Journal of Responsible Technology 16 (2023) 100070

Contents lists available at ScienceDirect

Journal of Responsible Technology

journal homepage: www.sciencedirect.com/journal/journal-of-responsible-technology

Data subject rights as a research methodology: A systematic literature review

Adamu Adamu Habu *, Tristan Henderson

School of Computer Science, University of St Andrews, Jack Cole Building, North Haugh, St Andrews, KY16 9SX, Scotland, United Kingdom

ARTICLE INFO

Keywords:
Data protection
Data subject rights
Research methodology
Systematic literature review

ABSTRACT

Data subject rights provide data controllers with obligations that can help with transparency, giving data subjects some control over their personal data. To date, a growing number of researchers have used these data subject rights as a methodology for data collection in research studies. No one, however, has gathered and analysed different academic research studies that use data subject rights as a methodology for data collection. To this end, we conducted a systematic literature review that searched, compiled, and analysed 32 academic studies that use data subject rights as a data collection method. We find that the right of access is the most commonly-used data subject right by researchers, most studies are interested in measuring data subject rights compliance, and that a variety of difficulties exist in conducting research studies with data subject rights. We conclude that researchers should explore other data subject rights for alternative purposes, ease the process of exercising data subject rights, and improve the scalability of these studies.

Introduction

In today's technological world, our personal data are at high risk of abuse, for instance, from unethical use and sharing of personal data, or privacy violations (Cellan-Jones, 2020; Hill & Mattu, 2018). The Cambridge Analytica scandal exposed how the personal data of millions of Facebook users can be improperly shared and harvested to influence democracy (Confessore, 2018). As a result of such abuses, there is a need to integrate fairness and transparency principles when designing and using technologies to collect, store, process, and/or share people's data. One way to implement these principles is to give citizens some control over who can access, process, and share their data. Many countries have done this through legislating data protection and privacy regulations to protect data subjects in respect to the use of their personal data and its movement (European Union, 2002). An example of such a data protection regime is the European Union's (EU) General Data Protection Regulation (GDPR) (Regulation (EU) 2016). Other countries (or regions) have, or are in, the process of legislating their data protection regulations (Greenleaf, 2021).

To address concerns about unethical use, retention, or sharing of personal data, data protection regulations afford some rights, known as data subject rights, to data subjects. For example, data subjects can use the right of access to obtain from data controllers data held about them, how their data are processed, and how (and to whom) their data are shared (Regulation (EU) 2016), Art 15). Another data subject right, the right to data portability, allows data subjects to obtain and reuse their data across different services (Regulation (EU) 2016), Art 20). Researchers can use data subject rights as a powerful tool for collecting data for studies (Ausloos & Veale, 2021). Consequently, a growing number of researchers use these data subject rights as a methodology for data collection in research studies; for example, studying compliance with data protection regulations (Ausloos & Dewitte, 2018; Mahieu et al., 2018), obtaining sensitive data to study corporate organisations (Uldam, 2016), and auditing pervasive systems (Zwiebelmann & Henderson, 2021). Other groups, such as activists, are using these rights to fight for workers' rights and demand fair processing of their personal data (Open Society Foundations, 2019).

In this paper, we explore how research studies employ data subject rights as a methodology. We term these studies as Data Subject Rights Driven Studies (DSRDS). As these research studies become more popular, it becomes timely to study how they are carried out. To our knowledge, no one has gathered and analysed different academic DSRDS. Our main aim of this paper is to collect and analyse academic DSRDS to identify gaps, shortcomings, and areas for research. We used a commonly-used set of guidelines (Kitchenham et al., 2007) to conduct a systematic literature review (SLR) and search, compile, and analyse

* Corresponding author.
E-mail address: aah5@st-andrews.ac.uk (A.A. Habu).

https://doi.org/10.1016/j.jrt.2023.100070

Available online 12 October 2023

academic DSRDS from six online databases. Our search covered the last 20 years and produced 3410 studies. A total of 32 studies pass our assessment criteria, as detailed in Section 3. From analysing these studies, we make the following contributions:

- We identify the data subjects' rights that researchers use in their studies.
- We categorise the purposes for which these rights are used by researchers and their findings.
- We highlight experiences and challenges associated with exercising these data subject rights in academic research studies.

From our SLR, we discover that DSRDS are hard to conduct in practice. We highlight research methodological challenges, for example, scalability and ethical issues. Our findings reveal a lack of data protection education on the sides of both data controllers and data subjects. As this field is still in its early stage (Ausloos & Veale, 2021), we believe that our review analysis will guide researchers, data subjects, data controllers, system designers, managers, and policymakers in assessing the practical implementation of data protection regulations. Stakeholders will find our review valuable in addressing existing, emerging, and future challenges in designing systems that fulfil these legal obligations and implementing the law in practice.

This paper is laid out as follows. We next briefly describe the history of data protection regulations and data subject rights. In Section 3 we outline how the SLR was designed and conducted. In Section 4 we present the results of our analysis of the 32 primary studies by answering our review questions. In Section 5 we discuss our results and provide some open questions. We discuss our review limitations in Section 6 and conclude.

Background

Data protection regulations

Data protection regulations are legal frameworks designed to protect the personal data and privacy of users. Historically, data protection has a strong connection with privacy (Lukács, 2016; Gellman, 2022). To give an account of data protection to date, we will start by stating some brief historical journeys of privacy. Privacy has evolved as part of fundamental human rights after its definition as "the right to be let alone" by Warren and Brandeis in 1890 (Lukács, 2016; Warren & Brandeis, 1890). In 1948, the Universal Declaration of Human Rights (UDHR) recognised privacy as a human right (Rudgard; Lukács, 2016), followed by the European Convention on Human Rights (ECHR) in 1950 (Rudgard; Lukács, 2016). In the 1970s, development in information society led to the birth of the right to data protection (Lukács, 2016). As a result, many countries, especially in Europe, enacted laws that regulate the use of personal information (Rudgard). In 1973 and 1974, the Council of Europe established a framework to prevent the unfair collection and processing of personal information (Resolutions 73/22 and 74/29) (Rudgard). The idea was to facilitate trade amongst member states. In the early 1980s, the Organisation for Economic Cooperation and Development (OECD) set up guidelines to protect the privacy and the transborder flow of personal data (OECD Guidelines); and the Council of Europe set up another framework to protect individuals from automatic processing (Convention 108) (Rudgard). Convention 108 was the first international treaty on data protection, and both the OECD Guidelines and Convention 108 integrate elements of the Fair Information Practices (FIPs) (Gellman, 2022). In 2018, the Council of Europe and other signatories modernised the Convention 108 treaty, which is now referred to as "Convention 108+". The modernisation was meant to enhance the protection of personal data in this digital age by, amongst others, strengthening the convention's follow-up mechanisms and addressing the emerging privacy challenges posed by the use of new information and communication technologies (De Terwangne, 2021; Council of Europe). Similarly, OECD guidelines were revised in 2013 in response to the tremendous developments since they were adopted (Organisation for Economic Co-operation & Development 2013; Kuschewsky, 2013).

The Fair Information Practices (FIPs), a set of principles developed to safeguard the privacy of personal data, play a vital role in shaping data protection regulations across the globe, to date (Gellman, 2022). FIPs were first codified following proposals from two different committees established by the British and U.S. governments in the 1970s.

Another major shift in European data protection regulation was the 1995 Directive 95/46/EC (the Data Protection Directive or DPD (Directive 1995)). The DPD aims to harmonise data protection legislation within member states. The DPD, however, was not a panacea, as it has its limitations. Van Biemen summarises diverging regulation, weak enforcement, failure to address technological innovations, and internalisation as the main critiques against the DPD (Van Biemen, 2018). To modernise and harmonise data protection across the EU member states, the GDPR was adopted on 14 April 2016 and came into force on 25 May 2018. Elsewhere around the globe, Canada has the Personal Information Protection and Electronic Documents Act (PIPEDA) (Government of Canada 2000), the state of California has the California Consumer Privacy Act (CCPA) (State of California 2018), amongst other data protection frameworks (Greenleaf, 2021). Looking at the historical progression of data protection regulations we can say that social, political, and technological factors have encouraged changes to data protection to date (Van Biemen, 2018).

The GDPR has been called the most remarkable reform in the history of data privacy (Edwards, 2018). The passage of the GDPR into law has changed companies' behaviour toward handling personal data within the EU and beyond (Mahieu et al., 2021). We thus choose to focus on the EU data protection legislative framework in this background chapter of our review and to drive part of the methodology of our SLR (see Section 3.4). For the purposes of this paper, we defined the following terms based on EU law:

- Personal data refers to any information concerning a person who can be identified directly from the information or indirectly in combination with other information (Regulation (EU) 2016, Art 4).
- Data subjects are people who can be recognised directly or indirectly by the data held by controllers (Regulation (EU) 2016, Art 4).
- Data controllers are natural or legal persons which determine the purposes and means of processing personal data (Regulation (EU) 2016, Art 4).
- Data processors are natural or legal persons responsible for processing personal data on behalf of the controller. (Regulation (EU) 2016, Art 4).

Data subject rights

Data subject rights are designed to provide data subjects with knowledge and potential control of their personal data. The GDPR delineates eight data subject rights in chapter III to data subjects and mandates respective obligations to data controllers (and data processors). The GDPR's data subject rights are as follows:

1 The right to be informed (Articles 13 & 14)
2 The right of access (Article 15)
3 The right to rectification (Article 16)
4 The right to erasure/'right to be forgotten' (Article 17)
5 The right to restriction of processing (Article 18)
6 The right to data portability (Article 20)
7 The right to object (Article 21)
8 The right not to be subject to automated decision-making (Article 22)

Giannopoulou et al. categorise the GDPR's data subject rights into two groups according to their mandate, passive and active (Giannopoulou et al., 2022). Passive rights, such as the right to be informed

(Articles 13 & 14), aim to facilitate the exercise of rights. Whereas, active rights, including Articles 15 to 18 and 20 to 22, deal with granting the rights to data subjects. To classify the active rights according to purpose, Articles 15 and 20 focus on transparency, Articles 16 and 17 focus on personal data, and Articles 18, 21, and 22 focus on processing (Giannopoulou et al., 2022).

Even though some of these data subject rights predate the GDPR, they are underused in practice (Giannopoulou et al., 2022; Ausloos & Dewitte, 2018; Van Biemen, 2018). Max Schrems, an Austrian Law student, revolutionised the use of data subject rights in practice (Van Biemen, 2018). In 2011, Schrems filed an access right request to Facebook for his personal data and received more than 1200 pages of PDF. Schrems, grieved by Facebook's action, initiated complaints proceedings before the Irish Data Protection Commissioner (DPC). Schrems' case led to the abolishment of the Safe Harbour agreement and was an eye-opener on Facebook's breaches of European data protection regulations (Ausloos & Dewitte, 2018; Van Biemen, 2018). Additionally, Schrems' case and other peoples' alike have demonstrated that activists, researchers, and other categories of data subjects can use the right of access, the right to data portability, or the right not to be subject to automated decision-making to correct injustice (Giannopoulou et al., 2022), or negotiate better working conditions (Open Society Foundations 2019), or audit pervasive systems (Zwiebelmann & Henderson, 2021). Invoking these rights can be done individually (Wong & Henderson, 2019) or collectively (Mahieu et al., 2018) to achieve targeted objectives. Collectively exercising data subject rights may be more powerful in righting wrongs (Mahieu et al., 2018; Ausloos & Dewitte, 2018).

To exercise data subject rights, three steps are involved (Ausloos & Dewitte, 2018). First, locate the controller that holds a data subject's personal data. Second, filing a request. The data controllers, upon receiving, are duty-bound to fulfil such requests from the data subject (or third party exercising on behalf of the data subject) except in certain circumstances. In cases where the request cannot be fulfilled, the data controller must communicate the reason(s). Third and finally, correspondence with the data controllers and analyses of their responses. In some cases, the data controller will provide no data despite answering the requests, for example, if the data has been deleted or violates some legal arrangements, or if the data controller fails to comply.

While data subject rights are designed to emphasise the rights of data subjects concerning their personal data, academics can use these rights to obtain access to enclosed datasets as part of their research studies (Ausloos & Veale, 2021). Legal, ethical, and methodological challenges can sometimes make these datasets held by data controllers hard to access and scrutinise (Ausloos & Veale, 2021). Thanks to the transparency requirements of the data protection regulations that allow access and study of internal data held by data controllers. Researchers have data subject rights as a legal framework to, amongst others, scrutinise algorithmic systems; identify and combat disinformation; and improve technological system designs. This is done by encouraging participants to collectively exercise their data subject rights in an attempt to answer promising research hypotheses. This legal framework provides valuable avenues for researchers to explore to quench their taste for access to sensitive (and large) dataset (Ausloos & Veale, 2021) and recent development has shown that researchers are increasingly adopting data subject rights as data sources for data-driven research (Wong & Henderson, 2019; Zwiebelmann & Henderson, 2021; Mahieu et al., 2018).

The need for the review

We propose that using data subject rights as a methodology for data collection should be studied extensively to identify gaps, shortcomings, and areas for future research. This is based on two observations: first, as argued by Ausloos & Veale, this research field is still in its infancy stage (Ausloos & Veale, 2021). Second, driven by the passage of the GDPR, companies' behaviour towards data protection has changed, not only in the EU but globally (Mahieu et al., 2021), which makes such a study timely. Having searched the literature, we found no existing reviews on academic DSRDS (Section 3.1 details our search effort). Consequently, we conducted an SLR using a commonly-used set of guidelines (Kitchenham et al., 2007) to collect DSRDS. We analyse how researchers use data subject rights as a methodology for data collection in research studies to identify gaps and challenges.

Review methodology

This section documents how we conducted this SLR, adapting our procedure from the guidelines provided by Kitchenham et al. (Kitchenham et al., 2007).

Planning the review

To plan this SLR, we first drafted a pre-review protocol by writing down research questions and an execution plan. Next, we tried to establish whether a SLR or tertiary review should suit our case (Kitchenham et al., 2007). In doing so, we searched six electronic databases for previous studies in the literature that gather and analyse DSRDS. The searched databases were as follows:

1 ACM Digital library (Association for Computing Machinery 2022)
2 IEEE Xplore library (IEEE 2022)
3 Scopus (Elsevier 2022)
4 Web of Sciences (Web of Science 2022)
5 LawArXiv (arXiv 2022)
6 ProQuest Dissertation & Thesis (ProQuest LLC 2022)

We searched all the databases on 2021–11–11 except 'ProQuest Dissertation & Thesis' which was queried on 2021–11–15. We used the following strings to search for any relevant literature reviews of DSRDS:

("Literature Review" OR "Mapping stud*" OR "Mapping Surv*") AND ("Data Righ*" OR "Right to be informed" OR "Right of access" OR "Right to access" OR "Right to rectification" OR "Right to erasure" OR "Right of erasure" OR "Right to be erased" OR "Right to be forgotten" OR "Right to restric*" OR "Right to data portability" OR "Right to objec*" OR "Right related to automated decisi*" OR "Right to avoid*" OR "Right not to be subject to automated decision-making" OR "Subject access right" OR "Right to withdraw consen*" OR "Right to complai*").

In the end, the search yielded 73 papers (with some duplication) from all six databases queried. We then read each of the studies using Keshav's guidance (Keshav, 2007), and found that none of the studies address our problem. This discovery paves the way for a SLR.

Having concluded that we need to conduct a SLR, the next stage of the review protocol was to develop the search strategy to answer our research questions (Kitchenham et al., 2007). This strategy entails a search string and search databases. For the search string, we used all areas of data subject rights (see Section 3.4), and for the search engines, we used the six online databases listed above.

Research questions

The aim of our SLR is to collect and analyse DSRDS to identify gaps and or opportunities. We therefore formulated the following research questions:

- **RQ1:** Which data subject rights do researchers use in their studies?
- **RQ2:** What purposes are data subject rights used in research studies, and what findings have been made?
- **RQ3:** How are researchers using data subject rights in research studies, and in which jurisdiction are people conducting research studies?

Inclusion/exclusion criteria

We used the following inclusion and exclusion criteria to screen the primary studies generated from our search queries. We intended to focus on academic studies driven by data subject rights, and we used a period of 20 years to cover other data protection regimes.

1 Academic studies from the last 20 years;
2 Research studies that use any data subject right as a methodology for data collection;
3 The expanded version of studies, if several version of a study was archived;
4 Each relevant study was treated differently when several studies were reported.

Exclusion criteria

1 Abstracts, tutorials, presentations (e.g., PowerPoint), errata, letters, retracted documents, or essays;
2 Studies associated with the Freedom of Information (FIO) request;
3 Studies not reported in the English language.

Search strategy

We adopt the search keywords and strings used in this study from data subject rights and their branches as enshrined in the GDPR. To capture all possibilities, we did our best to include alternatives for each area of data subject rights. We cannot, however, rule out the possibility of bias against some legal jurisdictions. We discuss this limitation in Section 6. The following paragraph details our search strings:

("Data righ*" OR "Right to be informed" OR "Right of access" OR "Right to rectification" OR "Right to erasure" OR "Right to be forgotten" OR "Right to restric*" OR "Right to data portability" OR "Right to object" OR "Right related to automated decision-making" OR "Rights in relation to automated decision making and profiling" OR "Right not to be subject to automated decision-making" OR "Subject access request" OR "Right to withdraw consent")

Conducting the review

We made three passes through our search: "First pass" through the metadata, "Second pass" through the abstracts, and then "Third pass" through the full text (Fig. 1). We base our procedure on the work of Meneely et al. (Meneely et al., 2013) and Keshav (Keshav, 2007). Our method differs, however, from Meneely et al. as we did not screen out primary studies using titles and authorship.

First pass

In the first pass, we queried the search strings from Section 3.4 through the metadata in each database. This procedure produced 3410 primary studies. Fig. 1 details the results according to databases.

Second pass

We considered all the 3410 sources returned by the six search databases for the second pass. Unlike Meneely et al. (Meneely et al., 2013), we did not filter out studies based on titles and authorship. Our reasons are two-fold: a researcher may use data subject rights to collect data regardless of their study domain Ausloos and Veale, (2021), and additionally, a title may not describe the methodology used for data collection.

For this pass, we used a spreadsheet to keep track of all the primary studies (see Section 3.7). We then read the abstracts to screen each study. Out of the 3410 sources, we considered only 55 studies for the final pass, representing 1.61 % (38 sources after removing duplicates, as some studies appeared in more than one database).

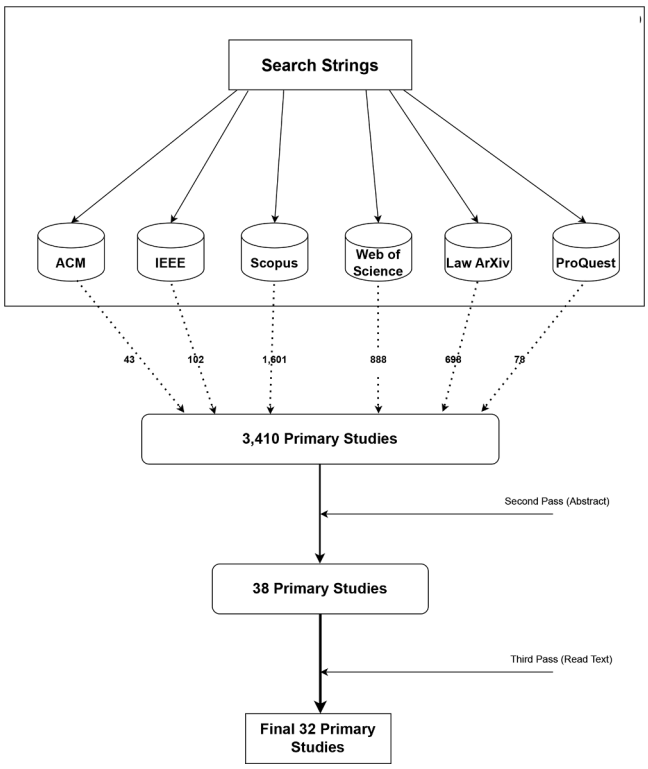

Fig. 1. The number of primary studies extracted from the six databases in each of the three passes of the search process. "First pass" through the metadata, "Second pass" through the abstracts, and then the "Third pass" through the full text. 32 primary studies pass all the SLR assessment criteria.

Third pass

In this pass, we used Keshav's procedure to read the 38 primary studies (Keshav, 2007). Specifically, we read the title, the abstract, the introduction, the conclusion, the Sections headings, and the sub-Sections headings (ignoring everything else). In the end, 32 primary studies pass all of the assessment criteria. In total, we dropped six studies from the second pass. One was a conference paper (shorter version) of another extended journal article. Two were the same work with different titles, a journal publication, and the other a book chapter; we dropped the book chapter. We rejected one study because the data right mentioned in the abstract has no connection to data protection regulations. We removed another because of its focus on FOI requests rather than data subject rights. Finally, we rejected two papers because they discussed data subject rights, but did not use them for data collection.

Quality assessment

We strictly applied our inclusion and exclusion criteria (Section 3.3) in selecting the primary studies. No study was included or excluded because of the quality or otherwise of its publication. This was because of the interdisciplinary nature of this study area: we found studies in Computer Science, Law, Social Sciences, and other knowledge fields repositories. Since quality requirements in Computer science may differ from those in Law, for example, setting metrics may lead to bias. Moreover, as the targeted area for this review is still in its infancy stage (Ausloos & Veale, 2021), we believe that including all valid research regardless of the stage of the work is important.

Data extraction

Following Kitchenham et al.'s guidelines (Kitchenham et al., 2007) and the work of Elhabbash et al. (Elhabbash et al., 2019), we extracted data using a table format showing data items from the primary studies

and their relevancy to the SLR research questions/objectives. Table 1 presents the data items we extracted and their connection to the SLR research questions. We designed the extraction table during the protocol design stage; this helps to minimize bias (Kitchenham et al., 2007).

Data synthesis

After the data extraction in Section 3.7, we coded the primary studies and tabulated the data from the primary studies in tables based on the research questions and other parameters, for example, the number of data subject requests sent per primary study. For some of our analysis, we sourced data from the meta-analysis by Norris and L'Hoiry (Norris & L'Hoiry, 2017). We grouped common themes for analysis. This procedure simplified the synthesis and the statistical analysis. Our synthesis and analysis approach was targeted to answer the research questions (Elhabbash et al., 2019).

Results

In this section, we present the results of our analysis framed around our research questions from Section 3.2.

Primary sources

Table 2 categorises publication venues for the 32 primary studies, using the Journal Citation Reports (Clarivate 2023), Web of Science (WoS Research Team, Web of Science Journal (WoS) Journal Info 2023), and ACM Computing Classification System (Association for Computing Machinery 2023). 17 (just over 50 %) of the primary studies were in Law, with a variety of papers in computer science and social sciences. This shows that research using data subject rights is multidisciplinary. But the high concentration of papers in the law category could be because data subject rights is a legal framework.

For the publication types, we see in Table 3 that the majority of the studies are articles in journals (13), followed by book chapters (11), and then conference proceedings (8). The high number of book chapters in this review is because of the work of Norris et al. (Norris et al., 2017); we treated each chapter in the book as an individual study.

Despite considering the period from 2001 to 2021, almost half (47 %) of the primary studies were published between 2018 and 2021 (see Fig. 2). This suggests that after the GDPR, more people are studying data subject rights. In addition, slightly over half of the papers in the 2018 to 2021 category are conference papers, and there were no technical reports in all the primary studies. This supports the claim that this subject area is still in its infancy stage due to the presence of many conference papers and no technical reports (Ausloos & Veale, 2021; Elhabbash et al., 2019). Conferences enable people to share observations, ideas, and solutions, and a transition between research and practice is evident from the presence of technical reports (Elhabbash et al., 2019).

Table 1
The data items extracted from each of the primary studies and their relevance to the study.

Data item	Description	Relevant to Study
Key	Key referencing the study	Documentation
Title	Primary study title	Documentation
Date	Year of publication	Demographics
Research question(s)	Research question/issues that needs to be identified	RQ1 & RQ2
Methodology	Method employed, how & why was the method used by the study	RQ1 & RQ3
Data subject rights	Data subject right used in the study	RQ1
GDPR status	Whether the study was conducted prior to GDPR Effective Date	RQ3
Scope	Study population and legal framework used	RQ2 & RQ3
Result	What the study found	RQ2
Open questions	Possible open questions to be explored	Discussion & open question Section
Limitations	Study limitations	SLR analysis
Future work	Possible future work	SLR analysis
Conclusion	Study conclusion	RQ2

Table 2
The primary studies according to publication venues for the 32 selected primary studies. The Law category has the highest number of primary studies (17).

Publication types	Number of studies
Communication	4
Computer science information systems	5
Information and library science	1
Law	17
Security and privacy	2
Social and professional topics	1
Social science and interdisciplinary	1
Urban	1
Total	32

Table 3
The 32 selected primary studies according to publication types. Most of the primary studies are journal publications.

Publication types	Number of studies
Book chapters	11
Conference proceedings	8
Journals	13
Total	32

Number of requests

75 % of the 32 primary studies sent fewer or 40 data subject requests, and a total average of 39 requests, as evidenced in Fig. 3. Wong and Henderson have the highest number of requests, with 230 data portability requests sent by the authors (Wong & Henderson, 2019). We attribute this low number of requests to the fact that data subject's rights are individual rights, which means that researchers have to recruit and guide participants to conduct such studies (Asghari et al., 2021).

Data subject rights in studies

RQ1: Which data subject rights do researchers use in their studies?

We are interested in understanding the data subject rights used by researchers. The most commonly used data subject right by researchers is the right of access: 29 of the 32 studies used this right as a methodology for data collection. One of those 29 studies used the right of access and the right to data portability in the same study (Sørum & Presthus, 2020). Relying on the GDPR, the right of access (Regulation (EU) 2016), Art 15) and the right to data portability (Regulation (EU) 2016), Art 20) are somehow related, as they both give data subject access to their personal data. The remaining three other studies used the right to data portability.

In answering our research question **RQ1**, our review shows that the right of access is the data subject right used in more than 90 % of the research studies.

Aims of the studies

RQ2: What purposes are data subject rights used in research studies, and what findings have been made?

First, we reviewed the reasons various researchers conducted research in the 32 primary studies. We find that the majority of the primary studies are interested in examining legal compliance and usability features.

Researchers have different reasons for conducting studies with data

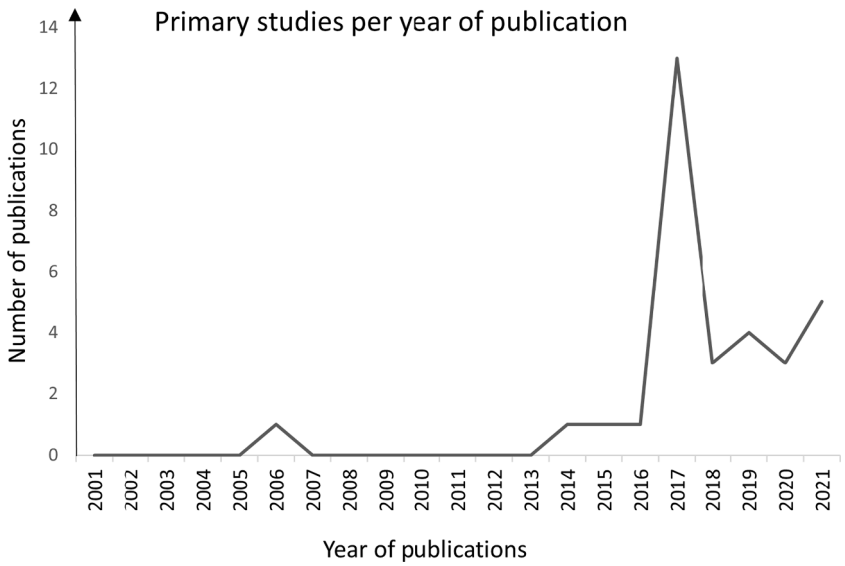

Fig. 2. The primary studies according to years of publication. The popularity of the studies after 2016 indicates how the GDPR influenced DSRDS.

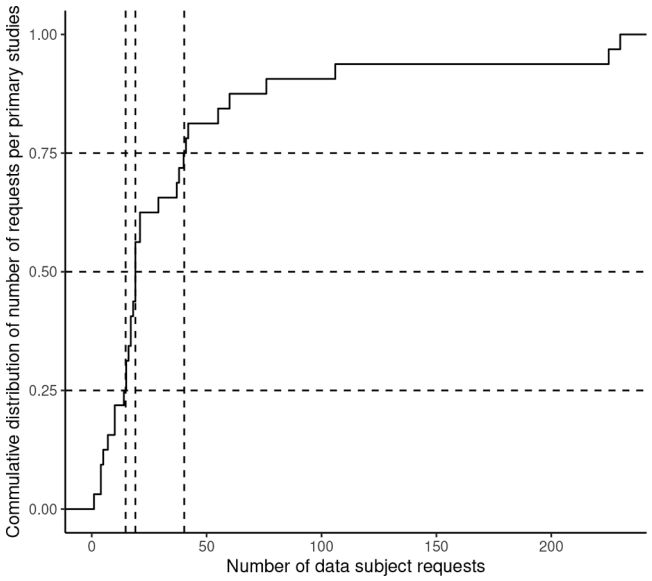

Fig. 3. The number of data subject requests sent in each of the 32 primary studies. The midpoint of the observed data distribution is 19. This suggests a scalability limitation in the number of data subject requests for DSRDS.

subject rights as a methodology for data collection. In the ten EU countries studied in the book, The Unaccountable State of Surveillance, the researchers were interested in testing how easy or difficult for a data subject located in a country of study to exercise their right of access (Norris et al., 2017). The studies monitor data protection regulation compliance and usability features. Mahieu et al. follow the trend of monitoring legal compliance by collectively examining how the right of access meets its transparency requirements (Mahieu et al., 2018). Ausloos and Dewitte also investigated how data controllers accommodate the right of access (Ausloos & Dewitte, 2018). Raento assessed how easy it is for data subjects to get their data and information on data processing using the right of access (Raento, 2006). The study further explored how easy it is for the data subject to verify the correctness of the response received following the requests or how the system supports transparency requirements. Turner et al. conducted an empirical study to test the exercisability of the right to data portability in the context of IoT and the information provided to simplify how data subjects can exercise their data subject rights (Turner et al., 2021). These 14 research studies measure compliance with legal requirements and usability.

Ten other studies were also interested in examining legal compliance. Veale et al. attempted to use the right of access and gain access to voice and audio data (Veale et al., 2018). Bier et al. set up a quantitative and qualitative study with a sample of 612 companies to evaluate the transfer of personal data, compliance with transparency requirements on commercial emails using honeypots, and compliance with the right of access requirements (Bier et al., 2017). Rothmann examined how the right of access is implemented in practice using experience on how data controllers react to subject access requests (Rothmann, 2017). Urban et al. evaluated how companies respond to the right to access and the right to portability requests (Urban et al., 2019). Tolsdorf et al. examined how privacy dashboards comply with the provision of Article 15 (Regulation (EU) 2016), Art 15) of the GDPR – the right of access (Tolsdorf et al., 2021). Spiller discussed his experience in an attempt to exercise the right of access in a UK city (Spiller, 2016). Using two EU countries as case studies, Galetta et al. examined the right of access from a theoretical perspective and its implementation in practice (Galetta et al., 2016). Sørum and Presthus investigated the practical exercisability of GDPR's Article 15 (Regulation (EU) 2016), Art 15) - the right of access and Article 20 (Regulation (EU) 2016), Art 20) - the right to data portability in information systems (Sørum & Presthus, 2020). Kr̈oger et al. examined how service providers complied with subject access requests using an undercover field study over four years (Kr̈oger et al., 2020). Finally, Wong and Henderson investigated how easy a data subject can exercise the right to data portability and the file format returned by the data controller by making 230 real-world portability requests (Wong & Henderson, 2019). Wong and Henderson's study focused on the practical implementation of the law (Wong & Henderson, 2019).

Four studies used data subject rights to collect data from data controllers for various reasons. Zwiebelmann and Henderson explored whether the right to data portability could be used to audit pervasive systems externally (Zwiebelmann & Henderson, 2021). Uldam (Uldam, 2016) and Uldam (Uldam, 2018) used the right of access to collect files from cor-porate organisations for monitoring of activists' activities. Similarly, Ebbers et al. used the right of access to retrieve data from vehicle manufacturers to reconstruct drivers' activities (Ebbers et al., 2021).

Two studies examined the security of data subject rights. Cagnazzo et al. created an attack model, GDPiRate (Cagnazzo et al., 2019), which sends subject access requests using a spoofed recipient's address.

Similarly, Martino et al. attempted to gain access to citizens' personal data without their consent using the right of access (Martino et al., 2019). This was done by forging authentication credentials or by impersonating publicly available information.

Researchers in the remaining three studies have different reasons for their studies. Rene et al. conducted a longitudinal study (before and after the GDPR) to examine the global effect of the GDPR on people who are not EU citizens (Mahieu et al., 2021). Rene et al. assessed the right of access using the EU and Canadian residents (Mahieu et al., 2021). Veys et al. examined the usability features of data download tools provided by data controllers (Veys et al., 2021). Data controllers provide download tools to meet their data protection obligations under the right of access and the right to data portability.

Summary of findings

In this section we report some of the findings of the primary studies. We find that researchers experience both, in the words of Norris and L'Hoiry, 'facilitative' and 'restrictive' experiences (Norris & L'Hoiry, 2017). For a facilitative experience, a data controller's practices, behaviours, and administrative policies are subjectively deemed satisfactory; otherwise, we say the experience is restrictive (Norris & L'Hoiry, 2017). A data controller can exhibit one or both practices.

We first look at studies that examined possible attacks on data subjects' privacy. Martino et al. (Martino et al., 2019) and Cagnazzo et al. (Cagnazzo et al., 2019) described a procedure that could use data subject rights for criminal conduct. Both studies used different techniques and attempted to obtain the personal data of individuals without their knowledge and consent. In both studies, the researchers have access to leaked personal data. It is, however, imperative that data controllers exercise caution when handling data subject rights requests.

In four studies, researchers successfully obtain data from data controllers for their analysis. Uldam (Uldam, 2016), Uldam (Uldam, 2018), Ebbers et al. (Ebbers et al., 2021), and Zwiebelmann and Henderson (Zwiebelmann & Henderson, 2021) obtain data from data controllers using the right of access and the right to data portability. Zwiebelmann and Henderson, however, fail to externally audit pervasive systems due to inaccuracies in the reported data (Zwiebelmann & Henderson, 2021). On the other hand, Veale et al. attempted to obtain a copy of voice data from a data controller, Apple but was unsuccessful (Veale et al., 2018). Apple refused the request citing an unsubstantiated claim of privacy by design (Veale et al., 2018). This trend of non-compliance with the legal requirements is observed in most studies that attempted to measure legal compliance (Norris et al., 2017). Though, some data controllers report good practices (Norris et al., 2017).

In reporting the findings of some studies that examined data protection compliance, we start with the study of Tolsdorf et al. (Tolsdorf et al., 2021). The study finds that most privacy dashboards and data exports lack vital information regarding the right of access. And there is a contradiction between the information in the privacy statement and the actual data export. On the positive side, however, data subjects' personal data are accessible, complete, and editable. Galetta et al. find a significant difference in the theoretical perspectives of exercising access rights in Italy and Belgium (Galetta et al., 2016). For example, theoretically, Italian law permits data subjects to access personal data processed by the police directly. While in Belgium police files are accessed indirectly. The EU harmonisation initiatives do not affect the exercise of the right of access. In addition, the study encounters no noticeable difference between the legal regime in practice in both countries. Urban et al. observe that data controllers offer an easy way to access personal data (Urban et al., 2019). They fail to provide the required data, however, in most cases. 58 % of data controllers do not provide the necessary information within the time limit stipulated by the GDPR. Data controllers provide obstacles to data subjects, such as obtaining signed affidavits and copies of ID cards (Urban et al., 2019). The data from the data controllers are heterogeneous. Though some facilitative practices have been recorded in these three studies, most findings are worrisome.

To report other studies that test data protection regulations compliance, we look at the ten EU countries' studies (Norris et al., 2017). For these ten studies, we use the meta-analysis by Norris and L'Hoiry (Norris & L'Hoiry, 2017). In locating data controllers, results show an overall success rate of 80 %. Researchers attempted to identify data controllers using three mediums, the web (63 %), phone calls (27 %), and in-person visitation (10 %). Email contacts have the largest share of the web medium of locating data controllers (Norris & L'Hoiry, 2017). Norris and L'Hoiry points out widespread restrictive practices in compliance with data protection regulations, as the process was terminated at the very beginning in a fifth of all cases, due to a failure to locate data controllers (Norris & L'Hoiry, 2017). The procedure for submitting access requests is complex and entails several correspondences. The average number of back-and-forth correspondence in Belgium and Spain is 3.1; in Austria, 1.3. Overall, results show the average number of back-and-forth correspondence before submitting subject access requests in the ten studies is 2.15 (Norris & L'Hoiry, 2017). For the data controller's responses to the subject access requests, in all cases, the study records 43 % positive outcomes, i.e., answering the researcher's query adequately (Norris & L'Hoiry, 2017). On the positive side, some data controllers, however, performed excellent facilitative practices. Some of the Data Protection Authorities (DPAs) helped to resolve complaints. The UK DPA resolved all complaints filed; the Spanish DPA resolved all but one of the 14 complaints. DPAs in Austria, Hungary, and Slovakia failed to resolve, successfully any of the complaints filed (Norris & L'Hoiry, 2017).

Other studies that look at legal compliance, as well as usability features, observe a general negative trend in the implementation of these regulations. Raento's study reports difficulty in the procedure adopted by data controllers in exercising the right of access (Raento, 2006). Similarly, the study finds that it is not easy to gain access and verify the correctness and or completeness of the data controller's responses. Ausloos and Dewitte find that the current approach to exercising the data subject's right of access is mired with a lack of awareness, organization, motivation, and harmonization (Ausloos & Dewitte, 2018). Turner et al. investigated the exercisability of Article 20 of GDPR (Regulation (EU) 2016), Art 20) in the context of consumer IoT devices (Turner et al., 2021). The study concludes that the right to data portability is not meaningfully explained to users and is not yet exercisable in the IoT world. In an attempt to measure the collective effectiveness of the right of access, Mahieu et al. engaged participants to evaluate the responses from data controllers, following the right of access requests (Mahieu et al., 2018). The study finds that noncompliance with the data protection regulation is widespread, and participants are unhappy with the practical implementation of the transparency requirements. Additionally, the study observes that citizens rarely exercise the right of access. Veys et al. explored the design, formats, and tools for data download provided by some data controllers in an attempt to fulfil their legal requirements (Veys et al., 2021). The study reports that generally, participants are not happy with either the format or content (or both) of their data download tools. They suggest improving the filtration, visualisation, and summarisation capabilities of the provided data download tools for higher usability and usefulness.

Other studies recorded both restrictive and facilitative experiences. In a longitudinal study conducted by Kröger et al., before and after the coming into force of the GDPR, 53 % of the examined vendors gave an acceptable response (Kröger et al., 2020). The result plummets to 15 % and 41 %, however, in the second and third rounds, respectively. Responses to queries about third-party sharing are not helpful. Bier et al. study finds that most companies do not transfer personal data without consent (Bier et al., 2017). The right of access requests reports restrictive responses as the study reports insufficient information. On the positive side, 85 % of data controllers send responses on time. For Sørum and Presthus, most data controllers respond within the legal time (Sørum & Presthus, 2020). Companies are better compliant with Article 20

(Regulation (EU) 2016), Art 20) than Article 15 (Regulation (EU) 2016), Art 15) of the GDPR. On the negative side, the study finds that data controllers fail to differentiate between the two data subject's requests. Other findings in that study reveal a variation amongst data controllers' attitudes, in terms of response time, differences in file format and content received, quality of feedback, and how they handle the requests. In Wong and Henderson's study, the findings corroborated the difficulty in exercising data subject rights (Wong & Henderson, 2019). 74.8 % of the right to data portability was completed and not all file formats meet the GDPR requirements. Like Sørum and Presthus's study (Sørum & Presthus, 2020), Wong and Henderson find that some data controllers fail to differentiate between data subject's rights (Wong & Henderson, 2019). Finally, Mahieu et al.'s study finds that slightly more than half reported facilitative responses and that the GDPR unilaterally changes companies' behaviour globally (Mahieu et al., 2021).

Our analysis shows that CCTV footage is challenging to obtain from data controllers. This is the case in the ten EU country studies (Norris et al., 2017), data controllers blocked requests using several denial strategies. A partial reading of the law; and a lack of technical and legal competence are examples of denial strategies to block such requests (Norris et al., 2017). This denial trend continues in other CCTV studies. Rothmann attempted to obtain video footage in 29 locations but successfully obtain a satisfactory response in only two and an incomplete answer in four (Rothmann, 2017). Spiller's study suffers a similar fate, as only six out of the 17 requests are successful (Spiller, 2016). Spiller points to the need for a more robust policy for handling data subjects' right of access requests (Spiller, 2016).

Accordingly, from our review findings, we suggest that the exercising of these data subject rights still requires much work going on the general level of negative trends observed in terms of the data controller's behaviour towards responding to data subject requests.

Research evaluation methods

RQ3: How are researchers using data subject rights in research studies, and in which jurisdiction are people conducting research studies?

In this part, we study the methods employed by researchers in the selected primary studies. We categorise the process of using data subject rights in research studies for data collection into three steps (Ausloos & Dewitte, 2018):

1. Identifying the data controller; where and how to send the data subject requests.
2. Submit the data subject requests to the data controller.
3. Back-and-forth correspondence with the data controller; data controller's response analysis; and correspondence with the DPAs, if there is a need.

We use these steps to characterise the primary studies research evaluation methodology. In this Section, we exclude papers that did not report their methodology (Uldam, 2016; Uldam, 2018; Ebbers et al., 2021; Veale et al., 2018).

Identifying data controllers

Researchers used three media to identify data controllers: online, including email; telephone contact; and in-person visitation. The ten EU countries' (Norris et al., 2017) and Galetta et al. (Galetta et al., 2016) studies attempted to locate data controllers using these three media. Meta-analysis for the ten EU countries reports that the studies successfully identify 262/327 data controllers; the majority through the web, 63 %; telephone, 27 %; and in-person, 10 % (Norris & L'Hoiry, 2017). Eleven primary studies report that they evaluated the privacy policy page of websites to locate the data controllers. In addition to the privacy policy page, one of the 11 studies searched for the terms and conditions page. Two of the 11 studies reported that they used other online media and the privacy policy page.

In the case of CCTV, the in-person visit to the camera site was the first contact. The signage may not provide useful information or provide misleading information, however. In that situation, the researchers attempted to search websites or use other media like email or phone calls before locating the data controllers. Other studies did not report how they identify data controllers

Overall, most researchers used online media to identify data controllers. With this result in mind, data subjects with limited or no access to the internet are disenfranchised in exercising their data subject rights, as the process terminates at the very beginning (Norris & L'Hoiry, 2017).

Submitting data subject requests

The majority of the studies used more than one medium to send requests. Several media were used in studies by Norris et al. (Norris et al., 2017) and Galetta et al. (Galetta et al., 2016) to submit requests: email; mail post; online forms; and fax machines. The process was marred with a lot of correspondence, combining several media in some cases.

Table 4 shows different media used by data subjects to send requests in 15 primary studies. In the table, we exclude the ten EU country studies, Galetta et al., the studies listed in the sub-Section 4.6, and two others that did state that the access request letter was sent to a provided address but did not report the medium of communication. Our results show that email is the most widely used medium for sending data subject requests.

Data controllers' responses

Data controllers responded through different media: email; mail post; online forms; fax machines; and DPAs. Some data controllers contacted the data subject after receiving the request to acknowledge the request recipient, others to seek clarification, others for additional information, and others to provide answers to the request. Norris and L'Hoiry, in their meta-analysis of the ten EU studies, and the study of Galetta et al. (Galetta et al., 2016), note that the process of exercising the right of access in those studies was complex to handle as researchers experience several restrictive, facilitative and or both practices (Norris & L'Hoiry, 2017). The process entails several back-and-forth correspondences (an average of 2.15) to receive personal data (Norris & L'Hoiry, 2017). In summary, data controllers contacted researchers in these 11 studies through different media: phone; email; postal mail; download tools; and DPAs.

Table 5 shows different media used by data controllers to establish contact with data subjects in 15 primary studies. In the table, we exclude the ten EU country studies, Galetta et al., the studies listed in the sub-Section 4.6, and two others that did not state how they receive responses from data controllers.

Legal framework

To answer the last part of **RQ3**, we provide descriptive statistics of the legal frameworks used by the researchers in sending the data subject requests. The 32 studies employed four data protection regimes, the DPD, the GDPR, the CCPA, and the PIPEDA. The DPD mandates member states to enact data protection regulations to protect citizens. We classify all requests under the scope of EU member states national laws that derive from the implementation of the DPD as part of the DPD. The

Table 4
The various media used by data subjects to contact data controllers in 15 primary studies.

Medium of communication	Number of studies
Email	12
Postal mail	5
Online forms	10
Fax	1
Other web forms	4

Table 5
The various media used by data controllers to contact data subjects in 15 primary studies.

Medium of communication	Number of studies
Email	13
Telephone	3
Postal mail	8
Download tools	6
Other web forms	9

majority of the studies used the DPD (22 primary studies), then the GDPR (12 primary studies); CCPA (one primary study); and PIPEDA (one primary study). The high number of DPD frameworks is attributed to the work of Norris et al. (Norris et al., 2017); we treated each chapter in the book differently. It is not uncommon for some studies to use more than one legal framework. One study used the GDPR and CCPA. One study used the DPD and GDPR. One study used three frameworks (DPD, GDPR, and PIPEDA). While the GDPR may have triggered most studies, some studies were conducted before the GDPR became effective, so the DPD framework was applied.

Discussion and open questions

In this section, we discuss our results and the primary studies, and identify some open questions.

Discussion of results

The right of access is the most commonly-used data subject right amongst researchers. From our analysis, three reasons explain researchers' interest in the right of access. First, the right of access is the primary driver necessary to enable data subjects to exercise other data subject rights (Norris & L'Hoiry, 2017). Most of the researchers in the 32 primary studies are interested in this right. Second, of all the data subject rights, the right of access allows easy monitoring of compliance. It is easier to measure compliance with data minimization, purpose limitation, accuracy, and storage limitations using the access right (Ausloos & Dewitte, 2018). Compliance with passive data subject rights (Giannopoulou et al., 2022), GDPR's Articles 13 (Regulation (EU) 2016), Art 13) and (Regulation (EU) 2016), Art 14) the right to be informed, can easily be monitored using the right of access (Raento, 2006). Third is the existence of relevant studies in the literature to motivate researchers. Mahieu et al. investigated the global effect of the GDPR by comparing existing data sets in the literature (Mahieu et al., 2021).

We also find that some studies conducted an undercover investigation. In one case, a longitudinal study covering three repeated rounds of data subject requests to data controllers (Kŕöger et al., 2020). The practice of sending multiple unfounded requests to data controllers deviates from the intention of data protection regulations (Regulation (EU) 2016), Art 12). It is common knowledge that data controllers spend time and resources to answer data subject requests. As such, these studies may overburden the data controllers without their consent. The result of that study shows how data controllers exhibit some restrictive behaviour toward the data subject. The data controller gave an impolite response, preferring to get rid of the data subjects than bear the excessive demand (Kŕöger et al., 2020). The effect of these undercover studies is that data controllers may be exhausted and discouraged from fulfiling their legal obligations, even when the requests come from legitimate citizens. Additionally, this may cause reputation damage to researchers if things go wrong. For example, the Princeton-Radboud Study on Privacy Law Implementation was stopped, and the researchers had to apologise (Mayer, 2021). In that study, researchers sent a bunch of unfounded requests, and the institutional review board did not properly evaluate the study to determine human subject involvement. Another issue raised from these undercover studies is the ethical issue of making the data controller waste excessive resources without consent.

Another insight from our review is that DSRDS are hard to conduct. Data subject rights are individual rights, and the studies may involve the personal data of participants. Researchers have to recruit, train and guide participants to conduct such studies (Asghari et al., 2021), in addition to ethical and privacy considerations. Our SLR shows research scalability limitations. Out of the 32 primary studies in our SLR, only three exceeded a hundred requests. In all three cases, fewer than 10 participants sent the requests in each case. A single data subject sent all the 230 requests in the study with the maximum number of requests (Wong & Henderson, 2019). Asghari et al. describe a novel method of delegating the right of access that could improve the scalability of research studies (Asghari et al., 2021). Delegation entails the exercise of data subject rights using mandates by third-party on behalf of data subject (Giannopoulou et al., 2022). This idea of mandate delegation is not explicit in the GDPR; its interpretation can be ambiguous (Giannopoulou et al., 2022). Despite the silence of the GDPR about mandate delegation, it is important to note that data subject rights can be mandated to a third party (Giannopoulou et al., 2022). Future studies could incorporate the idea of delegation into Community science/Citizen science to study scalability in a wider domain Vohland et al. (2021).

Another question that arises from our review is whether researchers can act in the role of data controllers. In one study (Mahieu et al., 2018), researchers recruited participants and as part of the study design, were asked to submit data subject requests to data controllers. The response received was shared with the researchers for analysis. In this case, the researchers are acting on the role of data controllers. There are, however, some concerns associated with this. For example, complying with Article 6 of the GDPR (or its equivalent in other jurisdictions) on further data analysis (Regulation (EU) 2016), Art 6).

From the results of the primary studies selected for this review, we observe a lack of data protection knowledge, expertise, and awareness amongst the representatives of the data controllers. In one such case, the data controller's representative erroneously referred the researcher to the police to submit a data subject request (Bellanova et al., 2017). In another case, the officer was polite and willing to help but stated that they had never received such type of a request before (Szekely & Vissy, 2017). These restrictive experiences, however, will discourage data subjects from participating in research studies or exercising their data subject rights. Literature shows that data subjects rarely exercise their data subject rights (Mahieu et al., 2018). The unwillingness of citizens, whom the data protection regulations protect, to exercise their data subject rights has a connection to the failure of data controllers to provide appropriate organisational mechanisms for implementing the data protection regulations in practice (Norris et al., 2017). Our SLR analysis affirms this position. Having said that, educating the data controller's members of staff will be highly valuable to the success of data protection. In the current situation, disclosure of personal data seems to rely heavily on the willingness and or training of data controller's staff members to handle the request rather than the legal rights of the subjects (Fonio & Ceresa, 2017; Galetta et al., 2016). Moreover, educating data subjects and data controllers on the importance of data protection regulations may simplify the process of exercising data subject rights.

Another issue we note in this SLR is the completeness of responses. Assessing responses from data controllers require legal and technical expertise. There are instances where the data controller can answer the requests without providing data. Zwiebelmann and Henderson attempted to audit pervasive systems using ground-truth data versus the data held by the data controllers (Zwiebelmann & Henderson, 2021). The study was, however, unable to successfully audit the system due to insufficient data from the data controllers. Some studies uncover how data controllers use legal interpretation as a denial strategy to deny requests (Rothmann, 2017; Norris et al., 2017). Others reveal how the controllers fail to provide answers to third-party sharing and automated decision-making questions (Fonio & Ceresa, 2017; Von Laufenberg, 2017). In one case, the DPA sided with the data controller to deny access

to the necessary data. After exchanges of correspondence with the researchers, the DPA rescinded its earlier position (L'Hoiry & Norris, 2017). Taking these experiences into account, how do we know that the response from the data controllers is satisfactory and complete?

We also note the usefulness and accessibility of responses from the data controllers. By usefulness, we here mean how the researcher (or data subject) can comprehend and use the data, that is, the content provided by data controllers. Is the response in an acceptable format that can ease comprehension and usage? In some instances, the data controllers provided the data in printable format (Krieger-Lamina, 2017), which raises the question of how to comprehend and use the data. Whether the data controllers provide the data in a structured or machine-readable form (Wong & Henderson, 2019)? As regards accessibility, we have to talk about the method of communication. Some data controllers attempted to respond in a language foreign to the requester (Bellanova et al., 2017; Szekely & Vissy, 2017; Fonio & Ceresa, 2017). This behaviour will certainly restrict data subjects from interpreting the correspondence. Another issue regarding accessibility is how some data controllers use some security measures to protect the data and instead impose some extra burden on the requester. In one case, the controller provided a password-protected response with the wrong password (Bellanova et al., 2017). The researcher could not access the document until after investing more effort and almost two months delay. While protecting the data is impressive, the data controllers must find usable ways to secure the data without burdening the requester.

Our results show that email correspondence is the most widely used medium for contacting data subjects following data subjects' requests. Email may have advantages, but it can lead to security breaches. Wong and Henderson lost two email correspondences in transit (Wong & Henderson, 2019). Further, Martino et al. demonstrate how an adversary can use an email to compromise the data of a target individual (Martino et al., 2019). Aside from security breaches, email correspondence can easily disenfranchise data subjects with low or no internet access. Data protection regulations aim to protect all categories of citizens. Data controllers should research and use appropriate and user-friendly communication channels accessible to data subjects.

Finally, we note that some data controllers complicate the exercise of data subjects' rights. Data subjects may find difficulty interpreting the procedure for submitting data subject requests. For example, following data subject requests, some data controllers responded by requesting additional documents (Bellanova et al., 2017). The documents are not part of the procedure for exercising data subject rights authored by the data controller. After sending the said documents, some data controllers argue that is the starting time for the 30 days counting (in the case of the GDPR). This practice will make the process cumbersome by not specifying all the required documents in advance.

Open questions and future work

We observe that the majority of the data subjects making requests in our reviewed DSRDS are experts in their fields. At a minimum, they understand the data protection regulations, including but not limited to the data subject rights and what to expect from the data controllers. In some cases, experts, for example, privacy experts or lawyers, assisted researchers in the process. On the other hand, data protection regulations aim to protect the rights of all citizens regardless of their background. Findings from this SLR show difficulty in exercising data subject rights and require legal and technical skills beyond average persons. The procedure entails back-and-forth correspondence, patience, perseverance, and, in some cases, resources. Our SLR analysis was on requests submitted by experts in the majority of cases. A layperson may face various challenges, starting from how to read and interpret the data protection statute, how to locate the data controller, how to submit the requests, when to expect the response, what to expect from the data controllers, how to further correspond with the data controllers and finally how to seek redress (if need be). In several instances, researchers from our primary studies (for example, Veale et al. (Veale et al., 2018)) fault the legal interpretation of the data protection laws by the data controllers and or the DPAs. They argue that the data controllers use a vague interpretation of the laws as denial strategies. Placing a layperson in the position of dealing with the challenges of data subject requests (as outlined above in this paragraph) may be disastrous, however.

We observe that none of the 32 studies focused on the views of data controllers concerning the data subject requests. As far as data subject rights are concerned, there are three key actors: the data controllers, who are obliged with the responsibility of fulfilling the data subject requests; the data subjects, whom the laws aim to protect; and the DPAs, who are vested with the powers to ensure compliance with the legal provision. The studies that we reviewed were interested in data subjects and DPAs; it may be worthwhile to study the data controller's sides concerning the data subject's rights implementation.

Limitations

The biggest limitation of this review study is that we considered only academic studies. Journalists, activists, and other categories of people are doing excellent work with data subject rights. The cancellation of the Safe Harbour agreement was a result of the work of an activist (Ausloos & Dewitte, 2018; Van Biemen, 2018) Another limitation is that our exclusion criteria discarded studies not reported in English. Further, we restricted our primary study selection to a search string that was inclined toward the GDPR and applied to some selected online databases. The search string may not favour all case scenarios as other legal frameworks may have different wording. For example, the recent Kingdom of Saudi Arabia's Personal Data Protection Law (PDPL) has a wording for "The right to knowledge" which our search string may not recognise. Therefore, we cannot rule out the possibility of bias against certain jurisdictions. Finally, researching with data subject rights is multidisciplinary; we chose six different online databases, but these may have been biased towards particular disciplines.

Conclusion

To promote transparency, data protection regulations grant data subjects some rights, which are referred to as data subject rights. A growing number of researchers are using these rights as a tool for studies (Ausloos & Veale, 2021). Exercising data subject rights in research studies (and in practice) is cumbersome and requires patience, perseverance, legal knowledge, technical skills, and resources (Norris & L'Hoiry, 2017). Data subjects interested in exercising these rights have to understand the legal statutes, how to exercise the rights, and what to expect in return. In this paper, we have conducted a systematic literature review that analyses 32 academic studies that use data subject rights as a data collection methodology. From our results, we find that the right of access is the most widely used data subject right in research studies; accordingly, there is a need for researchers to explore other data subject rights. Additionally, we find a variety of difficulties in using data subject rights in research studies. From our analysis, we identify scalability; ethical and privacy concerns; and lack of data protection knowledge, expertise, and awareness amongst stakeholders as the major problems associated with research studies that use data subject rights as a data collection method. Using data subject rights to conduct research is in its infancy stage (Ausloos & Veale, 2021). We believe that our contribution will be valuable to researchers, designers, policymakers, and other stakeholders in addressing new, existing, and future challenges associated with responsible ways of designing and using technologies that incorporate the practical implementation of data protection regulations. On a final note, we acknowledge that the GDPR was favored in our primary studies search, which can be interpreted as a possible bias against other data protection regimes. As many countries or regions of the world are legislating (or have) data protection frameworks, future reviews could look at the possibilities of integrating and analysing many

data protection regulation frameworks.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is part of the first author's Ph.D., supported by the University of St Andrews' School of Computer Science Handsel Scholarship and the Nigerian Government's Petroleum Technology Development Fund. We are grateful to Vicki Cormie, Senior Librarian (Academic Liaison) at the University of St Andrews, for the valuable contribution that helped ease the online database search. We thank Ethan Kenny, a former MLitt student at St Andrews University, and Aliyu Yusuf of the Department of English and Literary Studies, Bayero University Kano for reading and commenting on the first draft of the work. The work was presented at the 2nd Annual Scottish Law and Innovation Conference, "SCOTLIN 2023", but only in an early form, and the conference had no published proceedings. Valuable comments on the earlier work by the conference attendees are also greatly acknowledged. We would also like to thank our anonymous expert reviewers and the editors of this Journal for their time and feedback.

References

arXiv, LawArXiv Papers (2022). URL https://osf.io/preprints/lawarxiv.

Asghari, H., van Biemen, T., & Warnier, M. (2021). Amplifying privacy: Scaling up transparency research through delegated access requests. *5th Workshop on technology and consumer protection (ConPro '21)*. IEEE. URL https://www.ieee-security.org/TC/SPW2021/ConPro/papers/asghari-conpro21.pdf.

Association for Computing Machinery, Association for computing Machinery digital library (2022). URL https://dl.acm.org/.

Association for Computing Machinery, ACM computing classification system (2023). URL https://dl.acm.org/ccs.

Ausloos, J., & Dewitte, P. (2018). Shattering one-way mirrors - data subject access rights in practice. *International Data Privacy Law, 8*(1), 4–28. https://doi.org/10.1093/idpl/ipy001

Ausloos, J., & Veale, M. (2021). Researching with data rights. *Technology and Regulation, 2020*, 136–157. https://doi.org/10.26116/TECHREG.2020.010

Bellanova, R., Bergersen, S., Mirshahi, M., Moe-Pryce, M., & Burgess, J. P. (2017). Exercising access rights in Norway. In C. Norris, P. de Hert, X. L'Hoiry, & A. Galetta (Eds.), *The unaccountable state of surveillance: Exercising access rights in europe, law, governance and technology series* (pp. 257–296). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47573-8_10.

Bier, C., Kompf, S., & Beyerer, J. (2017). A study on corporate compliance with transparency requirements of data protection law. In R. Leenes, R. van Brakel, S. Gutwirth, & P. De Hert (Eds.), *Data protection and privacy: (In)visibilities and infrastructures* (pp. 271–289). Governance and Technology Series, Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-50796-5_10. Vol. 36 of Law.

Cagnazzo, M., Holz, T., & Pohlmann, N. (2019). GDPiRated – stealing personal information on- and offline. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (pp. 367–386). LNCS. https://doi.org/10.1007/978-3-030-29962-0_18, 11736.

Cellan-Jones, R.; Coronavirus: England's test and trace programme 'breaks GDPR data law', https://www.bbc.co.uk/news/technology-53466471, accessed: 2023-06-20 (2020).

Clarivate. (2023). *Journal Citation Reports*. URL https://jcr.clarivate.com/jcr/home.

Confessore N. , Cambridge analytica and Facebook: The scandal and the fallout so far, The New York Times (2018). URL: https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.

Council of Europe, Modernisation of the data protection "convention 108" (undated). URL: https://www.coe.int/en/web/portal/28-january-data-protection-day-factsheet.

De Terwangne, C. (2021). Council of Europe convention 108+: A modernised international treaty for the protection of personal data. *Computer Law & Security Review, 40*, Article 105497. https://doi.org/10.1016/j.clsr.2020.105497

Directive. (1995). 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. *Official Journal of the European Union*, 31–50. URL http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31995L0046.

Ebbers, S., Ising, F., Saatjohann, C., & Schinzel, S. (2021). Grand theft app: Digital forensics of vehicle assistant apps. In *ARES 21: Proceedings of the 16th international conference on availability, reliability and security, no. 30 in ARES 21, association for computing machinery* (pp. 1–6). https://doi.org/10.1145/3465481.3465754

Edwards, L. (2018). *Data protection: Enter the general data protection regulation, law, policy and the internet*. Hart Publishing. https://doi.org/10.2139/ssrn.3182454. may 2018.

Elhabbash, A., Salama, M., Bahsoon, R., & Tino, P. (2019). Self-awareness in software engineering: A systematic literature review. *ACM Transactions on Autonomous and Adaptive Systems, 14*(2), 1–42. https://doi.org/10.1145/3347269

Elsevier, Scopus digital library (2022). URL https://www.scopus.com/.

European Union, COMMISSION DECISION of 20 December 2001 pursuant to directive 95/46/EC of the European parliament and of the council on the adequate protection of personal data provided by the Canadian personal information protection and electronic documents act. (2002). URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32002D0002.

Fonio, C., & Ceresa, A. (2017). Exercising access rights in Italy. In C. Norris, P. de Hert, X. L'Hoiry, & A. Galetta (Eds.), *The unaccountable state of surveillance: Exercising access rights in Europe, law, governance and technology series* (pp. 181–218). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47573-8_8.

Galetta, A., Fonio, C., & Ceresa, A. (2016). Nothing is as it seems. The exercise of access rights in Italy and Belgium: Dispelling fallacies in the legal reasoning from the 'law in theory' to the 'law in practice. *International Data Privacy Law, 6*(1), 16–27. https://doi.org/10.1093/idpl/ipv026

Gellman R., Fair information practices: A basic history-version 2.22 (2022). URL https://www.bobgellman.com/rg-docs/rg-FIPShistory.pdf.

Giannopoulou, A., Ausloos, J., Delacroix, S., & Janssen, H. (2022). Intermediating data rights exercises: The role of legal mandates. *International Data Privacy Law*. https://doi.org/10.1093/idpl/ipac017, 00 (0).

Government of Canada, Personal information protection and electronic documents act (2000). URL: https://laws-lois.justice.gc.ca/ENG/ACTS/P-8.6/FullText.html/.

Greenleaf G., Global tables of data privacy laws and bills (7th Ed, January 2021), Privacy Laws & Business International Report 169 & 170 (2021) 6–19. . URL https://ssrn.com/abstract=3836261.

Hill K., S. Mattu, The house that spied on me, https://gizmodo.com/the-house-that-spied-on-me-1822429852, accessed: 2023-06-20(2018).

IEEE, IEEE Xplore digital library (2022). URL https://ieeexplore.ieee.org/.

Keshav, S. (2007). How to read a paper. *ACM SIGCOMM Computer Communication Review, 37*(3), 83–84. https://doi.org/10.1145/1273445.1273458

Kitchenham B., S. Charters, D. Budgen, P. Brereton, M. Turner, S. Linkman et al., Guidelines for performing systematic literature reviews in software engineering, Technical report, ver. 2.3 EBSE technical report. EBSE, School of Computer Science and Mathematics Keele University and Department of Computer Science University of Durham (2007).

Kröger, J. L., Lindemann, J., & Herrmann, D. (2020). How do app vendors respond to subject access requests? A longitudinal privacy study on iOS and Android Apps. In *Proceedings of the 15th International Conference on Availability, Reliability and Security, ARES '20, Association for Computing Machinery* (pp. 1–10). https://doi.org/10.1145/3407023.3407057

Krieger-Lamina, J. (2017). Exercising access rights in Austria. In C. Norris, P. de Hert, X. L'Hoiry, & A. Galetta (Eds.), *The unaccountable state of surveillance: Exercising access rights in Europe, law, governance and technology series* (pp. 359–404). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47573-8_4.

Kuschewsky, M., What does the revision of the OECD privacy guidelines mean for businesses? (2013). URL: https://www.cov.com/~/media/files/corporate/publications/2013/10/what_does_the_revision_of_the_oecd_privacy_guidelines_mean_for_businesses.pdf.

L'Hoiry, X., & Norris, C. (2017). Exercising access rights in United Kingdom. In C. Norris, P. de Hert, X. L'Hoiry, & A. Galetta (Eds.), *The unaccountable state of surveillance: Exercising access rights in Europe, law, governance and technology series* (pp. 359–404). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47573-8_13.

Lukács A., What is privacy? The history and definition of privacy (2016). URL http://publicatio.bibl.u-szeged.hu/10794/7/3188699.pdf.

Mahieu, R., Asghari, H., Parsons, C., van Hoboken, J., Crete-Nishihata, M., Hilts, A., et al. (2021). Measuring the brussels effect through access requests: Has the european general data protection regulation influenced the data protection rights of Canadian citizens? *Journal of Information Policy, 11*(1), 301–349. https://doi.org/10.5325/jinfopoli.11.2021.0301

Mahieu, R., Asghari, H., & Van Eeten, M. (2018). Collectively exercising the right of access: Individual effort, societal effect. *Internet Policy Review, 7*(3). https://doi.org/10.14763/2018.3.927. Jul.

Martino, M., Robyns, P., Weyts, W., Quax, P., Lamotte, W., & Andries, K. (2019). Personal information leakage by abusing the GDPR "right of access. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019), USENIX Association, USENIX Symposium on Usable Privacy and Security (SOUPS) 2019* (pp. 371–386). August 12–13, 2019 •URL https://www.usenix.org/system/files/soups2019-di_martino.pdf.

Mayer J., Princeton-Radboud study on privacy law implementation (dec 2021). URL https://privacystudy.cs.princeton.edu/.

Meneely, A., Smith, B., & Williams, L. (2013). Validating software metrics: A spectrum of philosophies. *ACM Transactions on Software Engineering and Methodology, 21*(4), 28. https://doi.org/10.1145/2377656.2377661, 24:1–24.

Norris, C., de Hert, P., L'Hoiry, X., & Galetta, A. (2017). The unaccountable state of surveillance. LGTS, volume 34 of Law. *Governance and technology series*. New York, NY: Springer Cham. https://doi.org/10.1007/978-3-319-47573-8.

Norris, C., & L'Hoiry, X. (2017). Exercising citizen rights under surveillance regimes in Europe – meta-analysis of a ten country study. In C. Norris, P. de Hert, X. L'Hoiry, & A. Galetta (Eds.), *The unaccountable state of surveillance: Exercising access rights in europe, law, governance and technology series* (pp. 405–455). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47573-8_14.

Open Society Foundations, Q and A: Fighting for worker's right to data (May 2019). URL: https://www.opensocietyfoundations.org/voices/q-and-a-fighting-for-workers-right-to-data.

Organisation for Economic Co-operation and Development, The OECD privacy framework (2013). URL: https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf.

ProQuest LLC, ProQuest dissertations and theses (2022). URL https://www.proquest.com/dissertations-theses/.

Raento, M. (2006). The data subject's right of access and to be informed in Finland: An experimental study. *International Journal of Law and Information Technology, 14*(3), 390–409. https://doi.org/10.1093/ijlit/eal008

Regulation (EU). (2016). 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union, L119*, 1–88. URL http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC.

Rothmann, R. (2017). Video surveillance and the right of access: The empirical proof of panoptical asymmetries. *Surveillance and Society, 15*(2), 222–238. https://doi.org/10.24908/ss.v15i2.6029

Rudgard S., Origins and historical context of data protection law (undated). URL: https://iapp.org/media/pdf/publications/European_Privacy_Chapter_One.pdf.

Sørum, H., & Presthus, W. (2020). Dude, where's my data? The GDPR in practice, from a consumer's point of view. *Information Technology and People, 34*(3), 912–929. https://doi.org/10.1108/ITP-08-2019-0433

Spiller, K. (2016). Experiences of accessing CCTV data: The urban topologies of subject access requests. *Urban Studies, 53*(13), 2885–2900. https://doi.org/10.1177/0042098015597640

State of California, Assembly Bill No. 375. California legislative information (June 2018). URL: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375.

Szekely, I., & Vissy, B. (2017). Exercising access rights in Hungary. In C. Norris, P. de Hert, X. L'Hoiry, & A. Galetta (Eds.), *The unaccountable state of surveillance: Exercising access rights in europe, law, governance and technology series* (pp. 135–180). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47573-8_7.

Tolsdorf, J., Fischer, M., & Iacono, L. Lo (2021). A case study on the implementation of the right of access in privacy dashboards. Vol. 12703. In *LNCS of lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer. https://doi.org/10.1007/978-3-030-76663-4_2.

Turner, S., Galindo Quintero, J., Turner, S., Lis, J., & Tanczer, L. (2021). The exercisability of the right to data portability in the emerging Internet of Things (IoT) environment. *New Media and Society, 23*(10), 2861–2881. https://doi.org/10.1177/1461444820934033. URL https://journals.sagepub.com/doi/pdf/10.1177/1461444820934033.

Uldam, J. (2016). Corporate management of visibility and the fantasy of the post-political: Social media and surveillance. *New Media and Society, 18*(2), 201–219. https://doi.org/10.1177/1461444814541526

Uldam, J. (2018). Social media visibility: Challenges to activism, Media. *Culture and Society, 40*(1), 41–58. https://doi.org/10.1177/0163443717704997

Urban, T., Tatang, D., Degeling, M., Holz, T., & Pohlmann, N. (2019). A study on subject data access in online advertising after the GDPR. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 61–79). LNCS. https://doi.org/10.1007/978-3-030-31500-9_5, 11737.

Van Biemen, T. (2018). *Personal privacy in practice: Putting the gdpr to test in a collective exercise of data subjects' right of access, master's thesis*. Netherlands: Delft University of Technology. Mekelweg 5, 2628 CD DelftOct.URL: http://resolver.tudelft.nl/uuid:ccea2ec8-5ecb-47e3-8fae-79733d765093.

Veale, M., Binns, R., & Ausloos, J. (2018). When data protection by design and data subject rights clash. *International Data Privacy Law, 8*(2), 105–123. https://doi.org/10.1093/idpl/ipy002

Veys, S., Serrano, D., Stamos, M., Herman, M., Reitinger, N., Mazurek, M. L., & Ur, B. (2021). Pursuing usable and useful data downloads under GDPR/CCPA access rights via Co-Design. In *Seventeenth symposium on usable privacy and security (SOUPS 2021), USENIX Association, USENIX Symposium on Usable Privacy and Security (SOUPS) 2021. August 8–10, 2021, Virtual Conference.* (pp. 217–242). URL https://www.usenix.org/system/files/soups2021-veys.pdf.

Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., & Wagenknecht, K. (2021). *The science of citizen science*. New York, NY: Springer Nature. https://doi.org/10.1007/978-3-030-58278-4

Von Laufenberg, R. (2017). Exercising access rights in Luxembourg. In C. Norris, P. de Hert, X. L'Hoiry, & A. Galetta (Eds.), *The unaccountable state of surveillance: Exercising access rights in Europe, law, governance and technology series* (pp. 219–255). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47573-8_9.

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review, 4*(5), 193–220. https://doi.org/10.2307/1321160

Web of Science, Web of science collections (2022). URL https://www.webofscience.com/.

Wong, J., & Henderson, T. (2019). The right to data portability in practice: Exploring the implications of the technologically neutral GDPR. *International Data Privacy Law, 9*(3), 173–191. https://doi.org/10.1093/idpl/ipz008

WoS Research Team, Web of Science Journal (WoS) Journal Info (2023). URL https://wos-journal.info/.

Zwiebelmann, Z., & Henderson, T. (2021). Data portability as a tool for audit. In *Adjunct proceedings of the 2021 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2021 ACM international symposium on wearable computers* (pp. 276–280). Association for Computing Machinery. https://doi.org/10.1145/3460418.3479343.