et al, *Spill-Overs in Data Governance: The Relationship between the GDPR's Right to Data Portability and EU Sector-Specific Data Access Regimes*, TILEC Discussion Paper DP 2019-005 (2021), available at https://tinyurl.com/23r7vrc7 (last visited 30 Jume 2022).

[89] Art 29 Working Party, Opinion no 7/2003 on the data protection concerns relating to PSI (2003). The objective of the Opinion was to providing guidance and examples on how to implement the amended PSI Directive with regard to the processing of personal data.

[90] Art 29 Working Party, Opinion no 6/2013 n 67 above.

[91] European Data Protection Supervisor, Comments in response to the public consultation on the planned guidelines on recommended standard licences, datasets and charging for the reuse of public sector information initiated by the European Commission [2013], available at https://tinyurl.com/zj3racrn (last visited 30 June 2022).

[92] Art 29 Working Party, Opinion no 6/2013 n 67 above, 3.

[93] ibid 3, 12.

[94] ibid 7.

[95] Art 29 Working Party, Opinion no 5/2014 on Anonymization Techniques (2014), 7-8; R. Ducato, 'La Crisi Della Definizione Di Dato Personale Nell'era Del Web 3.0. Una Riflessione Civilistica in Chiave Comparata' in M. Tomasi and F. Cortese eds, *Il Diritto e le definizioni* (Napoli: Editoriale Scientifica Italiana, 2016), available at https://tinyurl.com/rhfwe6hu (last visited 30 June 2022); S.

This is a central topic considering, for instance, that aggregated statistical data are presented as a typical example of PSI.

The WP29 mentioned that, when making data available under the PSI rules, public sector bodies will need a legal basis to make the personal data available for re-use (ie disclosure),[96] although in presence of a non-obligation to disclose, they would probably not be able to invoke the need to comply with the PSI Directive as a legal basis.[97] Under the GDPR, next to the necessity of the processing for compliance of a legal obligation (Art 6(1)(c) GDPR), another legal basis on which the public sector body may rely would be the consent of the data subject (Art 6(1)(a) GDPR) or necessity for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller (Art 6(1)(e) GDPR). Both the former and the latter would nevertheless require the legal basis to be laid down in Union or national law (Art 6(3) GDPR) and more specifically, for the performance of a task or exercise of authority, the purpose of the processing should be determined by the law or be necessary (Art 6(3) GDPR).

Another major issue is that the so-called disclosure likely qualifies as a further processing of the data, for purposes that are different from the ones for which the data was collected: this is one primary example of the tension between the guiding principle of open data and the data protection principle of purpose limitation,[98] which requires that the purposes of the further processing should be compatible with the purposes for which the data has been initially collected.[99] Conditions for further processing and assessment thereof are now included in Art 6(4) of the GDPR.[100] On this point, the WP29 strongly recommended the adoption of detailed national provisions that would specify the purposes for which public sector bodies would be able to disclose data, but also invited the public sector bodies to conduct a dedicated assessment.[101]

Stalla-Bourdillon and A. Knight, 'Anonymous Data v. Personal Data - A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data' 34 (2) *Wisconsin International Law Journal*, 284 (2017).

[96] Art 29 Working Party, Opinion no 6/2013, n 67 above, 6-7.

[97] ibid

[98] P. Guarda, n 83 above, 206-207.

[99] Art 29 Working Party, Opinion no 6/2013, n 67 above, 6.

[100] Regulation (EU) 2016/679, Art 6(4) recites: 'Where the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent or on a Union or Member State law which constitutes a necessary and proportionate measure in a democratic society to safeguard the objectives referred to in Art 23(1), the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, *inter alia* (…).'

[101] At the time, a Data Protection Impact Assessment was only recommended in the Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31, while it is today prescribed as mandatory in the Regulation (EU) 2016/679, Art 35. See Art 29 Working Party, Opinion no 6/2013, n 67 above, 6, 20.

Finally, the re-use of personal data by the users would also need a legal basis. The most appropriate legal basis for re-use is eventually identified by the WP29 in consent of the data subject or legal obligation.[102] Such processing would also need to comply with the principle of purpose limitation, although the WP29 specified that, when considering the compatibility of further use, the distinction between re-use for commercial or non-commercial purposes should not be decisive.[103] In particular, the WP29 underlined that even though the data would be available on the Internet, this would not mean that personal data could be processed for any purpose. As public sector bodies would be able to impose conditions for re-use, subject to a few requirements such as objectivity and non-discrimination between users, such conditions could limit the purposes of the re-use of personal data. Since the re-use could be difficult to monitor, however, this is another element that should fall into the dedicated data protection assessment.[104] For all these reasons, the WP29 supports the view that public bodies should put in place a rigorous licensing scheme that would specify purposes for which re-use is allowed[105] and foresee a data protection clause in their conditions, even when data is anonymized.[106]

More recently, the topic was tackled by the European Data Protection Board and European Data Protection Supervisor Joint Opinion 03/2021 on the Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act).[107] The document examines the relationship of the proposal for the Data Governance Act with the Open Data Directive and the GDPR. On this occasion, while critically examining the fact that data held by public bodies and protected on grounds of, *inter alia*, protection of personal data was included in the scope of the new proposed Regulation, the Opinion confirmed that the rules of the Open Data Directive appear consistent with the requirements governing protection of individuals' fundamental rights.[108]

For the purposes of the present work, there should be an investigation into how the elements hereby described would affect the context of re-use of research data according to Art 10 of the Open Data Directive. Numerous tensions characterizing data protection and public sector information are already mentioned in the WP29 Opinion of 2013[109] and indeed, the described data protection issues persist and continue to appear complex, compliance being

[102] ibid 19; the reference is to Directive 95/46/EC Art 7(a)-(f).

[103] ibid 21.

[104] ibid 20.

[105] ibid 19.

[106] ibid 25.

[107] European Data Protection Board and European Data Protection Supervisor (EDPB-EDPBS) Joint Opinion 03/2021 on the Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act) [2021] available at https://tinyurl.com/mtnvbmh9 (last visited 30 June 2022).

[108] ibid 18-20.

[109] Art 29 Working Party, Opinion no 6/2013, n 67 above, 23.

even more onerous, in the context of research data, as research activities frequently resort to personal data, involving a plurality of players acting in different capacities,[110] including public-private partnerships.

If research data contains personal data, the operations that are functional to allowing the re-use of this research data (ie the disclosure) would be tantamount to data processing activities that require an apt legal basis in Art 6 of the GDPR or equivalent in national laws. The same is true with regard to the re-use of research data by users, although limited purposes for the re-use of research data could be specified in the terms and conditions. It therefore seems helpful to consider the Data protection rules presenting a few specificities when personal data processing is for purposes of research, where research is defined under recitals from 157 and following of the GDPR. However, it should be acknowledged that the application of such provisions relies on the purposes of the processing, so they would impact data processing activities during the actual research phases. One first question is consequently whether the disclosure or even the re-use (eg when the conditions for re-use prescribe that data are re-usable for research purposes only) could be considered as falling under the research purposes.

As for the legal basis of personal data processing for purposes of scientific research in the GDPR, three of them are referred in the doctrine as the most relevant: the consent of the data subject (Art 6(1)(a) GDPR), the necessity of processing for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller (Art 6(1)(e) GDPR) and the necessity of the processing for the purposes of the legitimate interests pursued by the controller or by a third party (Art 6(1)(f) GDPR).[111] Both letter e) and f) would require the basis to be laid down in Union or national law (Art 6(3) GDPR). However, it can surely happen that personal data protection processed for purposes of research falls under the special categories of data (Art 9 of the GDPR), a primary example being medical or biological research, for, amongst others, data concerning health[112] and genetic data. Art 9(2)(j) of the GDPR

[110] F. Di Tano, 'Protezione dei dati personali e ricerca scientifica: un rapporto controverso ma necessario' 1 *BioLaw Journal – Rivista giuridica di Biodiritto*, 71, 80-81, (2022).

[111] P. Guarda, n 83 above, 145-149. Relevantly, considering the PSI rules, for public sector bodies only the first two legal basis mentioned would be applicable, due to GDPR Art 6.1(f) excludes that the legitimate interest basis shall apply to processing carried out by public authorities in the performance of their tasks.

[112] One relevant example could be disclosure of research data collected by public sector bodies during the pandemic of Covid-SARS-19; if not correctly anonymized, research data to be disclosed and possibly re-used may comprehend datasets that amount to special categories of data under the GDPR, ie data concerning health; on this point cf. E. Sorrentino and A.F. Spagnuolo, 'Dati sanitari: aperti, accessibili e riutilizzabili' *MediaLaws.eu*, 16 December 2021, available at https://tinyurl.com/59wdppau (last visited 30 June 2022); T. Fia, 'Access to and Ownership of Data to Tackle COVID-19: Some Lessons (IP) Law Should Learn for Good', (2020), available at https://tinyurl.com/48vd9s96 (last visited 30June 2022). The topic is politically charged due to the greater controversiality of both public and private control of information during the pandemic (eg

would apply in this case. This provision prescribes that the processing would be allowed where necessary for the purposes of Art 89(1) of the GDPR (processing of personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes), based in Union or national law, proportionate to the aim pursued, when it would respect the essence of data protection right and when appropriate and when specific measures are in place. For the sake of completeness, it should ultimately be remembered that processing of personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes falls within Art 89 of the GDPR, which ties such processing to a few safeguards[113] and derogations.[114] Essentially, the further processing of data for the purposes mentioned will not be deemed incompatible with the original purposes for which data was collected, at least when such processing happens in accordance with Art 89 of the GDPR.

To conclude, despite access and re-use of research data under the Open Data Directive bringing consistent data protection challenges, the possibility to refer to extensively harmonized data protection rules across member States, embedded in the GDPR, may ensure greater legal certainty in the implementation and application of these rules. In this respect, the scenario seems different from what has been described in relation to the limits concerning the intellectual property subject in para IV(2). Moreover, key elements to navigate the described

number of infections, deaths, vaccines and Covid-SARS-19 variants), especially considering Intellectual property laws.

[113] The safeguards provided by Art 89 GDPR are aimed at protecting the rights and freedom and of the data subject and they primarily consist in technical and organizational measures, particularly to ensure data minimization (eg pseudonymization). The prescription of such safeguards suggests very strong care should be adopted to decide whether research data containing personal data (although pseudonymized) should be made available and should be open for re-use.

[114] Derogations, instead, regard the exercise of a few data protection rights. More specifically: access (Art 15 GDPR), rectification (Art 16 GDPR), restriction of processing (Art 18 GDPR), notification (Art 19 GDPR), portability (Art 20 GDPR), objection (Art 21 GDPR). Derogations should also be established by Union or national law, be necessary to fulfil the aim pursued and be provided only when the rights would seriously in impair the aimed purposes. This however means that the public sector body that engages in research would be still be accountable for data subjects and ensure to respect their right to receive correct information (Arts 13-14 GDPR) and, in the few prescribed cases (eg revocation of consent, absent another legal ground for processing), the right to erasure (Art 17 GDPR), despite the right is additionally limited when processing is for the purposes of Art 89 of the GDPR. Art 17(3) (d) of the GDPR specifies the right to erasure would not apply when the processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Art 89(1) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing. This implies, on one hand, that the public sector body should provide information that data will be, even partially, disclosed, plus on potential re-use. On the other hand, it would also mean that in case of erasure of personal data, whenever data have been made public by the public sector body (as with public disclosure), Art 17(2) GDPR would also apply. Consequently, the public sector body, as the controller, would be obliged to take reasonable steps, including technical measures, to inform other controllers processing the personal data that the data subject has requested they also erase any links to, or copy or replication of, the personal data.

context are first the need to occasionally look at the national provisions for compliance of personal data processing for research purposes (eg considering the legal basis), and the fact that relevant uncertainties are likely to arise in the concrete re-use of research data, requiring a case-by-case assessment, as concluded by both the EDPB, EDPS, as well as the doctrine.[115]

V. Data from Cultural Establishments in the Directive (EU) 2019/1024: A Brief Overview

Although data from cultural establishments are not the focus of the present article, careful consideration of the applicable rules in the Open Data Directive is complementary to analysis sketched so far. This is mainly because research data and cultural data can be considered equally fundamental to the umbrella concept of open knowledge and growing attention, in time, to 'open cultural data' or what can be loosely defined as 'open access' in the cultural sector,[116] well exemplified in the OpenGLAM initiative born around 2010,[117] together with many others.

The PSI Directive of 2003 did not apply to data from cultural establishments and public broadcasting organizations.[118] The exclusion from the scope of the Directive, as reported in the first proposal of 2002,[119] was based on the idea that the administrative burden would exceed the advantages, the presence of materials characterized by third-party copyright, as well as the special position of such establishments in the society, due to their cultural and knowledge mission.[120] The exclusion was debated following the first publication of the Directive and became of major momentum when the reform of 2013 was discussed.[121] Respondents to the public consultation opted for the inclusion.[122]

Building on studies conducted in the meantime, the Staff Working Document of 2011 concluded in favor of the opportunity to extend the scope of the PSI Directive, as the scenario for the digital exploitation of digital cultural assets had profoundly changed.[123] In particular, what was explicitly acknowledged was the need to amend the PSI rules in order to overcome the differences in rules and

[115] P. Guarda, n 83 above, 209.

[116] M. Terras, 'Opening Access to Collections: The Making and Using of Open Digitised Cultural Content' 39(5) *Online Information Review*, G. E. Gorman and J. Rowley (eds) special Issue 'Open Access: Redrawing the Landscape of Scholarly Communication', 733, 735-736, 742-743, (2015).

[117] See OpenGLAM website, available at https://openglam.org/.

[118] Directive 2003/98/EC, Art 1(2)(f).

[119] European Commission Proposal for a Directive of the European Parliament and of the Council on the re-use and commercial exploitation of public sector documents [2002] COM (2002) 207 final.

[120] K. Janssen, n 15 above, 448.

[121] SEC(2011) 1152 final, n 26 above, 34-38.

[122] ibid 67-68.

[123] ibid 36-37.

practices across the member States relating to the exploitation of public cultural resources – differences that were barriers to realizing the economic potential of those resources in the Internal market.[124] Projects of digitization and availability of digital public domain were pointed out to hide great potential for developing products and services in the field of, amongst others, e-learning and tourism.[125] In doing so, the novel PSI Directive of 2013 was also recognized to reinforce the EU digitization policy for the cultural sector.[126]

At the same time, the document of 2011 acknowledged that ad hoc provisions had to be included due to the specificities of this sector –

> 'administrative complexities linked to IPR protection and the mission of public cultural institutions, which not only disseminate but also preserve the cultural heritage they hold'.[127]

One first principle consists in the fact that only public domain material with IPR clear status should be covered by the re-use, to avoid the administrative burden that would derive from right clearance activities. Second, cultural institutions should be able to recover their costs with a reasonable return on investment, to generate funds for making their collections available for re-use, as these are often insufficient.[128] As a result, the reform of 2013 extended the scope to the documents held by libraries, including university libraries, museums, and archives, while excluding other cultural establishments. This was in view of a performing arts specificity – the Directive currently cites orchestras, operas, ballets and theatres – and because almost all of the material detained by such establishments was reputed covered by third-party intellectual property rights.[129]

With regard to intellectual property rights, it is worthwhile noting that the general exclusion to documents in which third parties hold intellectual property rights would also apply.[130] On this point, authors argued about including in the scope of the Directive documents that were initially owned by third parties and that were only later acquired by cultural institutions, and thus questioned the reading of ex recital 9 of the PSI Directive of 2013, now recital 55 of the Open Data Directive (already mentioned in para IV(2)).[131] However, it was established

[124] Directive 2013/37/EU, recital 17.

[125] Directive 2013/37/EU, recital 15.

[126] SEC(2011) 1152 final, n 26 above, 27-28.

[127] ibid 37.

[128] ibid 37.

[129] Directive (EU) 2019/1024, Art 1(2)(j) PSI 2019, recital 65; Directive 2013/37/EU, Art 1(2)(f), recital 18.

[130] ibid Art 1(2)(c); Directive 2013/37/EU, Art 1(2)(b).

[131] P. Kelle et al, n 17 above, 4. On copyright and museums as subject to PSI rules see C. Sappa, 'Museums as Education Facilitators. How copyright affects access and dissemination of cultural heritage', in E. Bonadio and C. Sappa eds, *Art and Literature in Copyright Law: Protecting the Rights of Creators and Managers of Artistic and Literary Works* (Cheltenham: Edward Elgar Publishing, forthcoming).

that for documents in which cultural establishments hold intellectual property rights, the cultural institution could decide whether to allow re-use or not; member States shall ensure that these documents shall be re-usable for commercial or non-commercial purposes in accordance with the conditions set out in the Directive, where the re-use of such documents is allowed.[132]

Other ad hoc rules have been established for the relevance of strategic partnerships and the costs of digitization projects. Despite the general prohibition, cultural establishments are allowed to charge above marginal costs for the re-use; while not exceeding the cost of collection, production, reproduction, dissemination, preservation and rights clearance, a reasonable return on investment is possible.[133] In the new Directive, the possibility of charging is maintained for libraries, museums, and archives, and it would apply also in the case of high-value datasets.[134] What reasonable return on investment means has been further explained in the Guidelines of the EU Commission of 2014.[135] This would include a return rate, to be calculated not in reference to business risk, but being 'reasonable' instead, and placed slightly above the current cost of capital (ie considering the European Central Bank's fixed interest rate when in the euro-zone), while well below the rate for commercial players.[136] With regards to these conditions, a few scholars have argued for cautious interpretation and careful implementation of such a rule already under the previous Directive, since imposing conditions for re-use may alter the inner balance of copyright law, where there are examples of public domain works previously made available by cultural institutions without restrictions.[137] Next to ad hoc rules for charging, exclusive arrangements for digitization of cultural resources have been permitted, although subject to specific rules, as for the review of the exclusive rights duration or the provision of a copy of the digitized cultural resources.[138]

Time has passed, but regrettably the new Open data Directive still covers only certain types of cultural establishments. The relevant exemptions and limitations regarding intellectual property rights have also not changed, and it remains true that cultural establishments are subject to significant derogations. Amongst those, one that deserves particular attention in the Open Data Directive is on exclusive agreements. Contracts or other arrangements that would grant exclusive rights between libraries, museums, archives, and private partners concerning the digitization of cultural resources are allowed in order to give the

[132] Directive (EU) 2019/1024, Art 32; Directive 2013/37/EU, Art 3(2).

[133] Directive (EU) 2019/1024, Arts 6(2) and 6(4), recital 38; Directive 2013/37/EU, Arts 6(2) and 6(4).

[134] Directive (EU) 2019/1024, Art 14(4).

[135] European Commission Notice - Guidelines on recommended standard licences, datasets and charging for the reuse of documents [2014] 2014/C 240/01.

[136] P. Keller et al, n 17 above, 4.

[137] ibid 5-6.

[138] Directive (EU) 2019/1024, Art 12(2); Directive 2013/37/EU, Art 11.

private partner the possibility to recoup its investment (recital 49 and Art 12(2), second sub-para).[139] Nevertheless, it is far from obvious to assert what exclusive rights these provisions would refer to. The rights seem to be generally framed as rights to re-use the resources (eg in recital 48), but for the context of digitization projects, as also noted by other authors,[140] they seem to consist in the right to digitalize the resources, as it is in Art 12(3) and recital 49. Moreover, the same recital 49 may also be read as referring to IPR when it recites that the period of exclusivity should be as short as possible 'to comply with the principle that public domain material should stay in the public domain once it is digitised'.

While the new Open Data Directive does not meaningfully innovate the provisions on data from cultural establishments compared to the previous PSI Directive of 2013, its contents are remarkably complemented by the recent Commission Recommendation of 10 November 2021, on a common European data space for cultural heritage.[141] Following the previous Recommendation on the digitization and online accessibility of cultural material and digital preservation of 2011[142] and its evaluation in 2021,[143] as well as taking into account Covid-19 as a drive for digitization for cultural heritage institutions, the new Recommendation brings the cultural sector to the fore of the European Strategy for Data.[144] Relevantly, in provision no 18 the Recommendation affirms that the policies adopted by member States should seek to ensure that data resulting from publicly funded digitization projects become and stay FAIR. The result is that, despite having non-binding nature, the Recommendation provides persuasive elements that would deserve to be taken into account in both the implementation and application of the PSI rules. At the same time, the

[139] According to Directive (EU) 2019/1024, Art 12(3), first and second sub-paragraphs). Such agreements shall be transparent and public, and although the period should in principle not exceed 10 years, in case this happens the duration shall be reviewed during the 11th year and, if applicable, every 7 years after that. Since 'any public private partnership for the digitisation of cultural resources should grant the partner cultural institution full rights with respect to the post-termination use of digitised cultural resources' (recital 49 Directive (EU) 2019/1024) a copy of the digitized cultural resources shall be made available as the at the end of the exclusivity period (Art 12(3), third sub-paragraph, Directive (EU) 2019/1024).

[140] A. Wallace and E. Euler, 'Revisiting Access to Cultural Heritage in the Public Domain: EU and International Developments' 51(7) *IIC - International Review of Intellectual Property and Competition Law*, 823, 844 (2020).

[141] European Commission Recommendation of 10 November 2021 on a common European data space for cultural heritage [2021] C(2021) 7953 final.

[142] European Commission Recommendation of 27 October 2011 on the digitisation and online accessibility of cultural material and digital preservation [2011] OJ L283/39.

[143] European Commission Staff Working Document Evaluation Of the Commission Recommendation of 27 October 2011 on the digitisation and online accessibility of cultural material and digital preservation [2011] SWD(2021)15 final.

[144] In particular, SWD(2021)15 final, General Provisions no 10 recites: 'Where cultural heritage institutions enter into partnerships with the private sector, they should ensure that clear and fair conditions for reusing the digitised assets are laid down, in line with competition rules and with Directive (EU) 2019/1024, and in particular with the rules on exclusive arrangements laid down in Article 12 of that Directive, where relevant.'

PSI rules are confirmed to provide a substantial base of harmonization for realizing the EU Data Strategy in the field of cultural heritage.

VI. The Italian Implementation of the Directive (EU) 2019/1024

Moving on to the implementation of the new PSI rules in Italy, para 2 describes the provisions recently introduced by the decreto legislativo 8 November 2021 no 200, focusing on research data and providing a few insights into data from cultural establishments. Para 1 initially provides an introductory overview on the Italian regulatory framework on PSI.

1. Public Sector Information in Italy

The Italian rules on access and re-use of public sector information can be loosely described as being scattered across three main pieces of legislation.[145] Amongst those, the primary reference for the purposes of the present work is decreto legislativo 24 January 2006 no 36. This has transposed the Directive of 2003 and has been successively modified in accordance with the development of the EU PSI Directives.

Second, the decreto legislativo 7 March 2005 no 82, also known as Codice dell'amministrazione digitale (literally: code of digital administration), hereinafter CAD, should be considered, being the most important piece of legislation for the transition towards e-government.[146] Amongst others, a few provisions also target obligations of public entities for the access and re-use of data.[147] In particular, the CAD provides the main definitions of open data (more precisely, 'open-type' data), open format, and data ownership (*titolarità*),[148] as well as rules on licensing. Most notably, the principle currently enshrined in Art 52 is that in the absence of a general standard license, the documents and data that are published should be considered open data, according to the above-mentioned definitions of open format and open-type data, where the latter also implies that they can be re-used for commercial purposes.[149] This piece of legislation has included a provision on open data since 2012, when it was modified in accordance

[145] G. Luchena and S. Cavaliere, 'Il riutilizzo dei dati pubblici come risorsa economica: problemi e prospettive' *Rivista giuridica del Mezzogiorno,* 151, 160-166 (2020).

[146] F. Faini, n 22 above, 25. The CAD provides the key-provisions for the digitalization of information of the public sector, primarily considering the relationship with users and tools of 'digital citizenship', for instance digital identity, but also, more in general, rules for digital documents, signatures, transmission.

[147] Art 50 and following CAD.

[148] Art 1(1), (l-*ter*), (l-*bis*), and (cc) CAD. For further details on definitions provided in the CAD, see no 168 below.

[149] More precisely, under Art 1(1) (l-*ter*) of the CAD, data of open typology (*dati di tipo aperto*) are also available under the terms of a licence or regulatory provision that permits the use by anybody, also for commercial purposes, in a disaggregated format.

with the legge 6 November 2021 no 190, a delegation law that would have later converged in the other fundamental piece of legislation to be considered by the present overview, the decreto legislativo 14 March 2013 no 33, the so-called Decreto transparenza (literally: transparency decree). Afterwards, provisions on the re-use of data in the CAD were further amended in time, including by the legge 7 August 2015 no 124 – the so-called Legge Madia – that reshaped the digital administration. Conclusively, the link between the CAD and decreto legislativo 36/2006 is still particularly important today, and primarily regards the definitions of open data, open format, and others.[150]

Finally, the principle of transparency was already embedded as a principle in the legge 7 August 1990 no 241, detailing the rules on the administrative procedure and access to documents, but such national rules on administrative transparency have profoundly evolved in time[151] and now they are ultimately collected in the already mentioned Decreto trasparenza. This comprises the core rules for access to documents by citizens to protect their rights, promote participation, and favor distributed forms of control on the public. In particular, as a result of different reforms in time and more precisely after the decreto legislativo 25 May 2016 no 97 – possibly to be regarded as the Freedom of information Act of Italy [152] – Art 5(2) of the decreto legislativo no 33/2013 now provides further possibilities to access documents thanks to *accesso civico generalizzato*.[153] Aspects of the quality of the information, such as integrity and completeness, are mentioned in Art 6(2), while the re-use of data is targeted by Art 7 and 7-*bis*. In particular, Art 7 affirms that 'documents, information and data' that are subject to mandatory publication, made available also as consequence of the civic access, are published in open formats[154] and re-usable in accordance with, *inter alia*, the decreto legislativo no 36/2006. Art 7-*bis* contains a few limits concerning personal data protection.

Overall, it should be kept in mind that when discussing the national regulatory framework on PSI and open data, provisions of the decreto legislativo no 36/2006, the CAD and the decreto legislativo no 33/2013 overlap; this is in line with the parallel development of initiatives regarding public sector information and freedom of information and emerging trends on open data, also open

[150] For further details on definitions provided in the CAD, see n 168 below.

[151] R. Sanna, n 4 above, 37, 243.

[152] F. Faini, n 22 above, 87.

[153] Accesso civico generalizzato, provided by Art 5(2) d.lgs. 33/2013, is next to a simple civic access that regards documents subject to mandatory publication provided by Art 5(1) d.lgs. 33/2013. *Accesso civico generalizzato* covers documents which are not mandatorily published by public bodies, absent legitimization and motivation, and it is denied only in case of concrete prejudice to the protection of interests of public and private nature disposed by law and under circumstances detailed by Art 5bis of the d.lgs. 33/2013. F. Faini, n 22 above, 109-111; V. Pagnanelli, 'Access, Accessibility, Open Data. The Italian Model of Public Open Data in the European Context' *Giornale di Storia Costituzionale*, 205, 213 (2016).

[154] The definition of open format is provided by Art 1(1)(l-*bis*) of CAD; see note no 168.

government data, as described in para II. However, because the focus of the present work is the re-use of research data, the following analysis will focus on the related amendments to decreto legislativo no 36/2006 only.

2. Rules on Research Data and Data from Cultural Establishments Introduced by the Decreto Legislativo no 200/2021

A few days after the expiration of the implementation term for Directive (EU) 2019/1024, prescribed for 17 July 2021,[155] a draft of *schema legislativo* to implement the Directive was preliminary approved on 5 August 2021, in the meeting of *Consiglio dei Ministri* no 32, and subject to the approval of the Italian Parliament.[156] The decreto legislativo no 36/2006 has consequently been modified by the decreto legislativo no 200/2021, with amendments entered into force on 15 December 2021[157].

Art 1(2-*bis*) of the decreto legislativo no 36/2006 establishes that the rules of the decreto apply to research data under conditions described in Art 9-*bis*.[158] Importantly, this introduces in the legislative corpus the first binding rules to apply to the re-use of publicly funded research data, in the absence of other relevant national provisions in Italy. On this point, it should be mentioned that in the recent past the legge 7 October 2013 no 112 was enacted to implement the non-binding EU Commission Recommendation on access to scientific publications of 2012, promoting member States' actions as regard publicly funded research.[159] On the one hand, Art 4(1) of legge no 112/2013 has introduced an obligation for public entities to adopt, in their autonomy, measures to promote open access to the 'results' of publicly funded research when they are documented in articles published in scientific journals with at least two issues per year, and taking into account both the so-called Green and Golden OA opportunities.[160] On the other

[155] Directive (EU) 2019/1024 Art 17.

[156] Atto del Governo no 284, Schema di decreto legislativo recante attuazione della direttiva (UE) 2019/1024 relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico, documents available at https://tinyurl.com/mujzhbpm (last visited 30 June 2022). Such an approval was prescribed by Art 1 legge 22 April 2021 no 53, so-called European delegation Law 2019-2020.

[157] A first analysis of the amended decreto legislativo no 36/2006 is found in G. Cassano and M. Iaselli, 'Il riutilizzo dei dati pubblici: l'approccio del d.lgs. n. 200/2021' *Diritto di Internet*, 49 (2022).

[158] As a preliminary remark, the scope of application of the decreto legislativo no 36/2006 is defined in Art 1(1) as limited to documents which contain public data (*dati pubblici*) that are in the availability of public administration, bodies governed by public law and public and private enterprises (as further detailed by Art(2-*ter*) and (2-*quater*)). It should be remembered that the definition of public data (*dati pubblici*) (Art 2(d) of the decreto describes these are data which can be known by anyone) was instead removed in the CAD in 2016 (see Art 1(1)(n) CAD, now suppressed by decreto legislativo 26 agosto 2016 no 179). Exclusions follow in Art 3 of the decreto legislativo no 36/2006, while Art 4 provides for safeguards in respect to the compliance with relevant laws (including, *inter alia*, national data protection law, copyright law, industrial property law).

[159] R. Caso, 'La legge italiana sull'accesso aperto agli articoli scientifici: una prima panoramica' *Aedon,* (2013), available at https://tinyurl.com/2p98mf39 (last visited 30 June 2022).

[160] On further discussion on the legislative mandates for open access, and for particular reference

hand, Art 4(3) of legge no 112/2013 has prescribed that to optimize available resources and facilitate the retrieval and use of 'cultural and scientific' information, the Ministry of Cultural Heritage and Activities and Tourism and the Ministry of Education, Universities and Research would coordinate strategies for unifying the databases they manage. However, this law has resulted in the application of different practices across public bodies. Therefore, while fresh actions to enhance open science and open access are currently expected according to the national program for research (2021-2027), approved in 2021 but not yet implemented,[161] decreto legislativo no 200/2021 should be welcomed as having introduced groundbreaking elements in this backdrop.

The definition of research data now found under Art 2(1)(c-*septies*) of decreto legislativo no 36/2006 mirrors the one given in Art 2 of the Directive. Also, Art 3(1)(h-*sexties*) reiterates that the Directive would not apply to documents held by research institutions and organizations that fund research, including the research institutions that are engaged in the research results transfer, whenever different from documents that amount to research data.

Art 9-*bis* establishes specific rules for re-use of research data. Its para 1 first affirms that research data is re-usable for commercial and non-commercial purposes according to what is provided by the decreto. In this respect, it should be briefly mentioned that Art 5, concerning requests for re-use of documents, specifies in its para 6 that, as a way of derogation, educational establishments, organizations that perform research activities and those that fund research are amongst the subjects which define terms and conditions for re-use of data according to their regulations (*ordinamenti*). At any rate, Art 8 of the decreto replicates Art 8 of the Directive in prescribing that the re-use of all documents shall not be subject to conditions, unless these are objective, proportionate, non-discriminatory and justified on grounds of a public interest objective. Also concerning conditions for re-use, according to Art 7(9-*bis*)(b) the re-use of research data shall always be free of charge.

Art 9-*bis*(1) reiterates that research data is re-usable given the respect of laws on data protection, when applicable. On this point, it shall be considered that, in its Opinion on the implementation draft, the Italian Data Protection

to the Italian context and the proposal for a second moral right of publication in the so-called 'D.d.l. Gallo': disegno di legge proposal no 395 'Modifiche all'articolo 4 del decreto-legge 8 agosto 2013, no 91, convertito, con modificazioni, dalla legge 7 ottobre 2013, n. 112, in materia di accesso aperto all'informazione scientifica', documents available at https://tinyurl.com/2ttxmyhm (last visited 30 June 2022); R. Caso, La libertà accademica e il diritto di messa a disposizione del pubblico in Open Access 1(1) *Opinio Juris in Comparatione*, (2018); R. Caso and G. Dore 'Academic copyright, Open Access and the «moral» second publication right' *European Intellectual Property Review*, (2021), available at https://tinyurl.com/dt65m4r3 (last visited 30 June 2022).

[161] The National Program for Research (2021-2027) was approved with resolution no. 74 of 2020, Official Gazzette general series, 23 January 2021; R. Caso, 'Open Data, ricerca scientifica e privatizzazione della conoscenza', Trento Law and Technology Research Group Research Paper no 48, (2022), 24.

Authority (*Autorità Garante per la protezione dei dati personali*) asked to consider introducing in Art 9-*bis* a more precise reference to Art 105 of the Italian data protection act, decreto legislativo 30 June 2003 no 196 (known as *Codice di protezione dei dati personali*).[162] The referred provision prohibits the use of personal data processed for statistical purposes or scientific research in order to adopt decisions or measures concerning the person, or for personal data processing personal data for scopes of a different nature.

Art 9-*bis*(1) also affirms research data is re-usable in observance with the respect of commercial interests (*interessi commerciali*), and the respect of laws on intellectual property (legge 22 April 1941 no 633) and industrial property (decreto legislativo 10 febbraio 2005, no 30). Looking at these safeguards, one should remember that documents on which third parties have intellectual property rights and industrial rights, with reference to the same aforesaid laws, are already excluded by the scope of application of the decreto in light of Art 3(1)(h). The provisions in Art 9-*bis*(1) seem therefore to mirror the safeguards specified in Art 4(b) and (e) of the Decree, but for the additional reference to commercial interests. Such reference is worth further attention because the subject of trade secrets (*segreti commerciali*), as informed by the Directive (EU) 2016/943, is traditionally framed under the discipline of industrial property in Italy. Trade secrets are disciplined under Arts 98 and 99 of the decreto legislativo no 30/2005. For this reason, trade secrets are already mentioned in Art 9-*bis*(1). One possible interpretation is that the addition should be understood in relation to Art 1(2)(d) of the Open Data Directive, that excludes from the scope of application documents 'such as sensitive data when access is excluded by national access regimes on grounds of national security, statistical confidentiality and commercial confidentiality.' However, if this is so, the Italian transposition should be criticized in making no explicit reference to any specific national access regime. As corroborated by the Senate Dossier[163] the reference seems, however, to be to the final part of Art 10(2) of the Directive, that ambiguously concludes that in the context of research data 'legitimate commercial interests, knowledge transfer activities and pre-existing intellectual property rights shall be taken into account.' In this case, as in the first hypothesis, the Italian provision may be criticized for establishing a limit that appears excessively broad and introduces considerable legal uncertainty.

Art 9-*bis(*2) specifies the conditions under which the re-use rules would apply, in transposition of Art 10(2) of the Open Data Directive. The first requirement provides that research data is 'the result of research activities' that are financed

[162] Autorità Garante per la protezione dei dati personali, Provvedimento no 308 del 26 agosto 2021, Parere sullo schema di decreto legislativo recante 'Attuazione della Direttiva (UE) 2019/1024 relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico', 4.

[163] Dossier no 436, 9 Settembre 2021, 'Apertura dei dati e riutilizzo dell'informazione del settore pubblico', Atto del Governo 284, 20, available at https://tinyurl.com/4467btke (last visited 30 June 2022).

by public funds. Taking into consideration the aforementioned difficulties of interpreting the funding requirement at the national level, it should be considered that no provision within the decreto seems to support a more precise reading of it. However, the interpreter may resort to the legge no 112/2013 that refers to research funded by 50% or more by public funds in relation to (the promotion of) open access mandates for scientific publications.[164] The second requirement recites that data has already been made public, also by archiving in a public database (which represents an addition compared to the Open Data Directive), by researchers, organizations that conduct research activities and organizations that finance the research, by means of a database managed at the institutional level or subject-based database.

Finally, Art 9-*bis*(3) establishes that research data 'complies' with FAIR requirements: findability (*reperibilità*), accessibility (*accessibilità*), interoperability (*interoperabilità*), re-usability (*riutilizzabilità*). By incorporating the requirements in the provision, the Italian legislator seems to have gone beyond that prescribed by the Directive. On closer analysis of the Directive, Art 5 on available formats mentions almost coincident requirements to be applied 'when possible and appropriate', while the FAIR principles are only mentioned in Art 10(1) in relation to open access policies and actions that member States shall support for making publicly funded research data available. Since Art 6 of the decreto on available formats makes fewer requirements mandatory, it seems possible that the introduction of the FAIR requirements in Art 9-*bis*(3) reinforces the conditions for the re-use of research data as compared to other categories of data.

Looking at the first part of Art 6 of decreto legislativo no 36/2006, this prescribes that public administration, bodies governed by public law and public enterprises shall, in addition to making their documents available, make the metadata available 'when possible.' The absence of a more precise obligation in the Italian transposition always to make the metadata available can be considered a missed opportunity, although Art 5 of the Directive prescribes this merely 'when possible and appropriate'. What seems remarkable when comparing Art 6 of the decreto and Art 9-*bis*(3), the reference to FAIR principles in Art 9-*bis*(3) could be interpreted as prescribing an obligation to make metadata available in the context of the re-use of research data. This seems a desirable reading because the principles as originally conceived by their authors should be applied to both.[165]

Closer scrutiny of Art 6 of decreto legislativo no 36/2006 reveals that, other than prescribing the principle of open by design and by default (Art 6(4)), this affirms that data shall be made available according to the definitions of 'machine-readable format' and 'open format' (Art 6(1), referring to Art 2(c-*bis*) and (c-*ter*)), while complying with technical rules to be adopted by the Agenzia per l'Italia Digitale (literally: the Agency for Digital Italy, hereinafter AgID) (Art 6(1), referring

[164] L. 112/2013 was the conversion, with amendments, of decreto legge 8 agosto 2013 no 91.
[165] M.D. Wilkinson et al, n 74 above, 4.

to Art 12).[166] At the time of writing, these have not been updated accordingly but a series of seminars has been organized to prepare the launch of the open consultation on the new draft Guidelines.[167] This is worth mentioning since Art 5 of the Directive refers to formats that are not only open and machine-readable, but also accessible, findable, and re-usable. Finally, as for the other definitional provisions of the decreto, when comparing the decreto and the Directive, the references to the CAD provided by the decreto should also be considered.[168]

Overall, it seems that only the new detailed rules set out by the the AgID will allow for a comprehensive account of the standards, also technical standards, to be applied to research data and the Italian transposition. Therefore, the present contribution is limited to preliminary conclusions, while a more solid understanding of the new rules on research data should be deferred for future work and hopefully will be based on the practical application by relevant research bodies, ie considering empirical data and best practices that will follow. For the time being, the contents of Art 9-*bis* allow the consideration that the rules on research data seem to enhance re-use, compared to other categories of data. As

[166] Agenzia per l'Italia Digitale, 'Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico', (2017), available at https://tinyurl.com/2p86fatd (last visited 30 June 2022). The document is within the objectives of Art 52 CAD.

[167] The seminar series are named 'Linee Guida per l'apertura dei dati e il riutilizzo dell'informazione del settore pubblico nell'ambito della strategia europea e il contesto nazionale in materia di dati' and they are part of the project 'Informazione e formazione per la transizione digitale per l'attuazione del Progetto Italia Login – la casa del cittadino' – PON Governance e Capacità Istituzionale 2014-2020. The fourth and last seminar is currently planned on the 15 June 2022.

[168] The definition of open format is in Art 2(1)(c-*ter*) of d.lgs. 36/2006, that refers to Art 1(1)(l-*bis*) of CAD. The CAD defines open as a format made public, exhaustively documented, and neutral in respect to the technological tools for the fruition of data. This seems partially different from the definition of open format prescribed by Art 2 no 14 of the Directive that establishes the format should be platform-independent and made available without restrictions impeding re-use. The definition of open format is actually similar to the one of open standard format in the Directive, given in Art 2 no 15 and referring to a standard in written form, detailing specifications for the requirements on how to ensure software interoperability. Furthermore, the decreto, contrary to the Directive, also defines open data (*dati di tipo aperto*) referring to the CAD: Art 21(c-*quater*) d.lgs. 36/2006 refers to Art 1(1)(l-*ter*) of CAD. The CAD provides the definition of open data with three key characteristics. First, open data are data available for everyone to use, also for commercial purpose, in a disaggregated format, according to a license or law disposition. Second, they are accessible through means of information and communication technologies, including public and private telematic networks, in open formats (within the meaning of Art 1(1)(l-*bis*) of the CAD), they are suitable for automatic use by computer programs and are provided with the relevant metadata. Third, they are either available at no cost by means of information and communication technologies, including public and private telematic networks, or available at marginal costs for reproduction and divulgation, given Art 7 of the d.lgs. 36/2006, as reformed in 2021, would apply. The decreto also contains a definition of 'data ownership' (*titolarità*) that closely mirrors the one introduced in the CAD after 2016 (Art 2(1)(i)). Art 1(1)(cc) of the CAD affirms that the data owner (*titolare*) is the subject that originally created for its own use or commissioned to another entity the document which represents the data, or the subject that owns (*disponibilità*) the document; the decreto adds the subject is the public body, who may have commissioned the document to another public or private subject. Relevantly, both the definition of open data and data ownership are not prescribed in the Open Data Directive, but they appear to ensure the consistency between the decreto, the CAD and other relevant laws applicable.

for the terms and conditions of re-use, Art 5(6) seems to introduce potential limits, but Art 8 would still prohibit the application of discriminatory conditions.

To complement this analysis on the re-use of research data, it is useful to mention that the provisions on data from cultural establishments in the decreto legislativo no 36/2006 have also been slightly amended by decreto legislativo no 200/2021. One amendment seems to introduce a limit for the re-use of cultural data that is not apparently mirrored in the text of the new Open Data Directive. The reference is to Art 1(2) of the decreto legislativo no 36/2006. The provision reaffirms the principle that the documents should be re-usable for commercial and non-commercial aims. For the documents held by libraries, including university libraries, museums and archives, however, an addition states that the re-use should be authorized according to a series of provisions relating to the Italian law for the protection of cultural goods and landscape (decreto legislativo 22 January 2004 no 42, known as Codice dei beni culturali e del paesaggio, also Codice Urbani) and protection of personal data (decreto legislativo no 196/2003). More precisely, references to a specific authorization according to those two laws were already present in the decreto legislativo no 36/2006 before 2021. The references to the Italian data protection law in Art 1(2) of decreto legislativo no 36/2006 have remained the same, and they namely refer to part II, title II, chapter III of the decreto legislativo no 196/2003 and thus Arts 101-103 on the processing of personal data for historic purposes. The references to the law for the protection of cultural goods and landscape on the other hand have changed. The previous provisions linked to Part II, Title II, Chapter III of Codice Urbani and thus Arts from 122 to 127, regarding the possibility to consult archives and protection of privacy. However, today the link is to Part II, Title II, Chapter I and Chapter III and thus Arts from 101 to 110, regarding all the existing constraints for the fruition of cultural goods. These include most prominently the authorization for the use of the goods (Art 107 Codice Urbani) and fees for its concession and reproduction (Art 108 Codice Urbani).

This amendment can be questioned, since it is not clear the extent to which the reference to such rules – limiting the use of the cultural good – may impact the use of related data. What seems undisputed is that the nature, as well as the rationale, of the rules to be followed for the re-use of cultural data have changed: the mentioned provisions concern limits for the use of cultural goods that do not relate anymore to the protection of privacy, but refer to the need to protect cultural heritage. When the use of the good has commercial purposes, these imply relevant burdens. On the contrary, if the activities are for purposes of study, research, free thought and creative expression, promotion of knowledge of the cultural heritage, they are defined free (*libere*) by Art 108, after this was recently reformed.[169] It

[169] Legge 29 July 2014 no 106 (conversion, with amendments, of decreto legge 31 May 2014 no 83) and legge 4 August 2017 no 124, have modified Art 108(3-*bis*) Codice Urbani; F. Minio, 'La libera riproducibilità dei beni culturali dopo l'emanazione della legge 4 agosto 2017, n. 124 (legge annuale per

also bears emphasis that such limits operate independently from the copyright status of the work, and thus also when the work is in the public domain. For these characteristics, the same provisions of the Codice Urbani are also highly debated – and criticized – in relation to the implementation of the new Art 14 of the Copyright Directive in the Digital single Market, Directive (EU) 2019/790 (CDSMD) seeking to allow free reproductions of works of visual arts in the public domain.[170] Regrettably, the new Art 32-*quater* of the legge no 633/1941, introduced within the implementation of the CDSMD in 2021,[171] specifies that the rule is without prejudice to the provisions on the reproduction of cultural goods set out in the Codice Urbani. This appears to weaken the most recent Government initiatives that support the opening of images of the Italian cultural heritage,[172] and this work appreciates how the newly introduced limits in Art 1(2) of decreto legislativo no 36/2006 may be criticized for the very same reasons.

The relationship between the digitization of cultural heritage, including the circulation of images from the public domain, and the re-use of data from cultural establishments remains inconsistently addressed by the described national laws, as amended. At present, consistent efforts to elaborate guidelines for managing both the reproductions of works of cultural heritage and related data and metadata, while navigating the current framework, can be found in the plan and guidelines for the digitization of cultural heritage provided by the Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library (part of the national Ministry of Culture).[173] The public consultation of these documents, now open until the 15 June 2022, seems therefore a chance to elaborate more comprehensive policies on the topic.

il mercato e la concorrenza)' *BusinessJus* 76, (2018); M. Modolo and A. Tumicelli, 'Una possibile riforma sulla riproduzione dei beni bibliografici ed archivistici' *Aedon*, (2016); G. Gallo, 'Il decreto Art Bonus e la riproducibilità dei beni culturali' *Aedon* (2014).

[170] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L130/92; M. Arisi, 'Digital Single Market Copyright Directive: Making (Digital) Room For Works Of Visual Art In The Public Domain' 1(1) *Opinio Juris in Comparatione*, 1 (2020).

[171] Directive (EU) 2019/790 was transposed in Italy with decreto legislativo 8 novembre 2021, no 177.

[172] Risoluzione In Commissione Conclusiva di Dibattito 8/00126 of June 2021, available at https://tinyurl.com/2v2w8vyt (last visited 30 June 2022).

[173] Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library, Ministero della Cultura, 'Piano nazionale di digitalizzazione del patrimonio culturale 2022-2023 and, in particular Linee guida per l'acquisizione, la circolazione e il riuso delle riproduzioni dei beni culturali in ambiente digitale' (2022, version for public consultation), available at https://tinyurl.com/2s9dhf3s (last visited 30 June 2022). The initiative is part of the Piano Nazionale di Ripresa e Resilienza (also known as PNRR and named Italia Domani), that is the italian translation for the Recovery and Resiliency Facility part of the Next Generation EU program, Council Regulation (EU) 2020/2094 of 14 December 2020 establishing a European Union Recovery Instrument to support the recovery in the aftermath of the COVID-19 crisis, OJ L 433I/23. The branch M1C3 of the PNR, dedicated to tourism and culture, entails objectives of digitization of cultural heritage under the strategy 1.1. All documents are available at https://tinyurl.com/3uj2yfmz (last visited 30 June 2022).

However, despite growing interest in how to make cultural heritage more open, in view of the above, following the implementation of the CDSDM and the Open Data Directive in Italy, it seems that national legislator remains rather reluctant to open data from cultural establishments. This frustrates the hopes of those commentators looking favorably at the potential of the PSI rules for cultural digital heritage,[174] while it also seems to dismiss the convergence of policy objectives suggested by the recent Commission Recommendation on a common European data space for cultural heritage. A fundamental discrepancy to be solved in the near future seems to rely on the fact that while current laws on the protection of cultural goods limit the use and re-use of cultural goods for commercial purposes, the PSI rules embrace, and actually promote, the re-use for both commercial and non-commercial purposes.

VII. Conclusions: An Open Directive?

Legal mandates are crucial to fully realize the re-use of publicly funded research data to promote Open Knowledge, for the need to provide relevant subjects with clear obligations and rules that would help them to conduct the complex balance between rights and interests that characterizes the research environment, with primary reference to intellectual property rights and personal data protection rights. From this perspective, the inclusion of research data in the scope of the Open Data Directive should be welcomed as a positive amendment to the PSI rules in the European Union. The new Directive represents a stronger initiative to promote an increasingly harmonized access to publicly funded research, when compared to the previous open access and open science initiatives, lacking a binding nature. Crucially, it also seems that the Open Data Directive will be complemented by a series of even more impactful legislative initiatives on data within the EU Data Strategy that will also address PSI and research data.

Nevertheless, looking more closely at the new PSI Directive of 2019 and the provisions on research data, it may be argued that their open vocation, despite the Directive being entitled after open data, remains at times frustrated by significant and detailed limitations, especially with regard to the relationship with intellectual property law, with detriment to legal certainty. More specifically, while the open data definitions imply that data is free from *legal* and technical barriers,[175] this paper has tried to describe how the new EU PSI rules on research data and, to some extent, data from cultural establishments appear often complex or difficult to interpret. Finally, this entails their scope of application and safeguards largely depend on the national implementation.[176]

[174] M.C. Pangallozzi, 'Condivisione e interoperabilità dei dati nel settore del patrimonio culturale: il caso delle banche dati digitali' *Aedon*, (2020).

[175] F. Zuiderveen Borgesius et al, n 88 above, 2079.

[176] S. Gobbato, n 62 above, 159.

This was confirmed by the analysis of transposed rules in Italy, where relevant uncertainties remain as for the scope of application, ie addressed organizations, and limits of re-use of research data, as well as for the re-use of data from cultural establishments. However, it should still be viewed favorably that the national legislator has addressed research data adopting targeted provisions, to date in the absence of mandatory provisions aimed at opening research data. As mentioned, the detailed rules to be set out by the AgID will allow for a comprehensive account of this reform and its practical application, but it seems already plausible to conclude that the hereby described complex national regulatory framework should be subject to further study in the very near future to complement the analysis sketched by the present paper. The EU project of further promotion of a Data Strategy, including the proposals for the Data Governance Act and Data Act aforementioned, suggests the attempt to strike a balance between openness and closure of data in both the public and private sector, so the Open Data Directive would only be the starting point of a new discussion on open knowledge and public sector information, the interplay of decreto legislativo no 36/2006, decreto legislativo no 33/2013, the CAD, personal data protection and intellectual property.