Where is Open Data in the Open Data Directive?

1. Introduction

The re-use of non-sensitive government data continues to attract attention within the legal sphere of the European Union (EU). The latest regulatory iteration in this field is the Directive on Open Data and the Re-use of Public Sector Information, to be known as the Open Data Directive (ODD) of 2019 (European Parliament and Council, 2019). It is a recast of Directive 2003/98/EC on the re-use of public sector information, known as the Public Sector Information (PSI) Directive (European Parliament and Council, 2003), which is entrenched in the internal market and focused on reusing data for economic gain. Yet, the new directive further embraces the concept of 'open data', a concept rooted in openness and transparency for citizen participation.

At first blush, it may appear that the ODD is equally concerned with promoting both 'open data' and the re-use of 'PSI'. Indeed, in this marriage of concepts, the open data concept has been given the honour of obtaining the 'family name'- elevated to replace 'PSI' in the short title of this directive. However, despite enjoying a synergistic relationship (Janssen, 2011a), these two concepts have disparate origins and drivers, which should alert us to implicit dangers in seeking their co-location in a single instrument. Given the clear economic drivers behind the re-use of PSI, it is important to interrogate and elicit which aspects of open data the directive embraces, and which fall outside of its purview. There is a risk that open data will simply be appropriated and may impact and transform our conceptual understanding of the meaning of open data and constrain the possibilities for promoting and entrenching legally ideas of citizen participation. Therefore, the present paper undertakes a critical analysis of the ODD. It asks two questions.

First, how does the directive align with and diverge from the rationale and requirements of the movement for open data? Second, what are the implications of this for citizen participation? The analytical framework for the present paper is based on contrasting the disparate movements for open data and re-use of PSI with the directive's legal text.

The risks of conflating these two movements in a single legal order have hardly been researched. Indeed, there is a general dearth of research around the re-use of PSI or open data, and what exists is largely focused on technological, organisational and data-related issues (Wirtz et al., 2019). Few have taken a critical perspective on re-use, access, open data and the law. A key exception is Janssen, whose most recent writing dates back to the early 2010s. Janssen (2011a) identified a legal confusion between the regulation of access or freedom of information (FOI) and re-use at the member-state level. Moreover, several researchers at the time (Bates, 2012; Jaatinen, 2016; Kitchin, 2014) predicted that open data would become dominated by economic interests to the detriment of openness, transparency and citizen participation. However, the issue has lain dormant since then; hence, it is important to take a fresh look at the 're-use of PSI' in the context of open data.

Such an inquiry is not only necessary on account of the new and seemingly hybrid directive. It is also required because the 're-use' of PSI exists in a substantially different policy context than a decade ago. Data is now considered to be the lifeblood of economic development in Europe (European Commission, 2020). The ODD's direct impact alone predicted a rise from €52 billion in 2018 to almost €194 billion in 2028, with the number of jobs based on the PSI expected to reach 709,000 (European Commission, 2018a). Re-use of PSI regulation no longer exists on the side lines within a relatively specialist domain (Bates, 2012). Instead, it commands a central position in policy discourse, constituting a vital weapon in the EU's arsenal to nourish and expand the European data-driven economy.

The present paper is structured as follows: Section 2 traces the origins of the ODD in the distinct movements for open data, re-use of PSI and FOI, which is an often overlooked yet important protagonist in this terrain, given its role in facilitating access to data. In section 3, the main legal provisions of the ODD are outlined, followed by an analysis in section 4 that identifies which elements of open data are incorporated into the directive and discusses the implications this might have for citizen participation. The paper culminates with concluding remarks in section 5.

2. Understanding 'Open Data', 're-use' and 'FOI'

Words and concepts are often used without knowledge of their past, with their meaning taken for granted and rarely questioned (Bathelt et al., 2017). Yet they often contain the remnants, if not power, of their prior meanings, requiring a more diachronic approach to our analysis of concepts. To answer the questions posed in the current paper, it is critical to understand the past, trace the evolution of and question the meanings behind the central concepts of 're-use of PSI' and 'open data'. Although undoubtedly sharing many traits, 'open data' and 're-use of PSI' have different histories and are built on different foundations that have distinct identities and values. It is also important to chart the development of FOI because access and re-use are interlinked, especially when contrasted with open data. The re-use of PSI allows for the right to use information for both commercial and non-commercial purposes, as long as the use is outside of the public task (Art. 2(11) ODD). Access to information is a necessary precondition for 're-use' - without access there can be no re-use. The EU has limited competence to directly regulate general access to the information held by the public sector in member states; therefore, the decision about what is made public remains predominantly a domestic one (van Eechoud & Janssen 2012), controlled largely but not exclusively by national rules for access to documents in the form of FOI laws. Delineating and understanding the concepts of open data and re-use of PSI and FOI is necessary to facilitate a critical analysis of the ODD and its implications.

2.1. Open data

The open data concept emerged in the early 2000s; it was a movement spearheaded by an ecosystem of non-state actors seeking to pressure governments to proactively share non-sensitive data and make these data freely available to use, re-use and redistribute without any legal, technological or social restrictions (OKFN, 2012). The concept was not limited to government data but included data from other providers, such as the private sector and civil society. However, government data was the dominant focus, being seen as a cornerstone of a new model of open government; this was widely understood as leveraging data and information and communication technologies (ICTs) to generate participatory and collaborative dialogue between policymakers and citizens (Moon, 2020; Wirtz et al., 2016). The push for open data was not especially legalistic or envisaged an explicit or specific legal framework. In fact, there is no legal definition of open data in statutes or case law (De Filippi & Maurel, 2015). However, the concept builds upon the origins of FOI for transparency and accountability for public trust (Janssen, 2011b), which has legal dimensions.

Open data was firmly rooted in ideas of open government and citizen participation. Citizen participation has been deemed both a democratic right and a process through which citizens engage in the public sphere to shape policy (Kalandides, 2018), here based upon the fundamental principle that citizens are at the heart of open government (Wirtz et al., 2019). Giving the public access to information on initiatives and decisions taken by government actors could facilitate their participation and strengthen the accountability of governmental institutions (Harrison & Sayogo, 2014). Hence, open data has been deemed a vital ingredient for citizen participation, given that it ensures proactive access to information that citizens would otherwise not be privy to. The movement sought to break the interpretive monopoly of governments and allow the public to produce their own interpretations of public data (Kennedy et al., 2015), here driven by an optimism about what could be accomplished politically through the use of ICTs (Harrison & Sayogo, 2014; Moon, 2020). It is also recognised as an enabler for e-participation,

where ICTs are exploited to engage citizens in supporting democratic decision-making processes and strengthening representative democracy, bringing with it what Meijer (2009) has termed 'computer-mediated transparency'.

The first major international initiative around open government data emerged in 2007 when a group of open government advocates united to establish a 'more robust understanding of why Government Open Data is essential to democracy' (public.resource.org, 2007). They published a list of eight principles, the justification of which is as follows:

> *The Internet is the public space of the modern world, and through it, governments now have the opportunity to better understand the needs of their citizens and citizens may participate more fully in their government. Information becomes more valuable as it is shared, less valuable as it is hoarded. Open data promotes increased civil discourse, improved public welfare, and a more efficient use of public resources … By embracing the eight principles, governments of the world can become more effective, transparent, and relevant to our lives.* (public.resource.org, 2007)

The principles themselves entail that data should be *complete* such that all data should be made available; data should be *primary,* meaning that data should be collected at the source; data should be made available in a *timely* manner; *accessible* to the widest range of users and purposes; *machine processable* to allow for automated processing; made available in a *non-discriminatory* manner such that there are no requirements for registration; shared in *non-proprietary formats* and data should be made accessible under *free licences* (Public.resource.org, 2007).

The US and UK governments were frontrunners in sharing data on official open data portals under open licences, and many countries swiftly followed suit (Mellouli et al., 2014; Moon,

2020). This all occurred under the banner of signalling commitment to increasing transparency, accountability and participation, along with improving public services and value creation in the private sector (Moon, 2020; Wirtz et al., 2016; Worthy, 2015). In 2019, 27 of the 28 EU member states had an open data policy, with 20 stating that their policies were more ambitious than the 2013 PSI Directive (Pollock et al., 2019), which regulated 're-use'. Indeed, this is acknowledged in Recital 18 ODD: 'Member States have established re-use policies under Directive 2003/98/EC and some of them have been adopting ambitious open data approaches to make the re-use of accessible public data easier for citizens and legal entities beyond the minimum level set by that Directive'.

Multilateral initiatives such as the Open Government Partnership and international bodies such as the OECD and World Bank have long since prioritised open data. The OECD defines open government data as a set of policies that promote transparency, accountability and value creation by making government data available to all (OECD, 2022). The G8 launched an open data charter (G8, 2013), which subsequently became the International Open Data Charter and was adopted by over 100 governments and organisations. Although they have a direct normative effect (De Filippi & Maurel, 2015), these charters, recommendations and guidelines emerging on the international stage were non-binding, meaning that the ODD is of particular importance. The EU has been the first transnational body to enshrine, at least in name, open data into hard law.

2.2. Re-use of public sector information (PSI)

The concept of facilitating the secondary use of non-sensitive government data at the EU level, termed the 're-use of PSI', can be traced back to the 1970s (Bates, 2012). It became apparent that the public sector generated large amounts of non-sensitive data and was the single largest collector, user, holder and producer of certain kinds of information (Gonzalez-Zapata & Heeks, 2015). The EU set about providing a common legal framework for a European market for

government-held data that could ensure a level playing field between the public and private sectors and build a vibrant European information market capable of competing with the US (Andrasko & Mesarcik, 2018; Janssen, 2011b). Societal benefits were recognised, but the main target group was the information industry (Janssen, 2011a). Building on many years of consultation, the first official PSI document was published in the form of 'Guidelines for Improving the Synergy Between the Public and Private Sectors in the Information Market' (European Commission, 1990). These guidelines had a limited impact with fragmented adoption, partially because of their non-binding nature. Nevertheless, 13 years later, a binding minimum harmonisation legislative instrument was put in place in the form of the PSI Directive in 2003 (European Parliament and Council, 2003); this directive set out a general framework establishing a set of minimum rules for the governing of and practical arrangements for the facilitation of the re-use of PSI, which came in the form of removing the major obstacles for the re-use of information and introduced uniform rules on key issues such as pricing, licencing and exclusive arrangements (Dalla Corte, 2018).

In the early 2010s, it became clear that this directive had not realised its ambitions (Janssen, 2011b). Technological evolution had also progressed to facilitate the greater sharing and use of PSI (Jaatinen, 2016), and the economic potential of digitalisation and need for a vibrant digital single market became increasingly apparent to and prioritised by the EU. The Digital Agenda for Europe (European Commission, 2010) called for a review of the directive and paved the way for the amendments enacted in 2013 through Directive 2013/37/EU PSI Directive (European Parliament and the Council, 2013). Under the previous directive, the implementation of the main general principle of the right to re-use had been left to the discretion of the member states (van Eechoud & Janssen, 2012). However, in 2013, the right to re-use was codified, making all documents reusable unless access was restricted or excluded under national rules on access to documents (De Filippi & Maurel, 2015). Open data was also introduced as a concept in the preamble to the PSI Directive of 2013 in the form of acknowledging and encouraging

open data policies (Recitals 3 and 6), open licencing (Recital 26) and open data formats (Recital 26). The latest iteration of the directive comes as a recast in the form of the ODD, which is discussed in section 3 below.

2.3. Freedom of information (FOI) and access to PSI

FOI can be broadly defined as a presumptive right of individuals to access information held by public authorities so that they can hold government transparent and accountable, build public trust and, thereby, widen public participation (Worthy, 2010). FOI-type laws have been adopted in most countries throughout the world, being recognised as the legal backbone for creating and safeguarding a basic level of transparency (Grimmelikhuijsen et al., 2018). FOI is vital in understanding the context of 're-use'. As previously stated, the EU has limited powers to regulate 'access' because it falls outside their competence. Recital 23 of the ODD states, *'This Directive builds on the existing access regimes in the Member States and does not change the national rules for access to documents'.* Therefore, re-use sits on top of a variety of national access rules that differ in scope and field of application (van Eechoud & Janssen, 2012).

Today, FOI is seen as embodying 'old open government', in which citizens and businesses are information requesters and recipients (Moon, 2020). The concept of open data is deemed to be the epitome of 'new open government' because it extends FOI on a practical level through the proactive sharing of open and reusable datasets on open data portals, which are enabled by the innovations in and diffusion of ICTs. By enabling citizen participation and improved governance and challenging the mere reliance on reactive requests for information, as is prominent in FOI, open data approaches are deemed as going beyond the simple promotion of citizens' rights to know (Gonzalez-Zapata & Heeks, 2015; Lassinantti, 2019; Moon, 2020). The key distinction here is between 'passive' and 'active' transparency. Passive transparency entails legal compliance with requests to make data available, as opposed to active transparency, which means the simultaneous proactive disclosure of information for and to all who might be interested (Grimmelikhuijsen et al., 2018). The open data movement has been widely credited

with building upon and extending FOI traditions (Gonzalez-Zapata & Heeks, 2015) and catalysing the shift from 'old' to 'new' open government (Moon, 2020).

The more successful open data policies are at proactive sharing of data, the more data will be available for re-use. Herein lies a crucial component of the synergistic relationship between 'open data' and 're-use'. For example, a national government that proactively publishes its company register as open data will ensure that these data are available for *re-use*, regardless of whether that data is subject to *FOI* laws at the national level. We now turn to discussing the content of the new directive before analysing its coherence with the open data movement.

3. The Open Data Directive

Despite their disparate origins, the 're-use of PSI' and 'open data' have courted each other for many years. Indeed, they seem like natural bedfellows given their many similar traits. Open data is often credited with being instrumental in lifting PSI from relatively narrow academic and professional circles to being pervasive in government policy throughout Europe (Dalla Corte, 2018; Janssen, 2011b; Pollock et al., 2019). Open data's influence on PSI can be formally traced back to 2011, when the European Commission launched a communication on open data (European Commission, 2011), going on to incorporate the term 'open data' and adopt some of the technical requirements, such as open data formats, into the PSI Directive in 2013. This was followed in 2014 by Guidelines on recommended standard licences, datasets and charging for the re-use of documents, which relied heavily on open data terminology and recommendations (European Commission 2014). By 2018, data had been elevated to the higher echelons of EU policy. A review of the PSI Directive was initiated, culminating in the ODD.

There was, and still is, a palpable sense of urgency in the EU to capitalise on public sector data to feed the data economy, which is a key factor in why we have witnessed more developments in the PSI Directive during the past decade than the previous four combined. According to the European Commission (2020), a well-functioning and high value–creating single market for

data requires a strong legal framework. Therefore, the ODD must be seen as part of a package of measures aiming to facilitate the creation of a common European data space, with the EU progressively asserting dominance over national data policy. Key among these are the Data Governance Act (DGA) (European Parliament and the Council 2022a) and the Data Act (European Parliament and the Council 2022b) proposal. The DGA contains rules on the re-use of PSI but differs from the scope of the ODD in the sense that the rules apply to data held by public sector bodies, which are protected on grounds of commercial confidentiality, statistical confidentiality, protection of intellectual property rights of third parties and the protection of personal data (Art. 3(1)).

The ODD continues to recognise that access regimes are within the remit of the member states (Art. 1(2)). However, it does bring with it some notable new provisions on 'access'; it obliges public bodies to make dynamic data available by public bodies immediately after collection (Art. 5(5)), as long as this does not impose a disproportionate effort (Art. 5(6)). Most striking, though, is the new chapter on high-value datasets (HVDs), which provides for a list of HVDs by way of an implementing act, which will be discussed in section 4.3. Otherwise, the most notable changes include further limiting the possibilities for charging, which constitutes a market barrier for SMEs and start-ups (Art. 6), strengthening the transparency requirements for PSI-related public–private agreements and limiting agreements that could lead to exclusive re-use by private partners (Art. 12). Research data resulting from public funding (Art. 10) and public undertakings, such as transport operators and energy providers, are now also included in its scope (Art. 1(1)(b)) because these are deemed as having tremendous re-use potential (European Commission, 2023b).

However, although open data has been elevated into the directive's title, replacing PSI in the short name and is recurrent in the preamble, it is rarely mentioned in the substantive provisions. The most notable is Art. 1(1), which states, *'In order to promote the use of open data and stimulate innovation in products and services, this Directive establishes a set of minimum rules*

governing the re-use and the practical arrangements for facilitating the re-use.' However, this is essentially limited to the scope of the potential for open data. It establishes that the directive is an instrument to promote the re-use of open data rather than regulating open data per se. Otherwise, only two other mentions are made, here both in relation to HVDs, namely in Article 14(1) stating that *'investments made by the Member States in open data approaches, such as investments into the development and roll-out of certain standards, shall be taken into account and balanced against the potential benefits from inclusion in the list'*, and in Article 16, which provides for the establishment of the Committee on Open Data and the Re-use of Public Sector Information to assist in defining the measures for the HVD Implementing Act(s).

4. Alignment and Misalignment

Having outlined the key concepts, I proceed to answering the central questions of the present paper: How does the directive align with and diverge from the rationale and requirements of the movement for open data, and what are the implications of this for citizen participation? As with the boundaries between the concepts, areas of alignment and misalignment in the directive are not always clear-cut. Therefore, I group them according to where they predominantly converge or diverge. I begin by examining the definitions in the directive, discussing how the concepts are not synonymous. I then highlight the dominance of economic interests and the legal confusion stemming from the directive, which is followed by a discussion on how HVDs may indeed contribute to more open data, but that these data are more likely to serve economic needs rather than facilitate citizen participation. I then proceed to explore the main area of alignment, herein the adoption of the open data principles.

4.1. Defining the re-use of PSI and open data in the ODD

The EU has stopped short of defining open data in the ODD, instead choosing to describe it in Recital 16 as being *'generally understood to denote data in an open format that can be freely used, re-used and shared by anyone for any purpose'.* This echoes and aligns with the Open

Knowledge Foundations (OKFN) definition as *'data that can be freely used, shared and built-on by anyone, anywhere, for any purpose'* (James, 2013). Hence, there is still no legal definition of open data either in statutes or case law, as was identified almost a decade ago by De Filippi and Maurel (2015). We now turn to the type of data that the ODD regulates; it classifies documents that contain personal data, third-party intellectual property and sensitive data that are excluded from national access regimes as being outside of the scope (Article 1(2)). This corresponds with the open data movement, where the focus is on the sharing of non-personal data, also recognising other limitations such as security restrictions (OKFN, 2010).

There is, however, misalignment in what is understood by data. Bates (2012) pointed out that the movement for open data advocated that data be predominantly shared in the form of structured quantitative datasets. A good example is the global open data census, which was previously performed by the OKFN and that ranked countries according to the release of structured datasets, such as statistical, environmental, spending and company data (OKFN 2016). This differs from the ODD, which utilises the term 'documents' and has a broader scope than the structured data generally promoted by open data advocates; in Article 2(6), the ODD has defined documents as *'any content whatever its medium (paper or electronic form or as a sound, visual or audiovisual recording); or (b) any part of such content'.* This means that all forms of documents, be they structured or unstructured, even in paper form, are included within the scope of the ODD. It is interesting to note that, in this area, PSI is more compatible with FOI, with its broader understanding of data/documents, than the open data movement. Indeed, Janssen (2011a) warned that an emphasis on machine-readable raw data rather than understandable information might even replace the right to information under FOI, here given that most FOI acts require information to be disseminated in a legible and understandable form.

Another area of misalignment is the use of the term 'open data', not 'open government data', in the ODD, which many organisations, such as the OECD, are consequent about using. One could argue that, because PSI is in the full title, it is clear that the scope is limited to the public

sector. However, given that the directive is to be known by its short title of open data, the term matters. This may impact the conceptual understanding of open data, narrowing 'open data' to just government data instead of all types of providers, as was originally intended by the open data movement.

4.2. Economic interests and legal basis

The EU can only exercise the competencies that treaties confer upon the institutions (Van Den Brink, 2017). The European Commission has been consistently clear that 'access to information' is outside of their competence and that the re-use of PSI is an economic instrument. Nevertheless, Janssen (2011a) found that this was insufficient to avoid confusion between FOI and PSI at the member-state level, which she attributed to including both commercial and non-commercial use in the definition of re-use (Art. 2(11)). Various countries have chosen different methods of implementation. Some, such as Ireland, Germany and Greece, have specific PSI legislation, while others have chosen to incorporate it directly into FOI laws, such as in the cases of Norway, Estonia and Portugal. This confusion is not reserved for the member states, but it also extends to the open data community. Pollock et al. (2019), who are prominent open data actors, wrote that the PSI Directive is built on two main pillars of the internal market: transparency and fair competition. Transparency, however, does not refer to government transparency but instead to increasing transparency regarding the procedure for obtaining PSI and the conditions and charges under which it could be re-used (Art. 7). A decade ago, this confusion was largely deemed positive because, regardless of the drivers, it pressured governments to share more data (Cerrillo-i-Martínez, 2012). Janssen (2011b) found that many member states actually called for a stronger commitment to data sharing during their negotiations for the PSI Directive, with transparency and democracy in mind, not economic re-use (P.453).

It is fair to revisit the issue of economic interests and legal confusion a decade later, particularly now that the data economy has been designated the crown jewel in Europe's future economic

success. Pollock et al. (2019) noted that 'the EU focus on creating a digital single market cannot be ignored when seeking to understand the current state of open data in the region'. The ODD remains clear and does not try to disguise the fact that the legal basis remains in Article 114 of the Treaty on the Functioning of the European Union (TFEU), which relates to the establishment and functioning of the internal market. The realisation of the right to business, which is heavily attuned to the economic value of this information (Andrasko & Mesarcik, 2018), remains the primary objective of the ODD. It explicitly recognises the importance of promoting accountability and transparency; however, this must be understood as a 'nice-to-have' knock-on effect only making an appearance in the recitals (13 and 14). Turning to supporting documentation, the impact assessment of the ODD asks '*What is the value added of action at the EU level is?'* The response is the following:

> *[C]reate further incentives for businesses to use the scale of the single market to produce information products and services that cover different countries. At the same time, the changes will ensure that commercial re-users of different sizes and investment capacities will have similar opportunities in terms of access to public data.* (European Commission, 2018b, p. 2)

A further question posed was whether there would be other significant impacts. The answer was that the ODD is expected *'to significantly increase the current societal impact of the re-use of PSI'.* Indeed, the term 'social' is used frequently in the recitals and provisions (Art 5(6) and 18) and could be understood as encompassing openness and transparency. Therefore, it is crucial to explore what is actually meant by social impact in the ODD context. The assessment states that it would come in the form of benefits to consumers from products and services, such as quicker responses to cardiac arrest and a reduction in traffic deaths. The main social impact, then, is the creation of innovative products, with citizens as the end users or beneficiaries rather than a societal impact in the form of citizen participation. Or as Valli Buttow and Weerts (2022) stated, the well-being of citizens has been assumed simply as an automatic result of the data-

driven economy. Citizen participation is acknowledged in the assessment: *'Open data helps in reducing the distance between government and citizens as regards access to information and generates significant benefits in the areas of social inclusion and civic participation'* (European Commission, 2018a, p. 43). Notably, however, the term open data is used, not 're-use of PSI', hence acknowledging that this is a particular driver for open data, not re-use.

The above analysis shows that the ODD is clear regarding its legal basis and scope and has been since the original PSI Directive. This has, however, been insufficient to avoid legal confusion at the member-state level in the past. Renaming it the ODD and increasing references to open data is likely to perpetuate, rather than mitigate this confusion at the national level. And although a decade ago such confusion worked in favour of transparency and accountability for citizen participation, this is unlikely to persist now, given the economic policy context within which the ODD operates and the predominance of economic interests.

4.3. High-value datasets

It is increasingly recognised that simply publishing vast amounts of datasets does not guarantee realising the predicted value, hence leading to an increased international focus on HVDs. The ODD has adopted this concept, providing for HVD Implementing Acts (Art. 14). Article 2(10) of the ODD states, '*High-value datasets means documents the re-use of which is associated with important benefits for society, the environment and the economy, in particular because of their suitability for the creation of value-added services, applications and new, high-quality and decent jobs, and of the number of potential beneficiaries of the value-added services and applications based on those datasets'.* They are deemed to have a high commercial potential capable of speeding up the emergence of value-added, EU-wide information products and being a key data source for artificial intelligence (European Commission, 2023a). Society is regularly mentioned by the European Commission in relation to HVDs, but as established in the previous section, societal impact is regarded as coming in the form of access to improved services, not necessarily the empowerment of society to participate in government.

The first HVD Implementing Act was recently adopted in the form of an implementing regulation, providing a list of specific HVDs and the arrangements for their publication and re-use (European Commission 2022c). Member states are legally obliged to share multiple specific datasets to ensure that they are available across Europe; these datasets shall be available for free, utilise open standard licences, be machine-readable, provided via APIs and, where relevant, offered as a bulk download. Many states, such as Spain, France and Denmark, have prioritised the sharing of specific datasets; however, this is the first time that the PSI domain has demanded common ones across Europe. Interestingly, making this mandatory shifts the ODD from regulating purely for governing and facilitating re-use to also regulating access, thereby blurring the line between re-use and access. HVDs will increase the amount of data available as open data and have the power to shape the open data ecosystem of Europe for decades to come (Davies, 2019), hence requiring interrogation in the citizen participation context.

Datasets suitable for economic value creation can differ significantly from those that stimulate and facilitate citizen participation. It is here that the argument for the morphing of open data and re-use of PSI—because they share the same goal of an increased amount of data shared—becomes rather tenuous. The type of data shared matters. For example, in 2011, the UK government prioritised the publication of contracts, officials' salaries and crime maps to create an effective 'open data ecosystem' (Worthy, 2015), echoing the movement for open data's emphasis on sharing spending budgets and the activities and tasks of public bodies (Janssen, 2011a). These data were not in demand by the private sector but instead were made available for the purpose of improved citizen participation. Therefore, the datasets chosen matter.

The economic dominance in the definition of HVDs and supporting documentation has not gone unnoticed. Civil society organisations have reacted, raising concerns about a lack of public engagement, questioning how diverging economic and social values would be weighted and

warning that the European Commission should not simply rely on quantitative indicators, such as potential income generation (Pollock et al., 2019; Pugh, 2019). The process of deciding on HVDs may have taken heed to civil society concerns. The HVD impact assessment (Deloitte et al., 2020) adopted five categories from which to assess value: climate change and environment; economic; innovation and artificial intelligence; public services and public administration; and re-use and social. Notably, in this context, 'social' has been defined as *'improving transparency and accountability and creating other important benefits for society'.* The study stated that it relied upon open data literature as a basis for defining social value (Deloitte et al., 2020, p. 7). There was clearly a greater appreciation of citizen participation in the impact assessment; however, it cannot be disregarded that the business sector has remained the prime target for HVDs.

Turning to the chosen datasets, all are structured within the thematic categories of geospatial, earth observation and environment, meteorological, statistics, companies and company ownership and mobility (European Commission 2022d). None of the usual suspects deemed important for citizen participation, such as government spending, budgets, election results or legislation, were chosen. It is noteworthy that the chosen categories are the same as those identified in the European Commission's guidelines in 2014 (European Commission 2014), so much of this first round of HVDs may have been pre-decided. This implementing regulation is, however, just the start, with the commission empowered by Article 13(2) of the ODD to add new categories, but only ones that reflect *technological* and *market* developments.

A lesson learned from years of sharing government data is that there are significant costs involved. The impact assessment found that the costs for sharing HVDs would be substantial (Deloitte et al., 2020, p. 304), as evidenced in Table 1, which shows the estimated costs for just three of the datasets per member state. For statistical data the infrastructural costs alone deemed to be up to 3 million euros with a further operational cost of between 100-200k per annum. Member states will be legally obliged to share HVDs and devote significant resources to this

task, meaning that these datasets will likely top the list of data-sharing priorities at the national level. In turn, this may lead to the de-prioritisation of the release of other datasets that are more suitable for citizen participation.

4.4. Alignment with the open data principles

The various iterations of the PSI Directive have increasingly aligned with the open government data principles (Janssen & Hugelier, 2013), as described in section 2.1. There has been a cascading adoption of these found predominantly, but not exclusively, in Chapter III, 'Conditions for Re-use' in the ODD. In Article 5(1), we find that public sector bodies and public undertakings shall make their documents available in any pre-existing format or language where appropriate and possible, making documents, together with their metadata, available by electronic means in formats that are open, machine-readable, accessible, findable and re-usable. This corresponds with the principles of *complete*, *accessible*, *machine processable* and *non-proprietary formats*. Encouraging 'Open by design and default' (Art 5(2)) respects the *primary* and *timely* principles, as does dynamic data, which is to be made available immediately via APIs (Art. 5(5)). The principles call for *free licences*. Here, we turn to Article 8 on Standard licences, which reads that, for those member states using licences, they shall ensure they are standard and not be subject to conditions unless they are 'objective, proportionate, non-discriminatory, and justified on public interest grounds.' It stops short at specifying a particular open licence framework, such as Creative Commons (CC), despite calls from the academic and re-user community (Janssen & Hugelier, 2013; Pugh, 2019). Instead, the ODD continues to rely on non-binding guidelines on licencing from 2014 (European Commission 2014), which promoted the use of CC licencing. Notably, however, the HVD implementing regulation obliges HVDs be made available for use under CC Public Domain Dedication (CC0), CC BY 4.0 licence or an equivalent or less restrictive open licence (Art. 4). Practical Arrangements in Article 9 are also relevant for the principles because they facilitate the *accessibility* and *findability* of documents. This has largely materialised internationally in the form of national

open data portals, which the European Commission described as akin *'to library catalogues, they contain metadata records of datasets published for re-use, mostly relating to information in the form of raw, numerical data rather than textual documents'* (2023c). The outstanding principle, then, is *non-discrimination*, which has always been a feature of the PSI Directive and can be found in Chapter IV. The most extreme example of alignment, however, is in relation to the HVDs, where all countries must share data according to the open data principles.

Having established the alignment in these specific instances, it is fair to turn our attention to the potential implications for citizen participation. The experiences since 2007 have demonstrated that the predicted 'army' of auditors and a continuously engaged public have not transpired from open data sharing; instead, low-level use is dominated by small groups of intermediaries in the form of highly educated technology and data-savvy experts spread across businesses, a few members of the public, journalists and NGOs (Kennedy, 2018; Worthy, 2015). Data shared through APIs in machine-readable formats, here based on the assumption that information empowers and motivates individuals to exercise choice, has been disputed (Lassinantti, 2019; Worthy, 2015). Simply having access to raw datasets does not automatically lead to civic engagement and participation (Puussaar et al., 2018) because most citizens do not have the requisite skills to find, understand or interpret data. Even when faced with interpretations such as data visualisations from intermediaries, many citizens lack confidence in their ability to make sense of government data (Kennedy, 2018). Studies have also shown that open data portals do not contain the crucial structural and organisational elements necessary to fully support citizen participation (Lnenicka & Nikiforova, 2021).

Although the open data principles are ingrained in the ODD, these alone are unlikely to realise citizen participation at the member-state level. It is in this crux that we see that the importance of situating open data within its broader open government rationale is crucial; implementing the principles outside of this context is likely more conducive to feeding companies operating in the data economy.

5. Conclusion

The ODD is no longer on the periphery of EU policy but instead is a central instrument for the realisation of the European data economy and, therefore, requires interrogation. The present paper traced the origins of the ODD in the distinct movements for open government data and the re-use of PSI. It then proceeded to an analysis that indicates that, although these two approaches may seem like natural bedfellows, given that they share an abiding faith in the power of sharing non-sensitive data, there are fundamental tensions in this marriage. These stem primarily from the fact that the re-use of PSI is firmly rooted in the realisation of the European data economy, while open data is primarily anchored in ideas around government transparency and accountability as a way to facilitate citizen participation. The original largely technical principles of open data are somewhat respected and incorporated into the ODD but are disjointed from the rationale behind them. References to citizen participation in the directive and supporting documentation, while welcome, risk being little more than window dressing.

To be fair, the ODD and its predecessors have been clear on its legal basis. To ingrain the rationale of openness and transparency for citizen participation would be to go beyond its remit. Nor, indeed, is it correct to assign this responsibility to the directive. Nevertheless, there has been legal confusion at the member-state level, as was first identified by Janssen (2011a). The confusion in the PSI Directive was found to work in favour of citizen participation but given the importance of government data to feed the data economy, this is unlikely to still be the case. Moreover, replacing 'PSI' with 'Open Data' in the short title is only likely to exacerbate this confusion.

Bourdieu (2002) wrote that to allow:

> *texts to circulate without their context; they don't carry with them the field of production they come from, and the receivers, themselves integrated in a different*

> *field of production, reinterpret them in accordance with their position in the field of reception.*

The present paper has sought to give context to the original 'texts' that were integrated in the ODD so that the receivers, here in the form of academia, civil society and those tasked with implementing the ODD at the member-state level, can understand the field of production. In doing so, it serves to illustrate the role and scope of the ODD, highlighting potential areas of misunderstandings and confusion so that they can be avoided and/or resolved.

5.1. References

Andrasko, J., & Mesarcik, M. (2018). Quo Vadis open data. *Masaryk University Journal of Law and Technology, 12*(2), 179–220.

Bates, J. (2012). "This is what modern deregulation looks like": Co-optation and contestation in the shaping of the UK's open government data initiative. *Journal of Community Informatics, 8*(2).

Bathelt, H., Cohendet, P., Henn, S., & Simon, L. (Eds.). (2017). *The Elgar companion to innovation and knowledge creation*. Edward Elgar Publishing.

Bourdieu, P. (2002). The social conditions of the international circulation of ideas. *Actes de la Recherche en Sciences Sociales, 145*(5), 3–8.

Cerrillo-i-Martínez, A. (2012). The re-use of public sector information in Europe and its impact on transparency. *European Law Journal, 18*(6), 770–792.

Dalla Corte, L. (2018). Towards open data across the pond. In B. van Loenen, G. Vancauwenberghe, & J. Crompvoets (Eds.), *Open data exposed* (Vol. 30, pp. 11–32). T.M.C. Asser Press.

Davies, T. (2019). *High value datasets: An exploration*. https://www.timdavies.org.uk/2019/08/14/high-value-datasets-an-exploration/

De Filippi, P., & Maurel, L. (2015). The paradoxes of open data and how to get rid of it? Analysing the interplay between open data and sui-generis rights on databases. *International Journal of Law and Information Technology, 23*(1), 1–22.

Deloitte, ODI, The Green Land, & The Lisbon Council. (2020). *Smart 2019/0025 final report impact assessment study on the list of high value datasets to be made available by the member states under the Open Data Directive.* EU Publications Office.

European Commission. (1990). *Guidelines for improving the synergy between the public and private sectors in the information market*. Publications Office.

European Commission. (2010). *A digital agenda for Europe.* COM (2010) 245 final.

European Commission. (2011). *Open data — An engine for innovation, growth and transparent governance.* COM(2011)882.

European Commission. (2014). *Guidelines on recommended standard licences, datasets and charging for the re-use of documents*. 2014/C 240/01.

European Commission. (2018a). *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*. COM (2018), SWD (2018) 127 final.

European Commission. (2018b). *Executive Summary 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,* COM (2018), SWD (2018) 128 final.

European Parliament and Council. (2003). *Directive 2003/98/EC on the re-use of public sector information*. OJ L 245.

European Parliament and the Council. (2013). *Directive 2013/37/EU amending Directive 2003/98/EC on the re-use of public sector information*. OJ L 175.

European Parliament and the Council. (2019). *Directive (EU) 2019/1024 on open data and the re-use of public sector information.* OJ L 172.

European Commission. (2020). *A European strategy for data*. COM (2020) 66 final.

European Parliament and the Council (2022a). *Regulation on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)*.

European Parliament and the Council (2022b). *Regulation on harmonised rules on fair access to and use of data (Data Act)*

European Commission. (2022c). *Commission Implementing Regulation (EU) 2023/138 laying down a list of specific high-value datasets and the arrangements for their publication and re-use C/2022/9562*

European Commission. (2022d). *Annex to the Commission Implementing Regulation laying down a list of specific high-value datasets and the arrangements for their publication and re-use C(2022) 9562*

European Commission. (2023a) *Europa.eu.* https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data

European Commission. (2023b). https://digital-strategy.ec.europa.eu/en/policies/psi-open-data

European Commission. (2023c). https://digital-strategy.ec.europa.eu/en/policies/open-data-portals

G8. (2013). *Open data charter.* https://opendatacharter.net/

Gonzalez-Zapata, F., & Heeks, R. (2015). The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly, 32*(4), 441–452.

Grimmelikhuijsen, S., John, P., Meijer, A., & Worthy, B. (2018). Do freedom of information laws increase transparency of government? A replication of a field experiment. *Journal of Behavioral Public Administration, 1*, 1–10.

Harrison, T. M., & Sayogo, D. S. (2014). Transparency, participation, and accountability practices in open government: A comparative study. *Government Information Quarterly, 31*(4), 513–525.

Jaatinen, T. (2016). The relationship between open data initiatives, privacy, and government transparency: A love triangle? *International Data Privacy Law*, *6*(1), 28.

James, L. (2013). *Defining open data*. Open Knowledge Foundation Blog. https://blog.okfn.org/2013/10/03/defining-open-data/

Janssen, K. (2011a). The role of public sector information in the European market for online content: A never-ending story or a new beginning? *Info, 13*(6), 20–29.

Janssen, K. (2011b). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly, 28*(4), 446–456.

Janssen, K., & Hugelier, S. (2013). Open data as the standard for Europe? A critical analysis of the European Commission's proposal to amend the PSI Directive. *European Journal of Law and Technology, 4*(3).

Kalandides, A. (2018). Citizen participation: Towards a framework for policy assessment. *Journal of Place Management and Development, 11*(2), 152–164.

Kennedy, H. (2018). Living with data: Aligning data studies and data activism through a focus on everyday experiences of datafication. *Krisis : Journal for Contemporary Philosophy*, *2018*(1), 18–30.

Kennedy, H., Poell, T., & van Dijck, J. (2015). Data and agency. *Big Data & Society*, *2*(2)

Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1).

Lassinantti, J. (2019). Re-use of public sector open data. *International Journal of Public Information Systems*, *1*, 29.

Lnenicka, M., & Nikiforova, A. (2021). Transparency-by-design: What is the role of open data portals? *Telematics and Informatics*, *61*, 101605.

Meijer, A. (2009). Understanding modern transparency. *International Review of Administrative Sciences*, *75*(2), 255–269.

Mellouli, S., Luna-Reyes, L. F., & Zhang, J. (2014). Smart government, citizen participation and open data. *Information Polity*, *19*(1–2), 1–4.

Moon, M. J. (2020). Shifting from old open government to new open government: Four critical dimensions and case illustrations. *Public Performance & Management Review*, *43*(3), 535–559.

OECD. (2022). *Open government data*. https://www.oecd.org/digital/digital-government/open-government-data.htm

OKFN. (2010). *The open data handbook.* http://opendatahandbook.org/

OKFN. (2016). *Global Open Data Index*. http://global.census.okfn.org/

OKFN 2019. blog.okfn.org/2019/07/09/missed-opportunities-in-the-eus-revised-open-data-and-re-use-of-public-sector-information-directive/ .

Pollock, R., Lämmerhirt, D., Davies, T., Walker, S., Rubinstein, M., & Perini, F. (2019). Open data around the world: European Union. In *The state of open data: Histories and horizons* (pp. 465–484).

Pugh, S. (2019). *Missed opportunities in the EU's revised open data and re-use of public sector information directive*.

Public.Resource.Org. (2007). *Open government data principles*. https://public.resource.org/8_principles.html

Puussaar, A., Johnson, I. G., Montague, K., James, P., & Wright, P. (2018). Making open data work for civic advocacy. *Proceedings of the ACM on Human-Computer Interaction, 2*(CSCW), 1–20.

Valli Buttow, C., & Weerts, S. (2022). Public sector information in the European Union policy: The misbalance between economy and individuals. *Big Data & Society, 9(2).*

Van den Brink, T. (2017). The impact of EU legislation on national legal systems: Towards a new approach to EU–member state relations. *Cambridge Yearbook of European Legal Studies, 19,* 211–235.

van Eechoud, M., & Janssen, K. (2012). *Rights of access to public sector information* (LAPSI policy recommendation; No. 6). LAPSI.

Wirtz, B. W., Piehler, R., Thomas, M.-J., & Daiser, P. (2016). Resistance of public personnel to open government: A cognitive theory view of implementation barriers towards open government data. *Public Management Review, 18*(9), 1335–1364.

Wirtz, B. W., Weyerer, J. C., & Rösch, M. (2019). Open government and citizen participation: An empirical analysis of citizen expectancy towards open government data. *International Review of Administrative Sciences, 85*(3), 566–586.

Worthy, B. (2010). More open but not more trusted? The effect of the Freedom of Information Act 2000 on the United Kingdom central government. *Governance, 23*(4), 561–582.

Worthy, B. (2015). The impact of open data in the UK: Complex, unpredictable and political. *Public Administration, 93*(3), 788–805.

Table

	Statistics	**Companies**	**Geospatial**
Infrastructural costs	500 k–3 million	10 k–2.5 million	250–300 k
Data transformation costs	30–50 k	4–10 FTEs	100–200 k pa
Operational costs	100–200 k		150–250 k

Table 1: Estimated costs for opening three HVDs. Derived from Deloitte et al., 2020, pp. 24, 198, 304.