



Stamp—Standardized Data Management Plan for Educational Research: A Blueprint to Improve Data Management across Disciplines

PRACTICE PAPER

SEBASTIAN NETSCHER

ELKE C. BONGARTZ

ANNA K. SCHWICKERATH

DOMINIK BRAUN

KARSTEN STEPHAN

REINER MAUER

*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

To provide more tailored, discipline-specific guidance on data management, Science Europe suggested the concept of domain data protocols. Based on this concept, the project *Domain Data Protocols for Educational Research* developed a first domain data protocol for educational research, titled *Standardized Data Management Plan for Educational Research* (Stamp). Its multi-level approach includes minimal conditions on managing data according to the FAIR Data Principles and checklists with concrete activities to reach each minimal condition; also included are auxiliary materials to support researchers in educational research in planning, implementing, and realizing different data management activities. Although we developed the Stamp for educational research, its design and flexible structure enables transferring it to other (research) domains and communities. To investigate this flexibility, we organized two workshops, discussing to what extent the Stamp can be used beyond educational research, with representatives from other social science domains as well as from research domains beyond the social sciences. In sum, there was consensus among participants of both workshops on the usability of the Stamp outside educational research, at least if the same types of data are processed and analyzed with similar methods. For other types of data, the Stamp serves as a blueprint to develop further domain data protocols, in terms of standardized data management plans, according to the specific needs of the respective domain.

CORRESPONDING AUTHOR:

Sebastian Netscher

GESIS – Leibniz-Institute for
the Social Sciences, DE

Sebastian.Netscher@gesis.org

KEYWORDS:

(research) data management;
data management plans;
domain data protocols;
discipline-specific guidance;
standardization; cross-
disciplinary data management

TO CITE THIS ARTICLE:

Netscher, S, Bongartz, EC,
Schwickerath, AK, Braun, D,
Stephan, K and Mauer, R.
2024. Stamp—Standardized
Data Management Plan
for Educational Research:
A Blueprint to Improve
Data Management across
Disciplines. *Data Science
Journal*, 23: 7, pp. 1–10. DOI:
[https://doi.org/10.5334/dsj-
2024-007](https://doi.org/10.5334/dsj-2024-007)

1. INTRODUCTION

Data management is an integral part of good scientific practice. But the realization of adequate data management remains challenging, especially for researchers. Templates for data management plans, aiming to support researchers in planning and implementing data management, often lack concrete recommendations to support researchers in data management (Smale et al. 2020; Whitmire et al. 2015). Consequently, researchers ask for ‘much more tailored guidance and discipline-specific examples’ (Grootveld et al. 2018: 9) to manage data according to their needs and the common practice of their community. To address this challenge of more tailored and specific guidance, Science Europe established the concept of domain data protocols (DDP), i.e., ‘a “model DMP” for a given domain or community that shares common methods’ (Science Europe 2018: 8).

Based on this concept, members of the project *Domain Data Protocols for Empirical Educational Research* (DDP-Bildung) developed the *Standardized Data Management Plan for Educational Research*, briefly called *Stamp* (Perry et al. 2022; German Network of Educational Research Data 2022). As outlined in the second, the Stamp supports researchers in planning and implementing data management in educational research. It consists of eight content modules, each containing a so-called minimal condition on managing data, granular checklists to realize the minimal conditions and auxiliary materials, providing guidelines and examples on different data management activities. The eight content models reflect the FAIR Data Principles (Wilkinson et al. 2016), the idea of Open Science, and the rules of good scientific practice.

Initially, we developed the Stamp for the educational research community, aiming to provide guidance tailored to this domain. At the same time, we expected that the idea of activities to manage data adequately is rather a matter of the type of data processed, the methods of data analysis, and the information included in the data rather than of a particular community or domain (Netscher, Anders and Henzen 2022). We thus considered the transferability of the Stamp to other domains and its usability beyond educational research during the entire development process. We designed the Stamp with a flexible structure that allows it to evolve and adapt to changing data management practices, such as employing new types of data. This structure not only allows future development of the Stamp within educational research but also its transferability to other research domains.

To investigate the extent of transferability of the Stamp and its usability beyond educational research, staff members of DDP-Bildung organized two workshops, funded by KonsortSWD (2022), in August 2022. As discussed in section three of this paper, the outputs of these workshops highlight that the Stamp’s minimal conditions can be used as a basis for data management and its documentation across domains. Likewise, checklists can be transferred to other domains, at least to some extent, as exemplified in section four. In sum, transferring the Stamp to other domains is a matter of translating its terminology to match the terminology of respective communities.

2. THE STANDARDIZED DATA MANAGEMENT PLAN FOR EDUCATIONAL RESEARCH

Researchers are increasingly encouraged by different stakeholders, such as academic journals, professional associations, and research funding agencies, to make the research process as transparent as possible to enable reproducible results and to share data FAIRly and openly with others. Such requirements can be challenging. First, guidelines of good scientific practice often remain very general and at a high level of abstraction. Second, researchers often find data management a necessary evil. And third, not all of them are familiar with the FAIR Data Principles and the idea of Open Science. At the same time, existing tools to foster the creation of FAIR data—such as templates for data management plans (DMP)—largely vary, and rarely provide tailored, discipline-specific guidance (Grootveld et al. 2018).

To address this lack of more discipline-specific guidance, the project *DDP-Bildung* developed the Stamp. Funded by the German Federal Ministry of Education and Research (grant number 16QK01), the project brought together 12 research institutions in Germany with diverse areas of expertise on educational research to develop the Stamp within a funding period from June 2019 to May 2022. Based on the concept of Science Europe, DDPs are open, standardized, and

referenceable data protocols, serving as ‘model’ DMPs for a particular domain or community (Science Europe 2018). Accordingly, a DDP covers different topics of data management, such as ‘documentation and data quality,’ ‘legal and ethical requirements, codes of conduct,’ or ‘data sharing and long-term preservation’ (Science Europe 2021). For each of these topics, a DDP includes, e.g., ‘applicable regulations,’ ‘applicable standards,’ and ‘templates and examples’ (Science Europe 2018: 9). While the formal minimum conditions should ‘set a minimum standard’ for data management and its quality, standards, templates, and examples assist researchers in realizing their data management (Science Europe 2018: 10).

DDPs are regarded as a ‘pragmatic solution’ for the benefit of various stakeholders. They assist researchers in doing data management, preparing project proposals and funding applications, and offer support for data archiving and sharing. By describing activities to manage data throughout the data life cycle, DDPs also simplify the budgeting of such activities. They support replication of results by the community and the re-use of data by others in new (research) contexts. DDPs simplify review processes regarding data management, reducing the efforts of examining funding applications and reports on data management by implementing standardized procedures. Finally, DDPs foster data ingestion in data repositories and archives by assisting researchers in creating FAIR data. In addition, considering the various recommendations and attempts to reform research assessment and to improve the relevance of managing and sharing research data, we consider the Stamp to be an effective support here (German Research Foundation 2022; European Commission 2021a; Science Europe 2020).

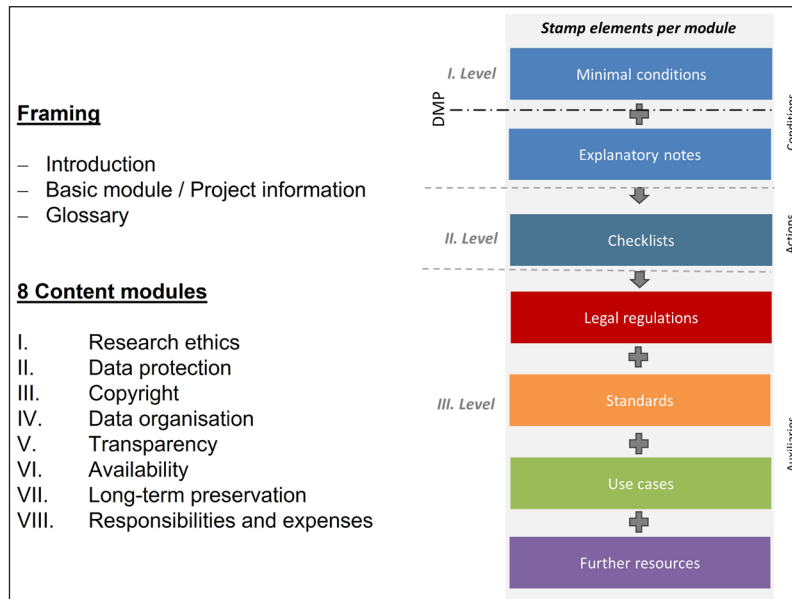
In the DDP project, we investigated educational research as a domain, characterized by common methods of data processing, analysis, use, and interpretation. Types of data in educational research are very heterogeneous, but there is a common terminology to describe such data. We applied this terminology in the development of the Stamp to improve its understandability and acceptance, paying close attention to the needs and requirements of the educational research domain by involving potential users throughout the project. The Stamp’s primary aim is to standardize and facilitate data processing and management in empirical educational research following the FAIR Data Principles. The data can be made available to third parties as openly as possible and as closed as necessary according to the idea of Open Science and in particular the ‘A’ of FAIR (European Commission 2016; Landi et al. 2020).

The Stamp consists of eight so-called content modules and a basic module providing information, e.g., on the project and its data, as well as of an introduction and a glossary, illustrated in Figure 1. The eight content modules comprise research ethics, data protection, copyright, data organization, transparency, availability, long-term preservation, and responsibilities and expenses. Each content module is hierarchically structured, subdivided into different elements. At the first level, the modules include minimal conditions on managing FAIR data and offer explanatory notes. In principle, these minimal conditions form the resulting DMP, and a reference to it should be meaningful enough for funders, according to the basic idea of Science Europe. The second level comprises checklists for managing data and complying with the minimal conditions (level of actions). The third level finally provides legal requirements, e.g., in the context of data protection regulations or copyrights, standards, use cases, and further resources (level of auxiliary materials). On this level, the Stamp contains domain-specific guidance in terms of its standards, referencing to, e.g., guidelines and best practice advise of associations, research centers, and repositories from educational research. Use cases illustrate possible applications, introducing challenges and solutions for the various data management activities in the context of educational research, employing the terminology of educational researchers and reflecting their work routines.

In comparison to the Stamp, most conventional DMP templates, e.g., in DMPOnline (2023), from the National Science Foundation (2023), or Horizon Europe (European Commission 2021b), consist of sets of questions about data management. But, they rarely provide guidance on how to find answers to these questions and thus on how to plan, implement, and realize data management activities. The Stamp sets out to address this need of researchers, giving concrete answers to a vast number of questions on how to manage their data, providing guidance in the form of checklists and auxiliary material. The Stamp thus differs from conventional DMP templates, going beyond simply naming data management activities and respective explanatory notes (see Figure 1, level 1). Researchers can apply the Stamp just by using its checklists. This option might be particularly fitting for experienced researchers interested in

giving their data management a clear structure, easy to compare with other projects and thereby promoting the idea of Open Science. According to their individual level of expertise, researchers can also choose to use more guidance. For example, an early career researcher, in his/her first research project, may want to consult legal regulations, a use case, or even use a training recommended in the further resources of the Stamp.

Figure 1 Overview of the Stamp.



Unlike traditional DMP templates, the Stamp, with its clear structure and step-by-step instructions, provides low-threshold support to assist researchers in helping themselves, and thereby also aims to increase data availability. As recent research shows, research data availability in the field of educational psychology was as low as 7.16% in 2020, independently of the scientific journals' and the research institutions' research data policies (Huff and Bongartz 2023). Thus, in cases where data management is not compulsory, introducing targeted incentives would significantly enhance the availability of research data. This requires that researchers comprehend the benefits of sharing data and have the capability and motivation to do so. The Stamp, rather than conventional DMP templates, is an example of providing such incentives.

3. THE CROSS-DISCIPLINARY CHARACTER OF THE STAMP'S MINIMAL CONDITIONS

Initially, we designed the Stamp to manage data of educational research. However, after consulting various stakeholders, we are convinced that the Stamp can be used outside educational research, expecting that data management is rather a matter of the types of data processed and methods employed than of a particular domain. With the aim of putting this approach to test, we organized two workshops with data management experts from the social sciences, on August 16th, 2022, as well as with experts from other research domains outside the social sciences, on August 23rd, 2023. In these workshops, we discussed the extent to which minimal conditions and checklists can be used beyond educational research. We announced the workshops via mailing lists and blog posts on the website of the project DDP-Bildung and further disseminated by DDP-partners into related domains. In total, more than 100 people applied for participation in one of the workshops. As places for each event were limited to 30 people, we selected participants according to the date of registration. In sum, 28 experts—representing social science domains such as political science, sociology, and psychology—participated in the first workshop. In addition, 26 experts joined the second workshop, representing, e.g., chemical science, agriculture, biology, physics, linguistics, philosophy, or medical science.

Most of these participants were from German universities, and were involved in general data management. Among the participants were particularly knowledgeable persons from the field of data management as well as leading experts. This had the advantage of receiving the statements of researchers who could provide insights into entire domains rather than speaking from a limited personal experience.

(European Organization of Nuclear Research 2019: 27). Such a volume cannot be preserved, or can be only with a considerable expenditure of resources, which can be very challenging, considering energy-saving measures and other issues (Bello and Lamont 2022).

Other minimal conditions are partially or even completely obsolete for some domains or empirical research projects, such as data protection (II.). Projects that do not process personal data do not need to consider data protection regulations. This may apply to entire domains, but even in medical or social science research empirical projects exist that do not process personal data, e.g., when examining organizations or associations, or when working with aggregated data, exclusively. In short, compliance with data protection regulations depends on processing personal data and the information covered in this data (Article 1 of the EU General Data Protection Regulation), regardless of the domain concerned. The flexible structure of the Stamp enables users to skip data management activities or even complete modules, if they are irrelevant for a given project. However, as summarized by one of the participants from natural sciences, once personal data is processed, such as when organizing an event, the Stamp’s minimal condition and respective checklists serve as a guideline to follow data protection regulations.

Finally, some of the Stamp’s minimal conditions do not meet the requirements of some domains completely. As an example, the minimal condition on copyright (III.) does not include patent rights, which might be relevant, e.g., in engineering or computer science. In addition, other legal regulations such as animal protection in veterinary medicine or the Nagoya Protocol on Biodiversity (2011) in biology are not considered, as highlighted by participants from the life sciences. Here, the respective domains are called upon to discuss possible additions to the Stamp’s minimal conditions required for their data.

4. APPLYING THE STAMP’S CHECKLISTS OUTSIDE EDUCATIONAL RESEARCH

In addition to the minimal conditions, we also explored the cross-disciplinary character of the Stamp’s checklists, discussing two checklists in the workshops. The first checklist focuses on consent management in the context of data protection, displayed in Table 2. For researchers in social sciences, data protection, informed consent and secure data handling are part of their everyday work routine. Regarding the checklist, there was consensus among participants of both workshops that consent management is a legal requirement for processing personal data, regardless of the domain.

NOT APPLICABLE	CONSIDERED	
	<input type="radio"/>	II.Ba.5 project members implement a consent management system , considering
	<input type="radio"/>	II.Ba.5.1 the purposes of processing consented
	<input type="radio"/>	II.Ba.5.2 further (categories) of recipients of data
	<input type="radio"/>	II.Ba.5.3 the duration of data storage as well as the criteria for data destruction
	<input type="radio"/>	II.Ba.5.4 the claiming of rights of the persons concerned, documenting
	<input type="radio"/>	II.Ba.5.4.1 who has claimed which rights, when, and how
	<input type="radio"/>	II.Ba.5.4.2 the timely communication with persons claimed their rights
	<input type="radio"/>	II.Ba.5.4.3 consequences of right claiming for further data processing see use case on <i>Objection Data Processing and Revocation of Consent (II.Ba.5.4.3.F)</i>
	<input type="radio"/>	II.Ba.5.4.4 restriction of rights of data subjects and its conditions see use case on <i>Restricting Participant’s Right (II.Ba.5.4.4.F)</i>
	<input type="radio"/>	II.Ba.5.5 possible deadlines , e.g., for destructing contact data, according to the consent form

Table 2 Checklist II.Ba.5 consent management (II. data protection).

Participants agreed that the checklist contains information needed to meet legal requirements of data protection regulations. Interestingly, while the checklist on consent management can be used in many domains, it requires some adaptations; for example, when collecting test results in the form of blood samples in medical science. Also, it is of importance that such legal requirements depend on the region in which researchers operate. It implies that all minimal conditions and checklists of the Stamp containing legal information must be adapted to regions outside Germany and the European Union, e.g., for international cooperative projects, as discussed with representatives from the social sciences in the first workshop.

In addition, participants of both workshops examined the usability of one of the checklists on data documentation ensuring transparency. The discussions focused on the documentation of measurement instruments for quantitative (V.Al.11) and qualitative (V.As.6) data, as shown in Table 3, and to which extent activities listed can be transferred to other domains. In general, participants agreed on the checklist's content as being of relevance for all domains. To ensure understandability, interpretability and (in part) replicability of data, data documentation is key, e.g., regarding data provenance and the measurement instrument employed.

NOT APPLICABLE		CONSIDERED	
V.Al.11 Documentation of the Measuring Instrument			
<input type="radio"/>	<input type="radio"/>	V.Al.11	The documentation contains information on the measurement instrument at the study level
<input type="radio"/>	<input type="radio"/>	V.Al.11.1	(documentation of the) original (digitalized) measurement instrument, including all questions, items, tasks, and answer options (as well as options for refusing to answer)
<input type="radio"/>	<input type="radio"/>	V.Al.11.2	(documentation of) digitalized interviewer instructions
<input type="radio"/>	<input type="radio"/>	V.Al.11.3	(documentation of) digitalized skip pattern structure in the measurement instrument
V.As.6 Documentation of the Measuring Instrument and the Method of Data Analysis			
	<input type="radio"/>	V.As.6	The documentation contains information on the survey instrument and the evaluation method on the study level
<input type="radio"/>	<input type="radio"/>	V.As.6.1	(documentation of the) original (digitalized) measurement instrument
<input type="radio"/>	<input type="radio"/>	V.As.6.2	(documentation of) digitalized documentation forms, socio-demographic questionnaire, interview guide, etc.
<input type="radio"/>	<input type="radio"/>	V.As.6.3	description of document collection or protocolling of interviews see recommendations on <i>Observation Protocols (V.As.6.3.E)</i>
	<input type="radio"/>	V.As.6.4	description of specifications regarding data storage, transport and preservation of recording and observations
	<input type="radio"/>	V.As.6.5	description of data collection and measures to ensure data quality
<input type="radio"/>	<input type="radio"/>	V.As.6.6	citation (source, persistent identifier, and mode of citation, if applicable) and licence conditions of all materials (re-)used in the project, such as
<input type="radio"/>	<input type="radio"/>	V.As.6.6.1	data, artefacts, and documents
<input type="radio"/>	<input type="radio"/>	V.As.6.6.2	further measurement instruments , like interview guides
<input type="radio"/>	<input type="radio"/>	V.As.6.6.3	supporting materials for interviewing, stimuli etc.
	<input type="radio"/>	V.As.6.7	(documentation of) method of data analysis

Table 3 Checklist V.Al.11 Documentation of the measuring instrument and Checklist V.As.6 Documentation of the measuring instrument and the method of data analysis (V. Transparency).

Of course, different domains employ different documentation standards, metadata sets and (controlled) vocabularies to describe, e.g., the process of data gathering. Such standards and vocabularies are usually to a great extent domain-specific. However, as discussed in the workshops, the domain-specific character of data documentation is rather about data

- Netscher, S, Anders, I and Henzen, Ch.** 2022. Activities and challenges in developing discipline-specific data management plan templates: From vertical to horizontal integration of RDM practices. *Bausteine Forschungsdatenmanagement*, 1(März): 13–25. DOI: <https://doi.org/10.17192/bfdm.2022.1.8371>
- Perry, A,** et al. 2022. Domain-Data-Protokolle für die empirische Bildungsforschung (DDP-Bildung): Schlussbericht des Verbundprojekts: Projektlaufzeit: 01.06.2019 bis 31.05.2022. DOI: <https://doi.org/10.2314/KXP:1859079814>
- Regulation (EU).** 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, GDPR). Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679> [Last accessed 13 December 2022].
- Science Europe.** 2018. Science Europe guidance document presenting a framework for discipline-specific research data management. Version v1. 18 January 2018. DOI: <https://doi.org/10.5281/zenodo.4925906> [Last accessed 16 February 2024].
- Science Europe.** 2020. Position statement and recommendations on research assessment processes. Available at www.scienceeurope.org/media/3twjxim0/se-position-statement-research-assessment-processes.pdf [Last accessed 13 December 2022].
- Science Europe.** 2021. Practical guide to the international alignment of research data management – extended edition. Version v1. 27 January 2021. DOI: <https://doi.org/10.5281/zenodo.4915862> [Last accessed 16 February 2024].
- Secretariat of the Convention on Biological Diversity.** 2011. Nagoya protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity. Available at <https://www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf> [Last accessed 13 December 2022].
- Smale, N,** et al. 2020. A review of the history, advocacy and efficacy of data management plans. *International Journal of Digital Curation*, 15(1). DOI: <https://doi.org/10.2218/ijdc.v15i1.525>
- Whitmire, AL, Boock, M and Sutton, SC.** 2015. Variability in academic research data management practices: Implications for data services development from a faculty survey. *Electronic Library and Information Systems*, 49(4): 382–407. DOI: <https://doi.org/10.1108/PROG-02-2015-0017>
- Wilkinson, MD,** et al. 2016. The FAIR guiding principles for management and stewardship. *Scientific Data* 3. DOI: <https://doi.org/10.1038/sdata.2016.18>

TO CITE THIS ARTICLE:
Netscher, S, Bongartz, EC, Schwickerath, AK, Braun, D, Stephan, K and Mauer, R. 2024. Stamp—Standardized Data Management Plan for Educational Research: A Blueprint to Improve Data Management across Disciplines. *Data Science Journal*, 23: 7, pp. 1–10. DOI: <https://doi.org/10.5334/dsj-2024-007>

Submitted: 15 December 2022
Accepted: 06 February 2024
Published: 23 February 2024

COPYRIGHT:
© 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.