ORIGINAL PAPER

Open research data, an archival challenge?

Charlotte Borgerud[1] · Erik Borglund[2]

Abstract

The Swedish government has decided that all research results in the form of research data and scientific publications financed with public funds should be openly accessible as far as possible. The question is whether the responsible actors and if the universities are ready to implement the change. The significance of open access has amplified in Sweden. Earlier research has brought to light that the collection and preservation of research data are often surrounded by ambiguous rules and lack a comprehensive structure. For example, archiving is not given enough consideration in connection to research projects and researchers often tend to save their material on platforms that are not persistent over time. This article is based upon a qualitative research approach where 15 semi-structured interviews have been used as primarily data sources to investigate the implementation of open access of research data and scientific publications. The article investigated how Swedish universities and public authorities were working with archiving and implementation of open research data and their opinions on open access. The results displayed a lack of coordination, resources and infrastructure but also that common agreed nomenclature were missing. The management of research data was not part of an overall recordkeeping strategy. One explanation could be differences in the information culture among researchers and archivists. Social sciences theory has been combined with archival theory in order to explain the reasons to this. These have been put in relation to the principles of the open data directive.

✉ Charlotte Borgerud
charlotte.borgerud@skl.se

Erik Borglund
erik.borglund@miun.se

1   Swedish Association of Local Authorities and Regions, 118 82 Stockholm, Sweden

2   Mid Sweden University, 85170 Sundsvall, Sweden

Introduction

This article investigates the utilization and accessibility of universities' research material and how the various authorities responsible for open science are working to implement open research data. The paper aims to address challenges in the management of open research data from an archival perspective. Open research data springs from the same ideological source as open data. In Europe the PSI Directive was the precursor of the European Open Data Initiative (Finansdepartementet 2009). The idea behind open data is that it should be possible to reuse public information for free to enable new ideas and innovations. In many situations open data are synonymous with records (Borglund and Engvall 2014); there is an obvious connection between archival institutions and open data. Schellenberg (1956/1998) introduced the concept of primary and secondary value of records, where the secondary value is the value records have for anyone other than the originating agency. Analysis of the PSI Directive makes it clear that reuse of public data is intended for users outside of the originating agency. Thus, it can be argued that open data is very strongly focused on the secondary value of public data. Another relevant and applicable categorization is presented by Shepherd and Yeo (2003), who also categorize the use of records into primary and secondary use, where secondary use is outside the original purpose of records creation, and which can be categorized further into the three purposes of using records: business, accountability and cultural. Yet, another argument for an obvious connection between archives and open data is through the records continuum model (Upward 2000), where the fourth dimension is "pluralize." In pluralization, Upward (2005) argues, records can be used in less predictable ways, but values other than the original are also found in this dimension. According to McLeod (2012), open data is a great opportunity for records professionals, but it is an opportunity that comes with a set of problems and challenges. Open data projects can contribute to records management at the same time as records managers' competences contribute to open data projects (Serra 2014).

Previous research on open data concludes that regulations concerning the collection and preservation of research data are unclear. Records professionals, such as records managers and/or archivists, are rarely in charge of the management of research data, despite having the necessary competence (Grant 2017). In addition, archival aspects are rarely taken into consideration in research projects, and there is a widespread lack of knowledge about how to preserve research data over time. Too often researchers store research data on media that are not trusted for preservation, risking loss of important data. Ethical risks of open data that deal with human participation and qualitative data have also been identified, but the involvement of archivists could well help to solve this problem (Childs et al. 2014). Therefore, there is a need to better understand open research data challenges from an archival perspective.

Little research has been conducted on archiving research data and making research data openly accessible. The opposite could be said about open access of publications. It is convenient to study the research conducted on open access to find arguments in support of open research data, since they are connected and both fall

under the umbrella term of open science. However, one problem with research on open science is that it is often advocated by those who have authored the articles, and therefore not always neutral. It can be noted that the main issue in the literature is not whether research should be made open access or not, but how it should be done. There are also articles that are strongly opposed to open access, albeit in the minority.

A master's thesis discovered that the university archivists considered the amount of research material they received to be limited, and did not think the researchers had any greater interest in or knowledge about archiving (Brinck and Leuhusen 2013). Many researchers considered archiving and access to research papers important, but at the same time, they did not consider it important to preserve their own research documents, believing the documents to be their own and not belonging to the university. Furthermore, the researchers did not consider their research material as public, but as working material. The public's right to request access to research material was seen as time-consuming and potentially expensive; partly because it could paralyze the research and partly because people could request documents without what the researchers considered a legitimate need for it. Some researchers also argued that a certain ego is required in order to submit documents to the archive (Brinck and Leuhusen 2013).

There are publications on interpreting the laws that surround archiving research material and others studying the structuring of research data. However, these works mostly focus on developing recommendations for working with the issues, rather than examining attitudes and working methods among archivists and records managers today (Bohlin 1997; Borgman 2015; Corti et al. 2014). The aim of this study is to investigate the archivists' role in managing open research data by investigating attitudes and working methods among different Swedish universities.

Analytical framework

The analytical framework used in our research is the Mertonian norms of science (Merton 1973).

The sociologist Robert K. Merton believed that the research community is governed by values and norms. These are expressed in regulations, proscriptions, permissions and approaches and are legitimized in institutional values. Merton divides the norm system into four institutional imperatives (Merton 1973, pp. 268–279):

- *Universalism* states that science should be objective, rational and accessible. Exclusion for social reasons such as ethnicity or religious belief must not occur. (Merton 1973, pp. 268–270)
- *Communism* is about the sense of common ownership. The research that is created is a product of cooperation and therefore belongs to the entire research community. The researcher has no ownership other than of the recognition and appreciation the researcher receives for the scientific discovery (Merton 1973, pp. 270–273)

- *Disinterestedness* means that the scientific institutions work for common benefit and not for personal gain. (Merton 1973, p. 273)
- *Organized Skepticism* is both a methodological and institutional norm. It states that scientific claims must be tested and objectively reviewed by other qualified researchers before the claim can be accepted. (Merton 1973, p. 268)

Merton's theories on ethical norms and a reward system in academia are useful for the investigation since the theories can explain why researchers are willing or unwilling to publish their research findings and data via open access. Publishing is one of the foundations of the research world, as a way to obtain confirmation of the research findings. At the same time, the quest for originality can explain why some researchers are unwilling to share their research data via open access. If data is published openly before the study is complete, there is a risk that another scientist will take advantage of a scientifically interesting discovery. However, open access and open data can contribute to wider dissemination, increasing the chances of recognition from the research community. It is important for archivists and others who work with preserving and making research available to be aware of the premises that drive the researchers in order to help make scientists more open to archiving and accessibility issues.

Previous research

When studying previous research on archiving of research materials, open research data and open access, several themes were identified. These themes provide the structure of the article and the interview study.

Preservation and reuse of research data

Data sharing and data reuse are not new for the scientific community, their positive aspects have been promoted for quite some time, but it is still challenging to identify "which data might be shared, by whom, with whom, under what conditions, why, and to what effects" (Borgman 2012). The use of the concept *data* is rather general, and there is not always a distinction between quantitative and qualitative data. According to (Corti and Thompson 2007) qualitative data is a source of research that is less used, and the reuse of such data is not without ethical and methodological challenges. Faniel and Zimmerman (2011) argue that the approach to the area of reuse and sharing research data is too narrow, and propose a wider more inclusive research approach for reuse and sharing of data. Wallis et al. (2013) ask which benefits shared data actually generate, they argue for a better understanding of reuse practice. There is a limited knowledge about the actual use of shared research data, and Faniel et al. (2016) identified a set of data quality attributes that was significantly related to the data reuse satisfaction. But the data that is shared also needs to be understandable, which is a job for the data repository staff (Faniel et al. 2012). The understanding of how other researchers can assess the quality of and possibility

to reuse data is also a challenge (Faniel and Jacobsen 2010), and it is an area that needs to be further studied to support curatorial decisions as well as actual reuse (Palmer et al. 2011). The preservation and reuse of qualitative data are identified as more challenging than preserving quantitative data (Corti et al. 2005).

In this article, the term data is used regardless of whether it is qualitative or quantitative data.

Arguments in favor of open research data

The most common arguments for open data were to facilitate research verification and increased transparency. Making data open facilitates reuse of data in new studies, which might speed up innovation, since it makes collaboration easier. In addition, other communities than those originally intended could use the data to create new things, for example, apps designed for smart phones (Borgman 2015; Corti et al. 2014; Doorn and Tjalsma 2007; Fransson et al. 2016; Fransson et al. 2017; Hammersley 1997; Piwowar 2011; Piwowar et al. 2007; The Royal Society 2012). The data is also more visible, which could encourage higher exploitation (Piwowar et al. 2007). Some emphasize that reusing completed research data is very cost effective as past results can be studied and meta-analyses made without collecting the same material again (Piwowar et al. 2007; Vines et al. 2013). This has been questioned on the basis that the workload to make data understandable to others requires a lot of time and money (Hammersley 1997). Open research data can be reused to test generalizations, and results can be compared using other contexts, variables or geographical areas. New research methods may also arise that could be interesting to apply to older data. Data could also be used in teaching where students can concentrate on analyzing data instead of collecting it. Furthermore, it has a historical value where the behaviors and attitudes of groups and individuals could be studied, as well as reveal how they worked at a certain time, what methods and theories were popular and so on (Corti et al. 2014).

In conclusion, the arguments are about cost-effectiveness and democratization. By making data openly accessible, more people, not only scholars, will have the opportunity to use and reuse the data, which increases transparency. This is important in order to discover research misconduct and facilitate cross-border strategies and approaches.

Arguments against open research data

There is lack of knowledge about the management of open data—both from researchers and from academies (Åhlfeldt and Johnsson 2015; Piwowar et al. 2007). For example, what is a project and how could the problems of conflicting publicity and privacy laws be resolved? The literature offers no real answers, but instead claims that a government review should be conducted (Hermerén et al. 2011). Some studies have shown that many researchers are initially strongly opposed to their research data being completely open, while others think that they have made data open by presenting the results in publications (Björklund and Eriksson 2007). Previous

studies have showed that researchers tend to see research documents as their private property, rather than belonging to the university or the public (Björklund and Eriksson 2007; Fransson et al. 2016). Legal restrictions concerning transparency, commercial interests, sensitive personal data and national security might also hinder the researcher from publishing the data as open access (The Royal Society 2012). Previous studies have also illustrated how many researchers do not want the research data to be open and accessible to anyone, in order to protect the individuals in the studies who have been promised strict anonymity (Björklund and Eriksson 2007; Mauthner et al. 1998). As mentioned in the previous section, some believe that data is difficult to interpret and is not understandable to third parties. Some also point out that researchers might provide raw data upon request, but are not open to publishing it. Researchers thus tend to confuse anonymity and confidentiality. A problem is that there may be a conflict between requirements for transparency and personal integrity. Weightings depend on the type of research, whether it involves sensitive personal data or not, since different regulations govern sensitive or non-sensitive data. The issue of protection of personal data is also an international issue where it cannot be guaranteed that people are completely protected through anonymization measures (Björklund and Eriksson 2007; Hermerén et al. 2011; Mauthner et al. 1998; The Royal Society 2012). The Swedish Research Council concluded that areas requiring further investigation include the management of sensitive personal data and research where there is commercialization potential, such as in collaboration with industry. They note that new users of the research data must observe the same legal and ethical standards as the first researcher (Vetenskapsrådet 2015). It is important to remember that accessibility and archiving research data are two different things. Regardless of whether the material should be made available or not, data management plans are important and should be included in research plans and applications (Fransson et al. 2016). It is also important that researchers' understanding of their legal obligations is increased—where a public authority is principally responsible for the research, researchers are obliged to archive their research data. However, it is not always clear how open data should be archived. Some documents are unavailable in digital or paper format, such as artistic artifacts and biomaterials.

The reward system in the research world is often based on the development of new data, which means it is easier to acquire funding for studies that do exactly that: develop new data (Borgman 2015). Therefore, many researchers feel that providing open access to their data risks their competitiveness with other research groups. A counter argument, however, is that researchers are rarely expected to publish data openly immediately. Most often, there are embargo times (also called proprietary periods), which may vary from a few months to a few years. This period should be long enough for the researchers to analyze the data and publish their results. In some research fields, for example, the humanities and the social sciences, it has become socially acceptable to hold on to data without making the material available (Borgman 2015). Since all experiments are unique and it takes time to understand other people's data—they may have used other questions, methods and theories—reusing the material in other studies could be difficult. Equally, it may be difficult to understand the context (Corti et al. 2014; Hammersley 1997). In one study, the researchers returned to their own research data, which they had not used for several years

(this applied to qualitative data), and the conclusion was that they did not have the same understanding of the data (Mauthner et al. 1998).

A further disadvantage of making data accessible is that it is time-consuming. The data must be formatted to enable understanding, and it is not always obvious where it should be published (Mauthner et al. 1998; Piwowar et al. 2007). It is also not easy to determine what should be classified as essential data and what is insignificant, and who should determine this (Parry and Mauthner 2004).

Another challenge is choosing persistent file formats for storage, as well as ensuring that they remain compatible over time, as operating systems and technologies change on a regular basis. Many researchers use widely different systems to save research data (Björklund and Eriksson 2007; Corti et al. 2014). Key questions that arise are whether it is possible to create a universal solution that works for several disciplines, or whether this is even desirable. In their study, (Björklund and Eriksson 2007) question whether general rules really can be applied to specific situations.

In summary, previous research is not against open research data as such, but there are many challenges that need solving, and one solution will probably not fit all research disciplines or studies. The greater question at hand is the archiving of research data, regardless of whether or not it should be made openly accessible. By starting to discuss these issues and how to store and map research data, we could gain an understanding of what steps to take next, toward accessibility.

The discipline matters

Many researchers point out that the way in which science is communicated differs between research fields. This also entails different concepts or notions, which complicate the work of e-publishing (Borgman 2015; Hedlund and Annikki 2006; Kling and McKim 1999). Some studies show that researchers' attitudes to open research data varies between scientific disciplines because of different needs (Björklund and Eriksson 2007; Corti et al. 2014). Studies have shown that researchers in physics, astronomy, computer science and mathematics are most positive and knowledgeable about open access, whereas medicine, law and some disciplines under the umbrella social sciences and humanities are not (Meyer Lundén 2008).

Archiving techniques for data retention have focused mainly on quantitative data (Fink 2000). Several of the articles point out that archiving qualitative data and making it available differs as making qualitative material open involves certain challenges. (Doorn and Tjalsma 2007). Experimental methods also require another type of archiving technique (McDermott 2014). In quantitative studies, the researcher's role is minimized in the data collection itself, while the opposite is true of qualitative studies (Mauthner et al. 1998). It is more difficult to anonymize people and places in qualitative studies, and it has been suggested that there is an interest in protecting the researcher as much as the informants in the study (Hammersley 1997). As respondents in qualitative studies do not have the same anonymity, there is a risk that researchers feel an extra responsibility to protect the information from third parties. Moreover, their working methods can make the researcher feel that the material cannot be fully understood by other researchers who may misinterpret data and

come to other conclusions (Fink 2000; Parry and Mauthner 2004). Qualitative data is collected using specific parameters, which may make the material difficult for other researchers to use. At the same time, data can be used to confirm generalizations and work in comparative studies. One way to make qualitative data more comparable is standardization, but the question is to what extent the research field would benefit from this (Hammersley 1997). The different approaches are an epistemological issue, which should be considered when archiving. Data have a historical value and provide information about how research was conducted at a certain point in time (Mauthner et al. 1998).

The message is that a single solution will perhaps not fit all types of research. A solution that has been suggested is to have specialist departments that work closely with the researchers at the beginning of a project to help them structure and create a plan for the management and preservation of the data. Archivists play an important professional role in this and could help the research community to decide on a suitable solution as needed. It can be noted that in the literature there is no clear and defined border between archiving research data and preserving research data for reuse.

Methods

An exploratory qualitative study was carried out in Sweden, investigating working methods and attitudes relating to open access and the availability of research data. Interviews were structured based on recurrent themes identified in previous research in order to facilitate a comparison (see above).

Fifteen semi-structured telephone interviews were conducted. Respondents were selected in part from the actors commissioned by the government to investigate the implementation of open data[1] and in part from state-owned universities with the right to issue doctoral degrees.

In the latter group, consisting of 27 universities, 6 universities were chosen with the aim to ensure a geographical spread across the country and to represent organizational differences and subject specializations. The universities include some of the largest universities in Sweden teaching all subjects, as well as specialist institutions, for example, in artistic artifacts, textiles, agriculture or medicine, demonstrating varying needs and problems when handling research data (Table 1). Nine respondents represent the universities, and five the various public authorities. The study is limited and does not claim generalizable results concerning archivists' working methods and attitudes. Our plan was that the respondents would address both the institution's and their personal policies, but the universities had no explicit policy. Therefore, the interviews mostly focused on respondents' personal opinions.

---

[1] The National Library of Sweden, the Swedish National Data Service, the Swedish Research Council and Vinnova. The National Archives is the authority responsible for transparency in government agencies and a key player in the implementation of open data but did not participate in the investigation. It was the most difficult of all the organizations to communicate with.

Table 1 Universities in Sweden

No.	Employees	PhDs	FTE students[a]	Students in total	Geographical area
1	500–800	50–100	5000–7000	10,000–12,000	South
2	5000–7000	2000–3000	5000–7000	9000–11,000	Middle
3	7000–9000	2000–3000	26,000–28,000	40,000–42,000	South
4	4000–6000	800–1000	17,000–19,000	33,000–35,000	North
5	2000–4000	250–450	3000–5000	4000–6000	Widespread
6	4000–6000	1000–2000	25,000–27,000	32,000–34,000	Middle

[a]Full-time equivalent

Table 2 Interview form

Respondent's knowledge, experience and work with open research data and open science
Respondent's opinion on open research data and open science
What does the respondent think the researchers' opinions are on open research data?
Is attitude to open research data dependent on science discipline?
What opportunities does the respondent see with open research data?
What challenges does the respondent see with open research data?
Infrastructure
Coordination and cooperation
Respondent's recommendations
Other reflections that have not been raised but may be of interest

The selected universities and actors were contacted by email, and snowball sampling (Williamson and Johanson 2013) was used to identify potential respondents. There were five representatives of public authorities (henceforth referred to as "actors"). From the universities were four archivists, four librarians, one research coordinator/project leader and one researcher in medicine as comparison. Each interview lasted from 30 min to just over an hour. The interviews were recorded with a dictaphone and transcribed word for word by the interviewer. Next, the transcription was sent to the respondents to give them the opportunity to review and comment. The interviews were open in character to let the respondent speak freely, but based on an interview guide with certain themes, shown in Table 2. The interview guide was based on the themes identified in previous research to facilitate a comparison of the results. It is noteworthy that despite the fact that each respondent could put their own mark on each interview, the same things were addressed and the same opportunities, challenges and recommendations were identified. These in turn were in line with the questions and ideas the literature raised about both open access publications and open research data. The question becomes whether there are a certain number of given answers to the questions about open research data or if the industry is so small that it is easy to be directly or indirectly affected by certain views and then unconsciously influenced. What contradicts the latter argument is that many of the respondents were unfamiliar with the questions and yet gave the same answer as those who

felt quite familiar with open access and open research data. The investigation was conducted in accordance with the Swedish Research Council's ethical recommendations (Vetenskapsrådetz 2015).

An inductive method was used when analyzing the interviews, and the Mertonian norms were applied to the analysis. The transcribed interviews were coded using keywords to facilitate comparison with previous research and then related to the Mertonian norms to see whether the study's results were consistent and could be explained by Merton's institutional imperatives. This was done by grouping the keywords into the themes identified in the review of previous research.

Research findings

In this section, the results are presented without applying the Mertonian norms, and descriptive headings are used.

The respondents' awareness of open research data

Familiarity with open data varied mostly among the universities, where six can be said to be very familiar and four acquainted with open research data. Among the actors, all respondents can be considered familiar with open data. It is not surprising that the concept of open data is known to many of the respondents; several of the participants in this investigation had been recommended for the study. We were sometimes referred to different people at the institutions and universities who said they were not familiar with open data. This reveals an ambiguous organization in which it is not always obvious who is responsible and who should be contacted and demonstrates a need to raise awareness within the organization.

It was clear that the actors and the university librarians generally had a higher understanding and experience of working with open research data than the university archivists had. However, this experience was most often limited to providing information about open science and research data. Only two of those interviewed (both actors) had worked with publishing open data (but not research data). The general pattern was that the archivists had not worked with open data, since open science was the library's responsibility. The archivists focused on archiving and storage of data, regardless of whether it was to be made open access or not. It is probably easier to motivate researchers and the professions working with data at the universities if the actors themselves have experience with making research data open. It was evident that the archivists did not work as actively with research support as the librarians did. The librarians had campaigns where they invited researchers, arranged workshops and lectures and had more digital channels for their support systems than the archivists had. The sample is too small to draw definitive conclusions, but it seemed that the more active the library, the more active the archive. In the universities where the librarians were running active campaigns in open science, the archivists were giving lectures on archiving data. While the archivists did

not provide information about making data open access, they at least informed the researchers about the necessity of archiving.

Another problem is archivists who refer researchers to the library, i.e., not taking a more active role. It is notable that none of them had worked with open research data and seemed to believe that the library should be in charge of this process.

There was no clear connection between the university's location, or orientation, and the employees' experience of open data—except in one case. One of the universities specializing in, among other things, geospatial data, had a collaboration with two public authorities, which meant that the authorities owned the data, but the university was charged with archiving the data and making it open access. This is not surprising as the data is comparable to public data that has been made open for years. That is, it seems there is a greater tendency to make data open when it is not sensitive data, and when the accessibility could benefit many. This is in line with previous research. It could be assumed that universities with sensitive research data, or data that is difficult to define, such as artistic objects, would have less experience or be more negative than universities with data that is easier to make open access, but no such patterns were found. Nor were there any patterns of larger universities, or universities with a lot of quantitative data, having more experience of open research data and more trained personnel than smaller universities or the medical university. This could be because the sample is small. Interestingly, the medical university had more experience than many of the other universities. It is likely that this was because they had a lot of funding and register-based research. But the archivists had not helped or worked with the databases that collected and made the medical data open and sharable. Instead, the researchers had employed external consultants from the IT community, or hired people with the right competence to their research group, that is with the money they received from funding agencies. This poses a great risk—what happens when there is no more external funding? Who will archive the databases? It also illustrates how the university has failed to employ staff to help the researchers to build databases or use a shared university-owned infrastructure.

Archivist 4 and Actor 3 commented that the concept of open data is not entirely obvious. Actor 3 noted that open data can have two meanings: open government data, i.e., an authority provides open access to their data, it can be downloaded freely, or it can mean open research data. These two things are not the same, which can cause problems.

It is remarkable that different professions at the same university can have such different experience of open research data, considering that only professional groups with a clear link to research data were interviewed. The results have shown that the libraries and research coordinators of various types of research, innovation and funding agencies at the universities generally have a greater insight into the university's work with open research data than the archivists.

Respondents' attitude to open research data and open science

The respondents were mainly positive about open science and open research data. At the same time, the interviewees pointed out that there are challenges that must be

solved, and that maybe not all research data should be made openly available. The respondents were generally more positive to open access of publications than data since it is easier to make publications open access than make data open access. The librarians were the most enthusiastic about open research data and the archivists the least. While both groups had nothing but positive words, and more or less identified the same opportunities and the same difficulties, the level of enthusiasm and propensity to see the positive differed, for instance in Archivist 2's comment *"Positive really but difficult to implement."* The reason could be that the librarians are more used to open science and open access, and therefore apply their experience on open research data. The archivists, who often seemed to have had difficulties receiving research material for archiving, might transfer this experience to the question of open data. It is also possible that they have more knowledge about laws regulating personal data and confidentiality, which could be a basis for their standpoint.

Archivist 3 thought a positive effect of the open science movement could be that more research material is archived as a result of increased awareness. This university had not archived any research data. Some highlighted that there may be a historical interest and that materials can be reused for completely different surveys than originally intended, for example, Actor 3: *"What may be a failed research material today could be extremely valuable as research material in 50 years."* Librarian 4 stated that it is very valuable for researchers to think about how to handle their data because it will increase the quality of the work. The respondents' reflections highlight the importance of involving archivists at the beginning of a research project. By organizing metadata, classified information, anonymized data and general data at the start, archiving at the end of the project will be much easier and might highlight if some of the data could be made open access.

Researchers' attitudes to open research data

Whether the respondents had been in contact with researchers about open research data varies, but the majority stated that they did not know since they were not familiar with the researchers' attitudes or worked close enough with them to know. Here, a potential conflict of interest can be seen: the respondents represent professional groups responsible for the implementation while not having enough knowledge of opinions of those most affected by the implementation. Many respondents seem to be aware of this problem. At the same time, the respondents in this study gave more or less the same answer: namely, attitudes differ from researcher to researcher. Some are very familiar with open research data and have or want to make data open, others do not. However, most respondents' experience is that the researchers generally have little or poor knowledge of open data. A reason could be that the researchers who are familiar with open data do not contact the support units for help. The universities' experiences are that it is primarily the researchers who are forced to publish open access who do it. That is, the inclination to publish open data increases when there is an external requirement. Many of the respondents emphasized that there cannot be too much administration if the aim is to make the researchers participate. The archivists' experiences were that the researchers are not involved in archiving, regardless

of whether it was analogue or digital—but the librarians had a positive experience where many researchers wanted to publish their data open access. Here, once again, we see differences between archivists and librarians. The difference could be based on the differences in mission. The librarians only provide information about the options available, without making any demands about collecting the data for storing and preservation. The archivists, however, are supposed to archive the data, which may make the researcher feel a loss of control. All the archivists agreed that some researchers understand the importance of archiving while others do not, but the archivists were not sure if some research fields were more positive than others. They guessed it was possible, because some fields are likely to be more accustomed to sharing data. Different datasets need different preparation, which is why it is important that the support units work closely with the researchers so they can give the scientists the right support in structuring their data. Also the actors were aware that very little research data is archived in Sweden, which is why they considered it a big step to making data open access. The reasons identified as to why it was difficult to collect research data for archiving were complicated administration and a lack of infrastructure. Archivist 1 also noted that one Italian researcher equated it to spying. Librarian 4 mentioned that most researchers who turned to them were positive about open research data, but those researchers could be met with resistance and doubt from their own institution.

The respondents' answers can be summarized as follows: (1) The universities in the study have no information about the researchers' attitudes or knowledge about archiving and open research data; (2) the researchers are not involved in archiving or making research data open access; and (3) there is a lack of knowledge among researchers, some of whom have not understood that research data belongs to the university and not the researcher. The third point is in accordance with findings reported in previous research.

Differences between research fields and scientific disciplines

While the respondents are unfamiliar with what the research community thinks about open research data, their experience was that attitudes toward and experience of open research data is dependent on the scientific discipline and research. The same answer was given regardless of profession or workplace. Some of the respondents argued that it is a linguistic issue where some research fields are not used to the concept of data.

In addition to the concept of definitions, it is not always clear exactly what data is. Librarian 1 was told by a researcher, "*But I'm just doing mathematical reduction.*" What is the data in a mathematical formula? How does it represent the work process? Another question is, for example, what constitutes data in artistic research? What constitutes data when a new item is created? What will be documented and saved? Librarian 1 reported that some researchers questioned the benefit of making data open access; if the researcher has worked in a specific program and if the data is removed, it loses its context and then what is the use of it? The respondents were also told that researchers wanted to ensure that the next person used the same ethical

guidelines. All of the respondents also felt that personally, they did not always have the answers, and more education was needed. This was especially the case among the universities. The actors, however, felt they needed more input from researchers. A solution could be that the same ethical permissions are required from studies using open research data collected by other researchers.

Several of the interviews make it clear that physics, astronomy, mathematics and various natural sciences are often more positive and used to making data open access, while researchers who work with personal data are more skeptical. Librarian 2 said that "*And they just say no, no it does not work and then they close that door and do not want to talk.*" The more structured the data the easier it is to make it open and understandable for others. Research with sensitive or qualitative data might learn from the various natural sciences how to make parts of their data open access. For example, studies with sensitive data might benefit from making parts of the data open access to register-based research. That way they might find results via multidisciplinary research they would never have found if researchers kept the data for themselves. What is important is to give credit in publications to the researchers who collected the data. It is also important that their method of collection is transparent and clear to facilitate credibility.

Librarian 4 mentioned an example where a researcher in psychology claimed that he was prepared to destroy his data rather than make it open. The researcher stored an encrypted version of his data on a hard drive. The researcher remarked that colleagues at his institution did not share his attitude, and that some might even have made data open access. It is evident from the librarians' example that the researchers have not understood that the data does not belong to them, and that transparency is in conflict with the aim to protect those participating in the study. Maybe parts of the data could be made open access, but it should at least be archived.

What are the opportunities and challenges of open research data?

The opportunities identified by the respondents are summarized in Table 3 and correspond with what previous research has found to be positive about open research data. That is, open research data enables reuse of data, which could make it time and cost effective, since it will not have to be collected again. This makes it easier to verify and validate research results, which could prevent research fraud. The respondents also mentioned increased accessibility and more possibilities to learn about other researchers' results; through digitization a lot could be done with the data. Comparative studies could be done nationally and internationally and through the collaboration and combination of data, facilitating interdisciplinary research, new discoveries and research fields can emerge, leading to increased quality. At the same time, it is pointed out that there are challenges that must be solved, and that all research data may not be made openly available. Archivist 2 had not thought about open research data to any great extent and replied that there might be great opportunities but provided no examples.

Table 3 presents a summary of the opportunities and challenges based on the analysis of the words used by the respondents, and the meaning of their words and

Table 3 Opportunities and difficulties of open research data identified in the interviews

Opportunities	Challenges
Clearer decision making	Globalization and cooperation across borders
Comparative studies can be done nationally and internationally	Jurisprudence, ethical and commercial interests can affect whether the data could be made open access (sensitive data)
Data can be combined in new ways, creating new fields, and new interesting discoveries and results can be obtained	Lack of resources and who should pay?
Digitally, a lot can be done with the information	Need of a common nomenclature
Increased accessibility → more possibilities to learn about other researchers' results	No metadata standards
Increased cooperation and interdisciplinary exchanges are facilitated	No infrastructure—long-time storage
Increased quality and prevention of research fraud—verification and validation	Insufficient information channels and a gap between researchers and administration → lack of information, knowledge and anchoring
Reuse and reiteration → saves time and resources (no need to retrieve the same data)	Willingness to share

sentences. In conclusion, there is no doubt that open research data is regarded positively—the problem is that it is difficult to achieve.

The respondents agreed that there are many challenges with open research data that must be solved. What they mentioned was very similar; their different backgrounds did not seem to affect opinions on possible obstacles or opportunities. The three things listed as the greatest challenges were insufficient resources and infrastructure, metadata standards and nomenclature. Indirectly, it also appears that information efforts are required from everyone involved. One challenge is not only that the infrastructure is not yet in place, but also how the infrastructure should be designed. Some respondents have suggested that maybe it should not be a single infrastructure, or one e-archive, but several. Archivist 3 summarized it as "*if you have no research support and no e-archive, what you do have are challenges.*" Another problem that both the universities and actors acknowledged is that data is not being archived, and it is not evident how to archive a research project (should you follow provenance?)—or even what a research project is. Funding from numerous agencies can lead to several different publications where different ethical permissions apply. Therefore, the archivists and researchers need to collaborate. The university archivists, however, felt the researchers often lacked the time, and the archivists could not make sense of the data on their own.

Librarian 4 had the opportunity to attend a course on open data given by a researcher for PhD students. Librarian 4 found the course quite useful, as the PhD students really addressed the questions, and Librarian 4 left feeling hopeful that this is a generational issue, at least in part. If today's PhD students are provided with information and training on data management plans, their attitude to open data could be different from that of the older researchers.

Interestingly, one of the identified challenges is also identified as an opportunity: increased globalization and collaboration with other countries. All the respondents identified the legal issues relating to sharing data. This is an international question where countries need to facilitate cross-border research by making laws that do not prevent this development, but instead make it easier to share data between countries at the same time as ensuring the integrity of the research subject. This could be done by developing common metadata standards and infrastructure in order to facilitate administration. The overarching laws in Sweden differ from those of many other European countries, making the processes very complex, particularly in international projects, not to mention the issues of storage and many different systems and programs. Once again, it becomes clear that archivists alone cannot solve this. The research community together with various types of information specialists also needs to find a solution for how to define a project, as well as how to define when it is finished; consensus is needed. Defining a PhD project is quite simple, but how should other, larger projects be defined? When the funding stops or when a larger publication has been made?

Many of the challenges really involve a lack of communication and information, rather than significant obstacles for open access. The gap between department and administration could be addressed by employing archivists or research coordinators. Major advances could be achieved by informing researchers about how to store data, how to name it and which file formats to use. Collaboration between the universities and the public authorities could help solve the problems caused by a lack of metadata standards and infrastructure. In this way, costs could be shared and duplication of effort avoided. Researchers might be more willing to share if they were aware that in cases with sensitive or unique findings, their entire dataset does not need to be made available immediately, nor would it have to be made open at the same time.

Infrastructure, coordination and cooperation

The majority of the respondents believe that the infrastructure is currently unsatisfactory; some going as far as saying that there is none. Actor 1, however, believes that there is quite a lot, but it is not widely known. Actor 1 thinks it is as much a technical as methodological issue. The majority also connect the insufficient infrastructure with a lack of financial resources. The costs of long-term preservation and open accessibility of data should probably be divided between the public authorities and universities. Either the authorities should provide the universities and researchers with funding—or create their own centralized data repository infrastructure in which the researchers could deposit their data.

Analysis through the lens of the Mertonian norms of science

When investigating open research data, one needs to start by looking at what the benefits are and who benefits from it. The results show that both the public and the researchers are expected to benefit from increased availability of research data.

However, it seems that researchers' attitudes vary, and this often (if not always) depends on the discipline. This means that it is important to investigate why a researcher may or may not be interested in archiving and making research data open access. If the researchers' point of view can be understood, it will probably be easier to create tools for open research data and start implementation.

The results of this investigation were analyzed by following the Mertonian norm system (Merton 1973). Merton's four pillars (as described above) represent an ideal based on how science *should* be communicated. Arguing that science is always rational and that the research community is based on shared ownership is a simplification of reality.

Universalism

Those in favor of open research data highlight the accessibility perspective and objectivity; their hope is that open research data can prevent research fraud and verify research results. This approach also goes hand in hand with the rational perspective where increased cooperation and coordination can result in cross-border and interdisciplinary studies that can create new research fields and new exciting discoveries and results that would not have been possible for a single researcher or research field to achieve. At the same time, inclusion requires some form of exclusion. It is impossible to include everyone in a group. In a research area, inclusion is automatically available to those who carry out the same type of research, as they can understand the research conducted.

One of the arguments in favor of open research data is that the public should be able to access publicly funded research. The idea is good and suggests that science is thought of as being accessible and open to inclusion. However, the question is, how well can a third party understand the research? Research often uses concepts and terms without definition, taking it for granted that the reader will understand since the audience is expected to be in the same research field. This means that large parts of the public are nonetheless excluded even if they can physically access the research. If the argument focuses more on increased transparency for the public, it would be better to emphasize that popular scientific articles or summaries should be made available to the public. However, if the concern is about interdisciplinary cooperation, it is important that the data is also made available. To ensure inclusion, the data must be well described, so that researchers from other fields or with other methodological starting points can understand its structure. This is where metadata becomes of utmost importance. However, the crucial question found in the interview responses is "what data should be made available and to what extent?" Some research will not be considered to have any major public interest. If so, is the time it would take to describe the data and the cost of storing it necessary? Also, how do you decide what is interesting and not interesting? Science is a changing process much like people's attitudes and interests. What we consider to be of particular interest today can be completely uninteresting in a hundred years. Today, much attention is paid to the fact that researchers have to decide if their raw data is be worth preserving. But do the researchers really have the ability to decide what can

be historically interesting in 100 years' time? It may not always be the data itself that is interesting but the administrative documentation around the data, how science is communicated and developed during the course of work. Therefore, extended collaboration between researchers and archivists, where the archivists should play a consultative part in what data should be kept and for how long, is essential. Since the archivist may not understand the data itself, a dialogue is needed.

Communism

Interestingly, Merton's principle of communism states that the researcher has no property rights other than the recognition received for the scientific discoveries (Merton 1973, pp. 273–274). This can be contrasted to some respondents' experience that researchers sometimes see the research material as "their" documents. When discussing communism, Merton states that research is a result of cooperation and therefore belongs to the entire research community (Merton 1973, pp. 273–275). Many researchers probably see their results in published form as a product belonging to the entire research community, since recognition can be gained in that way, while also acknowledging other researcher's results. Communism can be applied in the disciplines dealing with large volumes of data that are associated with different registers. Here, increased cooperation can provide large amounts of data to be analyzed by individual researchers. This means that faster gains can be made in the research, and recognition gained for the publication itself by sharing between groups. That is, the data collection in itself is not necessary for the researcher's career opportunities, nor to gain recognition for the research. In other disciplines, the collection itself is of great importance to the individual's research and results. Thus, publishing data may constitute a career obstacle for the researcher if others can use it to conduct their own research and to get ahead. Alternatively, the researcher can be criticized for the method used and how the data has been interpreted. This can make researchers reluctant to share data, considering the reward lower than keeping the material to themselves. Some disciplines work with their data for many years, and the material can be used as the basis for several studies. It is therefore difficult to apply the argument that publishing the data can wait until the researcher is finished and has published the study.

It must be emphasized that researchers do not only cooperate and hope to get recognition for a particular discovery—they also compete for the same money. This makes it conceivable that you do not want to make it possible for colleagues to use your data for their own studies.

Disinterestedness

The majority of researchers' drive is certainly a thirst for knowledge, and many have no desire for personal gain. A research career is usually not associated with any major financial benefit: the individual gains glory, while the research community as a whole gains knowledge. This is not to say that individuals cannot be driven by personal gain. Science is not just research and seeking knowledge, it is also a

social activity. In order for the research to be beneficial to society, it must be spread, which requires infrastructure and social networks. The researcher cannot be completely isolated if the aim is to disseminate what has been discovered. This means that social games and conflicts can arise in a struggle for resources and influence, as there are limited resources and limited places at a university. This means that there will always be individuals who seek to work for their own improvement, to gain as much recognition as possible in order to receive more funding. Furthermore, there are certainly individuals who engage with research in the hope of eternal glory and fame. In summary, attitudes to sharing data are linked to the discipline and the benefits gained from sharing or not sharing data.

Organized Skepticism

The norm organized skepticism states that scientific claims should be tested and reviewed, which is the case when a doctoral student submits a dissertation for examination or a researcher submits an article for peer review. This means that researchers must be able to accept that the results are reviewed. The idea rests on two pillars: the best motivated thesis is considered valid until there is another with a better supported conclusion, but it is also a way to prevent research fraud. Therefore, research must be considered as transparent and objective. This is likely true in most cases, but research fraud does happen.

As mentioned, verifying research and discovering fraud will be easier if data is openly available. If the examiner needs to ask the researcher for the data, it could take a long time and the reviewer may not receive all the data required. This could also happen if the data is made available in an open database, but if the researcher knows from the outset that the material will be more exposed the inclination to hide information may decrease. In addition, finding research fraud is not the only issue. One of our respondents, a researcher in medicine, planned to make his research data openly available to demonstrate the importance of the research. The researcher believed that some colleagues opposed and detracted from the results achieved. The researcher explained that they had published many articles, and that there was no one in the same field who distrusted their research, it was other doctors at the same institution in another specialization. As a defense, the researcher wanted to provide open access to the data to create transparency for its own sake. The researcher had something to gain from making his research data publicly available.

The Mertonian norms can also be applied to the ideas of open research data. The difference is that the Mertonian norms focus on the research community, while the advocates of OA and open research data also demand inclusion of the public. Most of the ideas of the Mertonian norms are in fact what the advocates of open research data propose, illustrated in Table 4.

It could be assumed that archivists would be advocates for open research data; however, none of the four archivists interviewed were willing to assume responsibility for an open data archive. The lack of knowledge among many and the long tradition of paper archiving are possible sources of this reluctance; an open data archive was regarded as difficult and time-consuming. The university archivists must

Table 4 The Mertonian norms applied to the ideology of open research data

Universalism	Communism	Disinterestedness	Organized Scepticism
•Openness •Accessibility	•Research conducted with taxpayers money should be openly available to the public •Open for everyone, not only academics	•Increased transparency & openness could create new discoveries and disciplines benefitting both the public and the research community	•Verification •Validation

take a more active role in collecting, organizing and educating researchers in archiving data and making sure that they are an obvious partner for the researchers. They should be able to give advice to the researchers both when it comes to archiving as well as open research data, and argue their case when procuring infrastructure and information systems. Archivists also need to increase collaboration with other information specialists, such as metadata specialists, system scientists, legal practitioners and research coordinators. A research data office is probably a good start, but it is important that the archivists take a leading and active role. The archivist should be an expert on information management, and as such should be an obvious partner to lead the open research data movement. It is troubling that none of the archivists considered it their responsibility or in their interest.

Is there a connection to the archive?

Archiving could be said to consist of the following pillars: accessibility, preservation, verification and reusability, which is in line with the Mertonian norms (Merton 1973, pp. 268–279). This section comments on the four Mertonian norms from an archival perspective. There is a close relationship between archival records and open data; sometimes records are synonymous with open data (Borglund and Engvall 2014). The core essence found in the Mertonian norms of Universalism and Communism is very similar to the core drivers for keeping archives. In Sweden, the whole concept of the archives is found in the Swedish Constitution: all citizens have the right to access public records (Tryckfrihetsförordning 1949:105, 1949). It is implicit that the citizens own the public records. In Table 4, where the Mertonian norms are applied to open research data, this is even clearer. Openness and accessibility for everyone are also the basis of public archives in Sweden.

When it comes to the Mertonian norm Disinterestedness, the focus is on science working for the common good. Archival institutions can be characterized as the places where evidential records are kept and stored which also could be related to Organized Skepticism which relates to peer reviewing and archive material as evidences.

When the records are made accessible to the public, they can also be verified and any errors can be identified. However, in an archive the evidential value is not about the content, it is about the completeness and transparency of the record itself, which in turn ensures the evidential value of the records. This differs from science, which is often more focused on the content (evidence from which the results are drawn).

Both Disinterestedness and Organized Skepticism cover the secondary value of records presented by Schellenberg (1956/1998) where records can be used for other purposes than they were created. Disinterestedness and Organized Skepticism are also very relevant when applying the categorization of secondary use introduced by Shepherd and Yeo (Shepherd and Yeo 2003), where the secondary use can be either for business use, accountability use or cultural use. For example, verification and validation are closely related to the "accountability" purpose of records. The Mertonian norm of Disinterestedness deals with innovation and new knowledge, a perspective very similar to the fourth dimension of Pluralization in the Records Continuum Model (A deeper explanation of pluralization is found in Upward 2000).

The science world is facing a paradigm shift for which it is unprepared: it has not sufficiently adapted to the digitization process. Frank Upward and colleagues write that in order for a paradigm shift to take place, a crisis situation must have arisen in the archives and in the information management sector (Upward et al. 2013). This article traces such a crisis in the university archives, on what we should characterize as data and how we should collect and process the generated information (Upward et al. 2013). Even the question of what to save and how to save it has emerged as a major challenge and concern in the interviews. Respondents' recommendations often match what has already been covered in previous research. The problem is that recommendations concerning what needs to be done have been formulated over a long period of time, but there is little or no development and implementation.

Conclusion

Further research is needed on how to describe the research data and larger collaboration is needed. This has been a small study and further research is needed to investigate whether the findings are applicable in an international context. Applying the Mertonian norms provides a new perspective on the problem, which can hopefully contribute to sharpening the edge of archival science in continued research.

References

Åhlfeldt J, Johnsson M (2015) Research libraries and research data management within the humanities and social sciences. Retrieved from University Library Lund University http://lup.lub.lu.se/record/5050462

Björklund C, Eriksson J (2007) Forskningsdata i öppna arkiv och universitetsarkiv: en förstudie vid Göteborgs universitet, Lunds universitet och Sveriges Lantbruksuniversitet. Retrieved from https://gupea.ub.gu.se/handle/2077/7379. English: Research data in open archives and university archives: a feasibility study at the University of Gothenburg, Lund University and the Swedish University of Agricultural Sciences

Bohlin A (1997) *Offentlighet & sekretess i myndighets forskningsverksamhet*. Stockholm: Riksarkivet. English: Publicity and confidentiality in State authorities research activities

Borglund E, Engvall T (2014) Open data?: data, information, document or record? Rec Manag J 24(2):163–180. https://doi.org/10.1108/RMJ-01-2014-0012

Borgman CL (2012) The conundrum of sharing research data. J Am Soc Inform Sci Technol 63(6):1059–1078. https://doi.org/10.1002/asi.22634

Borgman CL (2015) Big data, little data, no data : scholarship in the networked world. The MIT Press, Cambridge

Brinck J, Leuhusen K (2013) *Att arkivera forskning—Forskares inställning till arkivering och återanvändning av forskningsdokumentation*. Master thesis. Lunds Universitet, English: To archive research: scientists feelings to archiving and re-use of research documentation

Childs S, McLeod J, Lomas E, Cook G (2014) Opening research data: issues and opportunities. Rec Manag J 24(2):142–162. https://doi.org/10.1108/RMJ-01-2014-0005

Corti L, Thompson P (2007) Secondary analysis of archived data. In: Seale C, Gobo G, Gubrium J, Silverman D (eds) Qualitative research practice. SAGE, London, pp 297–313

Corti L, Witzel A, Bishop L (2005) On the potentials and problems of secondary analysis. An Introduction to the FQS special issue on secondary analysis of qualitative data. Forum Qual Sozialforschung/Forum: Qual Soc Res 6(1):1–6

Corti L, Eynden VVD, Bishop L, Woollard M (2014) Managing and sharing research data : a guide to good practice. SAGE, Los Angeles

Doorn P, Tjalsma H (2007) Introduction: archiving research data. Arch Sci 7(1):1–20. https://doi.org/10.1007/s10502-007-9054-6

Faniel IM, Jacobsen TE (2010) Reusing scientific data: how earthquake engineering researchers assess the reusability of colleagues' data. Comput Supported Cooperative Work (CSCW) 19(3):355–375. https://doi.org/10.1007/s10606-010-9117-8

Faniel IM, Zimmerman A (2011) Beyond the data deluge: a research agenda for large-scale data sharing and reuse. Int J Digi Curation 6(1):58–69. https://doi.org/10.2218/ijdc.v6i1.172

Faniel IM, Kriesberg A, Yakel E (2012) Data reuse and sensemaking among novice social scientists. Proc Am Soc Inf Sci Technol 49(1):1–10. https://doi.org/10.1002/meet.14504901068

Faniel IM, Kriesberg A, Yakel E (2016) Social scientists' satisfaction with data reuse. J Assoc Inf Sci Technol 67(6):1404–1416. https://doi.org/10.1002/asi.23480

Finansdepartementet (2009) Vidareutnyttjande Av Handlingar [Elektronisk Resurs] : Genomförande Av Psi-Direktivet. [Re-use of Documents [Electronic resource]: Implementation of the PSI Directive]. Stockholm: Finansdepartementet Fritze [distributör]. English: Re-use of records: implementation of PSI directive

Fink AS (2000) The role of the researcher in the qualitative research process. A potential barrier to archiving qualitative data. Forum Qual Sozialforschung. https://doi.org/10.17169/fqs-1.3.1021

Fransson J, du Toit M, Tapia Lagunas P, Kjellberg S (2016) Tillgängliggörande av forskningsdata : nulägesbeskrivning samt rekommendationer för uppbyggnad av stöd till högskolans forskare. Retrieved from Malmö högskolas bibliotek. English: Provision of research data: status report from Malmö university library

Fransson J, Kjellberg S, Sedvall K, Tengstam O, Tapia Lagunas P, du Toit M, Widmark J, Höie H (2017) På väg mot ett forskarstöd kring hantering och tillgängliggörande av forskningsdata : Slutrapport från biblioteksprojektet Tillgängliggörande av forskningsdata 2016. Retrieved from Malmö högskolas bibliotek. English: Towards research support concerning use and provision of research data: final report from the library project Provision of research data 2016

Grant R (2017) Recordkeeping and research data management: a review of perspectives. Rec Manag J 27(2):159–174. https://doi.org/10.1108/RMJ-10-2016-0036

Hammersley M (1997) Qualitative data archiving: some reflections on its prospects and problems. Sociology 31(1):131–142. https://doi.org/10.1177/0038038597031001010

Hedlund T, Annikki R (2006) Open access publishing in Finland: discipline specific publishing patterns in biomedicine and economics. Paper presented at the ELPUB2006 conference on Electronic Publishing, Bansko, Bulgaria, 2006

Hermerén G, Gustafsson B, Petterson B (2011) God forskningssed. Stockholm: Vetenskapsrådet. English: Good research custom

http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-53292. English: Open access for humanistics and legal scholars: a survey of publishing policies and praxis within Nordic edition. Final report

Kling R, McKim G (1999) Scholarly communication and the continuum of electronic publishing. J Am Soc Inf Sci 50(10):890–906. https://doi.org/10.1002/(SICI)1097-4571(1999)50:10%3c890:AID-ASI6%3e3.0.CO;2-8

Mauthner NS, Parry O, Backett-Milburn K (1998) The data are out there, or are they? Implications for archiving and revisiting qualitative data. Sociology 32(4):733–745

McDermott R (2014) Research transparency and data archiving for experiments. Polit Sci Polit 47(1):67–71. https://doi.org/10.1017/s1049096513001741

McLeod J (2012) Thoughts on the opportunities for records professionals of the open access, open data agenda. Rec Manag J 22(2):92–97. https://doi.org/10.1108/09565691211268711

Merton RK (1973) The sociology of science : theoretical and empirical investigations. University of Chicago Press, Chicago

Meyer Lundén K (2008) Forskare och parallellpublicering : forskares syn på, kunskap om och användning av den 'gröna' vägen till open access. (414 Student thesis). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-101869DiVA database. English: Scientists and parallel publishing: scientists view on, knowledge and use of the "green" way to open acess

Palmer CL, Weber NM, Cragin MH (2011) Analytic potential of data: assessing reuse value. Paper presented at the proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries, Ottawa, Ontario, Canada

Parry O, Mauthner NS (2004) Whose data are they anyway? Practical, legal and ethical issues in archiving qualitative research data. Sociology 38(1):139–152. https://doi.org/10.1177/0038038504039366

Piwowar HA (2011) Who shares? Who doesn't? Factors associated with openly archiving raw research data. PLoS ONE 6(7):e18657. https://doi.org/10.1371/journal.pone.0018657

Piwowar HA, Day RS, Fridsma DB (2007) Sharing detailed research data is associated with increased citation rate. PLoS ONE 2(3):e308. https://doi.org/10.1371/journal.pone.0000308

Schellenberg TR (1956/1998) Modern archives: principles and techniques (Repr. ed.). SAA, Chicago (Original work published 1956)

Serra LEC (2014) The mapping, selecting and opening of data: the records management contribution to the Open Data project in Girona City Council. Rec Manag J 24(2):87–98. https://doi.org/10.1108/RMJ-01-2014-0008

Shepherd E, Yeo G (2003) Managing records a handbook of principles and practice. Facet Publishing, London

The Royal Society (2012) Science as an open enterprise. Retrieved from https://royalsociety.org/-/media/policy/projects/sape/2012-06-20-saoe.pdf

Tryckfrihetsförordning, 1949:105 C.F.R (1949) English: Swedish freedom of the Press Act

Upward F (2000) Modeling the continuum as paradigm shift in recordkeeping and archiving processes, and beyond—a personal reflection. Rec Manag J 10(3):115–139

Upward F (2005) The records continuum. In: McKemmish S, Piggott M, Barbara R, Upward F (eds) Archives: recordkeeping in society. Charles Sturt University, Centre for Information Studies, Wagga Wagga, pp 197–222

Upward F, Reed B, Oliver G, Evans J (2013) Recordkeeping informatics: re-figuring a discipline in crisis with a single minded approach. Rec Manag J 23(1):37–50. https://doi.org/10.1108/09565691311325013

Vetenskapsrådet (2015) Förslag till nationella riktlinjer för öppen tillgång till vetenskaplig information. Stockholm: Vetenskapsrådet. English: the research council (2015), Proposal for national guidelines for open access to scientific information

Vines TH, Andrew RL, Bock DG, Franklin MT, Gilbert KJ, Kane NC, Yeaman S (2013) Mandated data archiving greatly improves access to research data. FASEB J 27(4):1304–1308. https://doi.org/10.1096/fj.12-218164

Wallis JC, Rolando E, Borgman CL (2013) If we share data, will anyone use them? Data Sharing and reuse in the long tail of science and technology. PLoS ONE 8(7):e67332. https://doi.org/10.1371/journal.pone.0067332

Williamson K, Johanson G (eds) (2013) Research methods: information, systems and contexts. Tilde University Press, Melbourne

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Charlotte Borgerud MA in history and M.Sc. in archival science is an archivist at one of Sweden's most high ranked universities. Charlotte has been a guest lecturer at Mid Sweden University for several years. Currently, she is working at the Swedish Association of Local Authorities and Regions.

Erik Borglund Ph.D. in computer and system science, is an associate professor at the department of information systems and technology at Mid Sweden University. Erik is involved in teaching on undergraduate and on graduate level in archival science and in information systems. His research interests cover the domains of digital recordkeeping, recordkeeping informatics, information systems in crisis management, information systems design and computer-supported cooperative work.