

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Digital Libraries: Supporting Open Science	
Series Title		
Chapter Title	The OpenUP Pilot on Research Data Sharing, Validation and Dissemination in Social Sciences	
Copyright Year	2019	
Copyright HolderName	Springer Nature Switzerland AG	
Corresponding Author	Family Name	Luzi
	Particle	
	Given Name	Daniela
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Institute for Research on Population and Social Policies, National Research Council
	Address	Rome, Italy
	Email	d.luzi@irpps.cnr.it
	ORCID	http://orcid.org/0000-0003-3404-3271
Author	Family Name	Ruggieri
	Particle	
	Given Name	Roberta
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Institute for Research on Population and Social Policies, National Research Council
	Address	Rome, Italy
	Email	
	ORCID	http://orcid.org/0000-0002-7051-4977
Author	Family Name	Pisacane
	Particle	
	Given Name	Lucio
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Institute for Research on Population and Social Policies, National Research Council
	Address	Rome, Italy
	Email	
	ORCID	http://orcid.org/0000-0001-6764-161X

Abstract



The paper presents the results of a pilot carried out within the European project OpenUp (Opening up new methods, indicators and tools for peer review, dissemination of research results and impact measurement). Aim of the pilot is to investigate the applicability of peer review and/or Open Peer Review (OPR) to datasets in disciplines related to Social sciences. Main emphasis is given to the characteristic and features of data sharing and validation in this heterogeneous scientific field, thus providing the basis for the selection of the community chosen for the pilot. Indications emerging from the analysis of the interviews carried out in the pilot can drive the adoption of data quality assessment, and hence peer review, as well as provide some principles that can incentivize other scientific communities to share their research data.

Keywords
(separated by '-')

Data quality - Open data - Open dataset review and validation - Open Peer Review (OPR) - Social sciences



The OpenUP Pilot on Research Data Sharing, Validation and Dissemination in Social Sciences

Daniela Luzi^(✉) , Roberta Ruggieri , and Lucio Pisacane 

Institute for Research on Population and Social Policies,
National Research Council, Rome, Italy
d.luzi@irpps.cnr.it

Abstract. The paper presents the results of a pilot carried out within the European project OpenUp (Opening up new methods, indicators and tools for peer review, dissemination of research results and impact measurement). Aim of the pilot is to investigate the applicability of peer review and/or Open Peer Review (OPR) to datasets in disciplines related to Social sciences. Main emphasis is given to the characteristic and features of data sharing and validation in this heterogeneous scientific field, thus providing the basis for the selection of the community chosen for the pilot. Indications emerging from the analysis of the interviews carried out in the pilot can drive the adoption of data quality assessment, and hence peer review, as well as provide some principles that can incentivize other scientific communities to share their research data.

Keywords: Data quality · Open data · Open dataset review and validation · Open Peer Review (OPR) · Social sciences

1 Introduction

The study presented is part of OpenUP [1] (OPENing UP new methods, indicators and tools for peer review, dissemination of research results and impact measurement), a European funded project that addresses the currently transforming science landscape on innovative peer review, dissemination and impact measuring. The project test the achieved results in a set of seven pilots [2, 3]. They are related to the three project's pillars and are applied to specific research areas and communities: arts and humanities, social sciences, energy and life sciences.

The results presented in this paper are related to the pilot study that investigated the applicability of peer review and/or OPR to datasets in Social sciences. The pilot aims at identifying strong and weak elements in the process of dataset review and validation and intends to outline best practices that facilitate transparency of the process as well as data dissemination, reliability and reuse.

In particular the paper reviews data sharing and evaluation practices in Social sciences, on which the selection of the pilot community is based, and reports on the interviews with the management team of the selected community. Lessons learned that

can help identifying requisites and best practices for OPR of research data are reported in the conclusions.

2 Methods

To reconstruct the context of data sharing and evaluation in Social sciences, a landscape scan was performed to identify relevant literature, guidelines and scientific communities' practices. This first phase provided the basis for the selection of potential communities to be involved in the pilot, representative for this heterogeneous scientific field, as well as the identification of specific characteristics and problematic issues. For these reasons the landscape scans focused on surveys related to the researchers' motivation and constraints towards data sharing and on desk research enabled the identification of types of research data, dataset providers and modes of validating and sharing/publishing datasets in Social sciences.

After considering the involvement of different research communities in Social sciences, a collaboration was formalized with the Human Mortality Database (HMD). HMD is an open database that provides detailed, consistent and high quality data to researchers, students, journalists, policy analysts, and others interested in the history of human longevity and its prospects for the future (<https://www.mortality.org/>). HMD constitutes a good example of a well-known source of information providing open access to data in the scientific community of demographers. Moreover, the collaborative organization of HMD, as well as the high number and variety of data users worldwide, represented a good premise to get further insights into data managing practices and reuse.

The second phase of the analysis was devoted to the development of an interview schema that aimed at exploring the following HMD features: origin, motivations and organizational features of the scientific community, data quality assessment process, opinion on Open access of data. A further step of the project concerned the development of a users' survey carried out in collaboration with HMD to get insights into users' perspectives and feedback on data reuse. The results of the HMD survey are not included in this paper and will be presented in further publications. The interviews with the selected community were conducted on the 31st of January and 1st of February 2018 at the Max Planck Institute for Demographic Research in Rostock, Germany. They were performed according to interviewees' role in HMD, and for the selection of interviewees we also considered gender balance. The two directors, two researchers in their role of country responsible (in charge of analysing data for specific countries) were interviewed. These interviews covered the majority of HMD staff (4 out of 7). The summary content presented in this paper was revised, commented and approved by the interviewees.

The paper presents the results of the landscape scans related to the main characteristics of data management and validation in Social sciences. Moreover, a summary of the interviews is provided highlighting in particular scientific motivations for data sharing, organizational features as well as the quality assessment process as important success factors for data publishing.

3 Data Sharing and Evaluation Practices in Social Sciences

Similar to peer review of publications, data peer review is a quality assessment process of a dataset performed by experts in the field. Data quality assessment is a complex process that has to consider the different phases of the data lifecycle, starting from the development of a Data Management Plan (DMP) at the initial stage of a scientific project to the publication of its results. Data publication (some authors speak about Publication with capital letter, [4, 5]) that undergoes a peer review process validating data quality is currently performed in data journals and data repositories. However, analyses by Candela et al. [6], and Carpenter [7] show that there is room for improvement, as peer review activities in data journals vary widely and are mostly focused on metadata rather than data themselves, aiming at assessing the documentation and metadata description that facilitates data reuse. Assante et al. [8], in the analysis of generalist repositories (Zenodo, Dryad, Figshare etc.), also come to the conclusion that different criteria and quality control mechanisms are implemented based on various policies and/or guidelines. Best practices of data validation can be

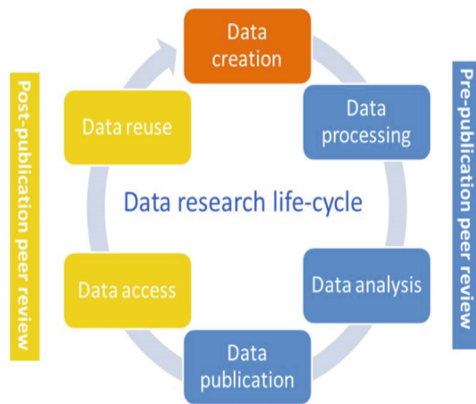


Fig. 1. Pre and post-publication peer review within the data research life-cycle

taken from the publication in trusted data repositories [9], in which the data quality control can be considered a form of review by experts in the field carried out in a pre-publication phase. This form of pre-publication review is then confirmed by the use of the dataset by data consumers in the post-publication phase through formal citations and/or statistics of use (see Fig. 1).

3.1 Researchers' Motivation and Constraints

Several surveys have investigated researchers' attitudes to analyse barriers and facilitators towards data sharing. General surveys allow us to compare social scientists' attitudes with other disciplines. Tenopir et al. [10] found that social scientists have a

lower propensity to share the data they produce compared to the STEM researchers. These results have been confirmed in other surveys [11, 12]. This may depend on the nature of data, especially when qualitative data are involved, privacy and confidential issues, lack of technical standards and easy-to use platforms. Social scientists share the same concerns as scientists in other disciplines, such as not being recognised for making the data available, misuse of data, costs and time-consuming activities required.

Concerning peer review, Kratz and Strasser's [13] survey shows that researchers of different disciplinary fields are still unsure on how data peer review should work and in which context it should occur, even if they expect that data in repositories are subject to validation. The OpenUP survey [14] indicates that social scientists together with ICT researchers are the ones that are less satisfied with the traditional peer review process of publications.

3.2 Types of Dataset Providers

A specific characteristic in this field is that a significant portion of data is produced for purposes other than research [15]. These are data created by governmental bodies that have to comply with transparency regulations (such as the UK Freedom of information act) and make data they collect publicly available. Examples of these data comprise census figures, cohort and longitudinal studies, cross national surveys, economic indicators, etc. Among governmental bodies it is worth mentioning the data produced by national statistical offices that apply standard procedures to collect and process data, provide detailed supplementary documentation to describe the dataset, and also guarantee long-term preservation. These data collection constitutes trustworthy information, on which many other studies are based representing an important data source not only for social scientists.

While these sources of information can be compared to the big data produced by STEM, long-tail data are produced by social scientists to investigate local phenomena in small collaborative groups, often within interdisciplinary projects or individually. They are usually facing privacy issues that make the dataset sharing more complex.

3.3 Modes of Sharing/Publishing Datasets

In Social sciences, data publication model is mostly related to dataset submission in a data repository. In fact, there is only one data journal covering Humanities and Social sciences: "Research Data Journal (RDJ)" [16]. It was created by DANS [17] in 2016 with the aim to increase the visibility of data stored in the archive and to provide more extensive and detailed documentation. This journal conforms to well-established data journals in other disciplines such as Earth System Science Data, Geoscience Data Journal, and Scientific Data, it assigns a DOI to the article and provides the related DOI assigned to the dataset stored. On the other hand, DANS archive does not provide a standard description to cite the article. Currently eleven data papers have been published and three papers refer to the field of Social sciences.

Considering trusted data repositories, their main feature is that of data centres that act at national level as main information sources in this field. Worth mentioning are the UK Data Archive [18], GESIS [19] and DANS. The majority of these national

centralized data centres are also part of two consortia, CESSDA [20] at a European level and ICPSR [21] at international level. These consortia provide a single access to international data and also develop and coordinate initiatives on standards, protocols and best practices to support data management and dissemination. Most of them provide access to data produced by governmental bodies and by research groups.

3.4 Modes of Validating Datasets

The above-mentioned data repositories are certified by the Data Seal of Approval [22] that has identified 16 requirements based on 5 criteria: data availability on the Internet, accessibility (clear rights and licenses), usability (format), reliability and identification of dataset through a persistent identifier. Note that these also correspond to the criteria used to evaluate the data themselves, similar to the FAIR principles. Given that data validation represents an iterative process that encompasses the entire research lifecycle, these trusted repositories provide guidelines on how to develop a data management plan at the very beginning of a research to assure data quality. Moreover, they require data producers to establish copyright and appropriate licenses, to use proper data formats and metadata schemas to facilitate access and reuse.

Trusted data repositories provide guidelines and/or template for a correct data ingestion according to the metadata schema of DDI [23], a standard supported by the Social sciences community that facilitate data replication and/or reproduction. They assure long-term data preservation and curation, develop data discovery tools (such as landing pages; [24]), suggest users a data citation format that acknowledge data provenance. Moreover, there are different initiatives to increase data re-use and make it traceable. The suggestion of a data citation format represents an important means to support data citations that are an indirect appraisal of the quality of the dataset at the post-publication phase. Some trusted repositories have adopted tools to track data use. For instance, DANS provides data users with a validation template to rank data set available in EASY: users can provide the rating (up to five stars) to data quality, quality of documentation, completeness of the data, consistency, structure and usefulness of the file format [25].

4 Interviews with the Human Mortality Database Management Team

The scientific community that manages the Human Mortality Database (HMD) was selected for the pilot as it fulfilled the requirements set up in the OpenUP methodology, and not least, for their willingness to an active collaboration on the analysis. The Human Mortality Database documents the longevity revolution of the modern era and facilitates research into its causes and consequences providing open data on human mortality. HMD was launched in 2002 as a result of a collaborative project that involved researchers of the Department of Demography at the University of California Berkeley (UCB) and the Max Planck Institute for Demographic Research (MPIDR). The following paragraphs summarise the key points of the interviews carried out with

the two directors and two country specialists (CSs). As previously mentioned, its content was revised and approved by the interviewees.

4.1 Origin, Motivations and Organisational Features

Two previous relevant experiences guided the development of the database: the Kannisto-Thatcher Database on Old Age Mortality (KTD) at the MPIDR and the Berkeley Mortality Database (BMD), founded by John Wilmoth at UCB. Both experiences were concerned with what was at that time an emerging phenomenon of low mortality at young and adult ages, falling mortality at old ages, and greater survival to an advanced age, leading to a potential increase in the number of people exposed to degenerative diseases, which are difficult to treat or prevent. To understand this phenomenon, it was necessary to analyse and model longevity and survival of humans with a special emphasis on advanced (frontier) age over a long period of time. This research needed reliable data at international level providing long-term and continuous series without gaps, running up to the highest ages, providing fine details according to age, time, and cohort dimensions, ensuring sufficient quality and comparability across time and populations. HMD was therefore developed to answer this scientific question providing a methodology based on the previous mentioned experiences as well as freely available high-quality data [26].

The two HMD directors explained that *“the collaboration was originally, (and still is), based on a small, very well-established group of internationally based demographers who were willing to serve the scientific community interested in demographic studies”*. The workload is equally distributed among the team that comprises CSs who have high-level competences on demographic development of a set of specific countries and are responsible for collecting and analysing data from the related national statistical offices. Other tasks comprise the development of computer codes, which are also made freely available to the end user who wants to reproduce the analysis, as well as the management of the website. Strong collaboration pertains the data quality process performed before data are publicly available, which constitutes a form of internal pre-publishing peer review process. During the interview the two directors agreed that *“trust among the team and scientific curiosity are the drivers of this successful cooperation”*, that only recently was formalized by a Memorandum of understanding.

4.2 Goals and Main Features of the Database

The main goal of HMD is to support research on human mortality and longevity providing open data on 39 countries and some sub-areas and sub-populations with series starting as early as 1751 (i.e. Sweden) and covering more than 100 years for 16 populations. Birth and death counts are generally based on data from national vital registration systems, while data on population are based on the national census and estimates between censuses. However, differences may exist among countries in the periodicity of census, methods and definition used as well as in data format. Moreover, some countries have experienced changes in their territorial boundaries, have suffered substantial loss during war periods and/or faced substantial consistent migration over the period covered by HMD. For these reasons, as underlined by the two directors,

HMD has developed a methodology to produce detailed death counts and population estimates, to correct mortality estimates at old ages, and to build high quality life tables (as described in detail in the Methods protocol). “All HMD data are prepared using this standard methodology. This assures comparability in time and across countries”. The two County specialists explained, that “when special methods are needed to accommodate issues in data availability, this is documented in the country-specific documentation as well as reported in summary tables” [27]. Country-specific details related to the data quality and statistical system in each country are therefore documented in the country-specific Background and Documentation file accessible from each country webpage. The application of these thorough procedures, “the punctual explanation of the estimations and refinements of data sources make this database different from other sources providing mortality rates”. These procedures guarantee a uniform analysis of raw data, facilitating the comparability across time and space, while the detailed documentation and the availability of source data allow end user to reproduce the analysis. The HMD team has also developed software code that guide them in the evaluation of data quality as well as software packages that facilitate end user to import and working with HMD data. These tools are freely available to end users along with technical reports explaining how to use these scripts [28]. This is another value-added feature of HMD.

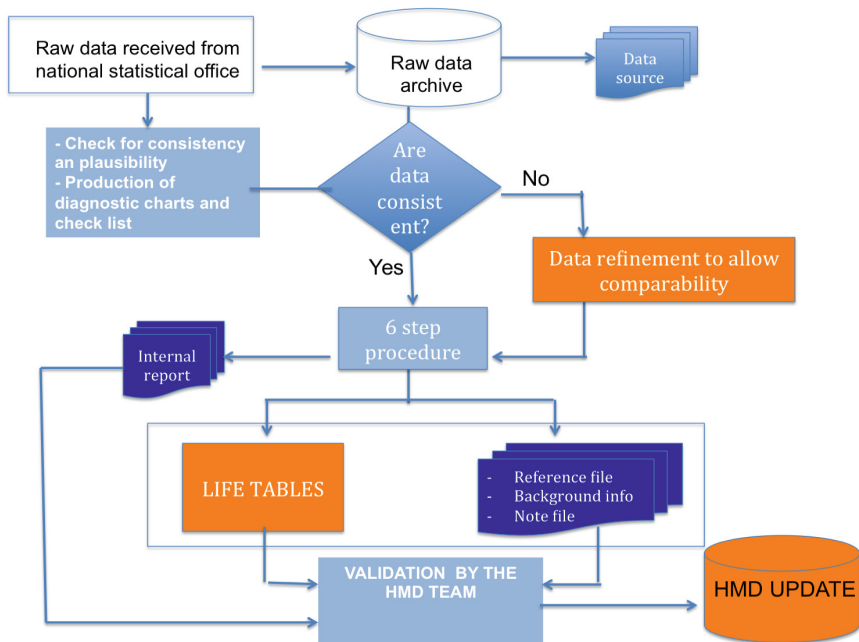


Fig. 2. Data quality assessment process

4.3 The HMD Data Quality Assessment Process

The HMD team has developed a set of procedural steps to ensure data quality. This important topic was addressed in the interviews with the two directors and particularly explored in the interviews with the CSs. An activity diagram that reconstructed the workflow of the activities performed before data publication was presented to the CSs and discussed to have further insights on the procedures adopted to assess data quality. This intended to explore whether collaborative activities resembling a peer review process could be tracked in HMD data quality assessment. A high-level description resulting from the interviews is provided in Fig. 2.

During the interviews the CSs explained that each country or area is assigned to an individual researcher, a CS, who maintains a close relationship with a local expert generally at national statistical offices, and has an extensive knowledge of the population dynamics as well as how data are collected at national level. A CS is responsible for the first quality checks that evaluate consistency and plausibility of input data, prepares pre-calculation file (Lexis files) and analyses the results on the basis of a pre-defined data quality checklist and diagnostic charts that help him/her to explore unusual fluctuation and/or any other issues in data sources. The results of this analysis are shared within the HMD community via an internal report and are the basis for the application of the six-step procedure to produce the complete data series (exposures to risk, death rates, life expectancy and other life tables). Before data are published, the HMD team perform an additional phase of validation. These activities are crucial especially when a new country has to be included in HMD. However, they constitute a routine procedure every time data are updated. *“In cases of unexpected changes in national statistical systems or in regimes of national statistical registration, the updating procedures are non-trivial”*.

All steps in the computing of data analysis are documented in detail and made available to end users in the different files (*Background and documentation, Data source and Explanatory Notes*). According to the CSs interviewed, this is the distinctive feature of HMD: *“Data refinements and harmonization that allows comparison across countries are documented in detail so that researchers in this field are aware of possible problems in the data and know how these issues have been solved”*.

4.4 Opinion on Open Access of Data and Peer Review

HMD management team declared that open access and open data in particular are very important for the development of demographic studies. Although they have no official statements on open policy, since its beginning, HMD provided open access data, based on a user agreement indicating that the data in the HMD are provided free of charge to all individuals who request access to the database” [29]. Moreover, users are required to cite the database in their publications, following the citation guidelines provided by HMD [30]. Citations tracked through Google scholar are also reported in the website, and further steps to improve their collection are going to be planned in the next future.

When asked about long preservation of data, it emerged that the two HMD directors are dependent on funds. At the moment MPIDR support their activities (*“MPIDR researchers are allowed to spend half of their work time on HMD”*), while the UCB

team has to provide its own funds. A clear commitment of the organisation would therefore be very important and would also mean a clear recognition of their activities. Between the lines, it emerged that publication of scientific papers are generally considered more important than managing a database. In their opinion, *“the analysis of data, their quality check is not only a service for the community of reference but is a researcher activity in itself.”* The majority of the interviewees has heard about open review of journals but has little knowledge on all its traits. If they see a similarity with peer review of data, this is associated in particular with transparency as a means of reconstructing the methods and procedures used for the data analysis.

5 Lessons Learned

Some important indications emerged from the analysis of the interviews that can drive the adoption of data quality assessment, and hence peer review, as well as some principles that can incentivize other scientific communities to share their research data. As stated by the HMD interviewees *“the guiding principles to create an open access database were: comparability, flexibility, accessibility and reproducibility”*. *Comparability* was reached using a uniform, scientific methodology to calculate the various statistics of the 39 countries included in the database. *Flexibility* was achieved in the analysis of results using a uniform set of procedures for each population, but at the same time giving significant attention to each population in terms of its history and socio-political development. This is also reflected in the available formats of output data series. This is achieved thanks to the experiences and knowledge CSs, that is persons in charge of collecting data from a specific number of countries, who interact with statistical offices, check data consistency and provide population statistics together with a country report that explains specificity and motivation of analysis. *Accessibility* was guaranteed from the beginning by free of charge access of data, as well as by the provision of data in an open, no-proprietary format. *Reproducibility* is provided by the reconstruction of the data lifecycle that includes the availability of raw data, the method applied, the related results as well as the explanatory documentation. One of the main successful features of HMD is its transparent way of data managing and sharing that has two central phases of data validation. The first one is carried out by the CSs, who analyse the raw data according to a common predefined checklist that verifies consistency and plausibility of data. The second one is carried out in a collaborative way within the HMD team that validate the statistics before their publication, each time the database is updated.

Moreover, another successful component of HMD was its collaborative approach that is based on a strong scientific interest in the field as well as on the trust among the involved community that only recently has formally signed a Memorandum of understanding. The interviews also highlighted some indications that confirm some concerns already mentioned by other surveys. Interviewees stressed the importance of having a strong commitment of the organization in supporting the development of data infrastructures. This pertains different aspects: a long-term financial support (beyond the project duration), a policy endorsement on open data as well as a formal recognition of scientists for the efforts in data curation and quality assurance.

Implementation of other pilots should be further promoted in different sub-disciplines of Social sciences to get a deeper insight into sharing and evaluation practices of research data, focusing in particular on the procedures that document data validation and scientific assessment. It is necessary to promote transparency in the process of data evaluation, similar to the one adopted by HMD, so to facilitate the reproducibility of the research.

Acknowledgments. This study is part of the Horizon 2020 OpenUP project. Grant agreement no. 710722. The authors acknowledge the support and the collaborative efforts of the Human Mortality Database management team, namely Magali Barbieri (University of California, Berkeley and INED, Paris), Vladimir Shkolnikov (Max Planck Institute for Demographic Research (MPIDR) and Dmitri A. Jdanov, Head of the Laboratory of Demographic Data at MPIDR. A great thanks goes to our CNR colleague Cristiana Crescimbeni for the valuable technical support during the OpenUP Pilot.

References

1. OpenUP. <http://openup-h2020.eu/>. Accessed 31 July 2018
2. Vignoli, M.: Project OpenUP-Deliverable D6.2 – Interim Use Case Evaluation Report, 30 November 2017. http://openup-h2020.eu/wp-content/uploads/2017/01/OpenUP_D6.2_Interim-Use-Case-Evaluation-Report.pdf
3. Blümel, C., et al.: Project OpenUP-Deliverable D6.3 – Final Use Case Evaluation Report, 14 September 2018. http://openup-h2020.eu/wp-content/uploads/2018/09/OpenUP_D6.3_Final-Use-Case-Evaluation-Report.pdf
4. Lawrence, B., Jones, C., Matthews, B., Pepler, S., Callaghan, S.: Citation and peer review of data: moving towards formal data publication. *Int. J. Digital Curation* **6**(2), 4–37 (2011). <https://doi.org/10.2218/ijdc.v6i2.205>
5. Mayernik, M.S., Callaghan, S., Leigh, R., Tedds, J., Worley, S.: Peer review of datasets: when, why, and how. *Bull. Am. Meteorol. Soc.* **96**(2), 191–201 (2015). <https://doi.org/10.1175/BAMS-D-13-00083.1>
6. Candela, L., Castelli, D., Manghi, P., Tani, A.: Data journals: a survey. *J. Assoc. Inf. Sci. Technol.* **66**(9), 1747–1762 (2015). <https://doi.org/10.1002/asi.23358>
7. Carpenter, T.A.: What Constitutes Peer Review of Data: A Survey of Published Peer Review Guidelines, April 2017. <http://arxiv.org/abs/1704.02236>
8. Assante, M., Candela, L., Castelli, D., Tani, A.: Are scientific data repositories coping with research data publishing? *Data Sci. J.* **15**, 1–24 (2016). <https://doi.org/10.5334/dsj-2016-006>
9. Callaghan, S., et al.: Guidelines on recommending data repositories as partners in publishing research data. *Int. J. Digital Curation* **9**(1), 152–163 (2014). <https://doi.org/10.2218/ijdc.v9i1.309>
10. Tenopir, C., et al.: Data sharing by scientists: practices and perceptions. *PLoS One* **6**(6), e21101 (2011). <https://doi.org/10.1371/journal.pone.0021101>
11. Kim, Y., Adler, M.: Social scientists’ data sharing behaviours: investigating the roles of individual motivations, institutional pressures, and data repositories. *Int. J. Inf. Manage.* **35**, 408–418 (2015). <https://doi.org/10.1016/j.ijinfomgt.2015.04.007>
12. Faniel, I.M., Kriesberg, A., Yakei, E.: Social scientists’ satisfaction with data reuse. *J. Assoc. Inf. Sci. Technol.* **67**(6), 1404–1416 (2015). <https://doi.org/10.1002/asi.23480>
13. Kratz, J.E., Strasser, C.: Researcher perspectives on publication and peer review of data. *PLoS ONE* **10**(2), e0117619 (2015). <https://doi.org/10.1371/journal.pone.0117619>

14. Stančiauskas, V., Banelytė, V.: OpenUP survey on researchers' current perceptions and practices in peer review, impact measurement and dissemination of research results survey, 19 April 2017. <https://doi.org/10.5281/zenodo.556157>
15. Borgman, C.L.: *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. MIT Press, Cambridge (2007)
16. Research Data Journal. <https://brill.com/view/journals/rdj/rdj-overview.xml>. Accessed 19 Sept 2018
17. DANS - Data Archiving and Network Services. <https://dans.knaw.nl/en>. Accessed 19 Sept 2018
18. UK Data Archive. <http://www.data-archive.ac.uk/>. Accessed 19 Sept 2018
19. GESIS - Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen. <https://www.gesis.org/en/home/>. Accessed 19 Sept 2018
20. CESSDA - Consortium of European Social Science Data Archives. <https://www.cessda.eu/>. Accessed 19 Sept 2018
21. ICPSR - Interuniversity Consortium for Political and Social Research. <https://www.icpsr.umich.edu/icpsrweb/>. Accessed 19 Sept 2018
22. Data Seal of Approval. <https://www.datasealofapproval.org/en/>. Accessed 19 Sept 2018
23. DDI - Data Documentation Initiative. <http://www.dcc.ac.uk/resources/metadata-standards/ddi-data-documentation-initiative>. Accessed 19 Sept 2018
24. Callaghan, S.: Data without peer: examples of data peer review in the earth sciences. *D-Lib Mag.* **21**(1/2), 9 (2015). <https://doi.org/10.1045/january2015-callaghan>
25. Data review in Easy. <http://datareviews.dans.knaw.nl/>. Accessed 19 Sept 2018
26. Barbieri, M., Wilmoth, J.R., Shkolnikov, V.M., et al.: Data resource profile: the human mortality database (HMD). *Int. J. Epidemiol.* **44**(5), 1549–1556 (2015). <https://doi.org/10.1093/ije/dyv105>
27. Special methods used for selected population. <https://www.mortality.org/Public/Docs/SpecialMethods.pdf>. Accessed 19 Sept 2018
28. Max Planck Institute for Demographic Research Technical reports. https://www.demogr.mpg.de/en/projects_publications/publications_1904/mpidr_technical_reports/all.htm. Accessed 19 Sept 2018
29. User Agreement. <https://www.mortality.org/Public/UserAgreement.php>. Accessed 19 Sept 2018
30. Citation guidelines. <https://www.mortality.org/Public/CitationGuidelines.php>. Accessed 19 Sept 2018

Author Query Form

Book ID : **479208_1_En**

Chapter No : **20**

Please ensure you fill out your response to the queries raised below and return this form along with your corrections.

Dear Author,

During the process of typesetting your chapter, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query Refs.	Details Required	Author's Response
AQ1	The closing parenthesis does not have a corresponding opening parenthesis in the sentence 'Although they have no official...'. Please insert the parenthesis in the appropriate position.	