ZEF Data Management Policy

10/2014

Antonio Rogmann, John P.A. Lamers, Tina Beuchelt, Saravanan Subramanian, Dominik Wisser, Franz Gatzweiler, Marc Müller

Center for Development Research, University of Bonn

Content

Abbreviations and acronyms 3
Definition of technical terms 3
Preface 5
1 Introduction 6
1.1 Goals of ZEF Data Management 6
1.2 Principles in the ZEF Data Management 6
1.3 Scope 6
2 Data Rights Management 7
2.1 Data shared by the ZEF Data Information System 7
2.2 Data Exchange Regulation 7
2.2.1 Legal Framework 7
2.2.1.1 Intellectual Property 7
2.2.1.2 Data Privacy 8
2.2.2 Re-use of Data 8
2.2.3 License 9
2.2.4 Data Users 9
2.2.5 Data Providers 9
2.2.6 Data Privacy Protection 12
2.3 Citation and Acknowledgments 12
3 Data Management Framework 13
3.1 Basics 13
3.2 Responsibilities 13

Abbreviations and acronyms

- URL (Uniform Resource Locator): Unique address for a resource in the internet
- ZEF-DM (ZEF Data Management System): The entire set of technical, organizational and legal components of the ZEF Data Management
- ZEF-DMP (ZEF Data Management Policy): Legal framework of the ZEF-DM. Regulates rights and commitments of every actor in the ZEF-DM. They need to be accepted by all actors participating in the ZEF-DM.
- ZEF-DIS (ZEF Data Information System): Technical component of the ZEF-DMS, mainly consisting of the ZEF Data Portal and the ZEF Research Databases, implemented in a web-service-based client-server architecture hosted by ZEF. The ZEF-DIS is managed by the ZEF Data Management Unit (ZEF-DMU)
- ZEF-DMU (ZEF Data Management Unit): The Group in charge of implementing and maintaining the ZEF-DM
- ZEF-DP (ZEF Data Portal): Web-portal application allowing users to access the ZEF Research Databases (ZEF-RDB) by a searchable metadata catalogue
- ZEF-RDB (ZEF Research Databases): Databases at ZEF where data, metadata and other information is stored

Definition of technical terms

For the purposes of this document the following definitions have been adopted:

- Data: „Data“ is defined as *„a reinterpretable representation of information in a formalized manner, suitable for communication, interpretation, or processing“*[1] represented by distinct units of information such as facts, numbers, letters, symbols, usually formatted in a specific way, stored in a data base and suitable for processing by a computer
- Metadata: Data providing information on data available and availability, enabling the user to assess its fitness for its own use and purposes
- Primary Data: Data collected/ created by the data user
- Secondary Data: Data collected and provided by someone other than the user
- Raw Data: Data coming from data loggers or other field instruments or surveys without any further reviewing in terms of data quality or any other processing
- Data User: An individual authorized to access and use data
- Data Provider: An individual or an institution which distributes data to other parties
- Data Manager: The person(s) responsible for the operational management and processing of the data in the information system
- Data Standard: Norm or requirement for an item (e.g. service, process, technique) based on a common agreement
- Data Management: The development and execution of architectures, policies, practices and procedures that properly manage the full data lifecycle needs of an institutional unit (following data management definition of the international Data Management Association, i.e. http://en.wikipedia.org/wiki/Data_management)
- Data Rights Management: The regulation of access to data under the consideration of intellectual property rights and other legal restrictions

[1] ISO/IEC 2382-1. Source: http://de.wikipedia.org/wiki/Daten. Accessed: 12/2012

- Data Management Plan: Formal document created at the beginning of a research project describing planned
    - data management workflows (quality assurance, formatting, coding, metadata creation, etc.) in the life cycle of data from data acquisition, processing, sharing, archiving to deletion;
    - standards in use;
    - technical facilities;
    - organizational data management framework (responsibilities in data management workflows);
    - data sharing and exchange rules (data exchange policy, licenses).

Authors: Antonio Rogmann*, John Lamers*, Tina Beuchelt*, Saravanan Subramanian***, Dominik Wisser*, Franz Gatzweiler**, Marc Müller**

Acknowledgement: We gratefully acknowledge the assistance in legal and work-contract related issues from Ms. Franziska Hadaschik, Contract Management Respresentative (Dpt. 7.2 – Knowledge Transfer and Trade Mark Rights), and Mr. Thomas Rütten, In-house counsel (Dpt. 1 - Law, Academic and Student Affairs), both from the University of Bonn.

For the linguistic revision we like to thank very much Christian Borgemeister* and Tyler Frazier*.

* ZEF, Dpt. C

** ZEF, Dpt. B

*** ZEF, Dpt. A

Preface

The Center for Development Research (ZEF) carries out inter and trans-disciplinary research aiming at gaining findings on social, economic and ecological systems and their interactions also in "developing countries". Increased understanding is compulsory for sustainable development and helps solving challenges to reach better living conditions.

Commonly, these findings are transmitted to a broader audience such as the science community or the broad public by scientific papers as interpretations and conclusions of research data, which in most instances is not accessible for verification or re-use purposes. But there is a strong need to make research data underpinning any publication available, not only to the science community as a whole but also to the public, as emphasized by the Alliance of German Science Organizations (2010, p. 2):

„*Quality-assured research data are a cornerstone of scientific knowledge and, independent of the purpose for which they were originally obtained, can often serve as the basis for further research. (….). Preserving research data over the long term and making them available therefore does not only serve the verification of prior results, but also, to a large extent, the obtaining of future ones. It is a strategic task to which science and the humanities, politics as well as other parts of society, must contribute*".

In the European as well as German scientific landscape awareness raising for the need to share research data within and outside the scientific community is rapidly emerging. Scientific leading organizations, such as the German Council of Science and Humanities (Wissenschaftsrat, 2012) or the High Level Expert Group on Scientific Data (2010), foster this process by outlining the significant benefits gained from the interaction between society and science, the interdisciplinary collaboration in research and the benefits for individual scientific careers e.g. by reputation as high quality data provider.

In the national and international ongoing debates on the improvement of research data management many challenges have been addressed. Funding schemes, necessary changes in institutional data management cultures, workflows and individual motivation, data and system interoperability, citation and data documentation standards, open access and the management of intellectual property crossing jurisdictional boundaries, are part of these continuous debates.

However, with the implementation of an own Data Information System, ZEF embraces the challenge, and contributes to the global trend of growing national and international scientific data infrastructure developments. Furthermore, ZEF considers itself with this view in the particular commitment to share research data and its products with communities worldwide, but in particular in regions where data availability and accessibility is even more problematic than in other parts of the world.

The present Data Management Policy (ZEF-DMP) provides the organizational as well as legislative framework for ZEF's Data Management. Actually, the ZEF-DMP seems to be unique in the consideration of legal conditions for data sharing at the University of Bonn, so that ZEF can claim to be a trendsetter in providing trust and confidence in the management of researchers' data. It is further hoped that this example finds followers within Bonn University and beyond.

The ZEF Directorate

C. Borgemeister A.-K. Hornidge J. v. Braun

1 Introduction

1.1 Goals of ZEF Data Management

The goal of the ZEF Data Management System (ZEF-DM) is to provide permanent archival services for high quality research related to data and to make these data sets available over the web for the following purposes.

Within ZEF, to

- save time and costs by streamlining data search, access, and distribution workflows within and between ZEF projects as well as in project partner networks;
- avoid duplication in data collection and knowledge generation;
- support research project development by providing comprehensive overviews of research that has already been completed;

outside of ZEF, to

- support the greater scientific community in the area of development research by making available well documented research data that have been generated by projects at ZEF; and
- better inform stakeholders and decision makers in support of efforts to improve the living conditions in developing countries through user-friendly access to research data and data products such as surveys, statistics, research reports or maps.

1.2 Principles in the ZEF Data Management

To achieve ZEF-DM goals, clearly defined organizational and legal principles are set forth and outlined in this policy as follows:

- Data must be of high quality, in accordance to the best practices of the respective scientific disciplines. Quality management and documentation is the responsibility of the scientist who is providing the data;
- Data must be well documented, such that individual users can assess the suitability of different data sets for their purposes;
- Data must be recorded in a manner that can be easily searched and accessed while being protected against its misuse;
- Data rights that serve to protect intellectual property and the confidentiality of sensitive personal information as well as other requirements for restricting access to data[2] must be considered consistently, both during data preparation and distribution;
- Data authors must be cited when data is re-used and contributors to the generation or processing of the data must be acknowledged; and
- The use of data from the ZEF Data Information System (ZEF-DIS) without proper citation is strictly prohibited.

1.3 Scope

The ZEF Data Management Policy (ZEF-DMP) defines usage and accessibility-rights with regards to research data that have been stored in the ZEF Research Databases (ZEF-RDB). This policy governs data that are managed by the ZEF-DIS, the integration of data into the ZEF-DIS and the delegation of necessary tasks resulting from data management workflows.

At ZEF, every scientist is required to manage his or her research data following a written policy. In some projects the management of data is predetermined by the funding agency, the organization and/or management of a project or other factors. Appropriate guidelines need to be documented in a project-specific

[2] Publication in progress, etc.

data management policy, especially since project-specific data management policies may have a priority higher than as the here outlined ZEF-DMP. It is suggested that such project-specific data management policies should refer to the ZEF-DMP in whole or to specific parts, as appropriate.

Projects, which have their own policies and data sharing regulations may still use the ZEF-DIS. In this case, close collaboration with the ZEF Data Management Unit (ZEF-DMU) will be needed to organize data integration and to manage data access rights in the ZEF-RDB by means of the ZEF-DIS[3].

Any other project or individual scientist participating in the ZEF-DM has to follow the instructions laid out in this policy.

Whether or not a project or individual scientist becomes part of the ZEF-DM has to be decided by the respective project managers and the ZEF directorate.

2 Data Rights Management

Data rights management concerns the legal framework used in the distribution and re-use of research data and includes:

- Protection of intellectual property and other data related rights;
- Protection of sensitive information; and
- Acknowledgement of intellectual and practical contributions during the generation of the data.

In the following the scope of intellectual property rights and other rights, is referred to as „data rights".

These data rights provide general directives on how to manage the sharing, re-use and citation of all research data managed by the ZEF-DIS such that the German law is not violated.

2.1 Data shared by the ZEF Data Information System

In general, the ZEF-DM applies to all data created, collected and used by projects carried out or coordinated by ZEF and participating in the ZEF DM, independently from its processing status or source.

2.2 Data Exchange Regulation

2.2.1 Legal Framework

Given that the University of Bonn of which ZEF is part of, is a German institution, German law is being applied in all cases including thus the terms of intellectual property, data privacy protection, etc.

2.2.1.1 Intellectual Property

In the Federal Republic of Germany, intellectual property is protected by the Copyright Act. An important aspect of this law is the absence of a copyright protection for bare facts in the form of data. The Copyright Act can only protect the form in which these facts/data are presented. In particular, collections of data, which by virtue of the selection or arrangement of data represent a personal intellectual creation, are protected by the Copyright Act[4]. With regard to research findings, the German Copyright Act[5] ensures protection for:

- Technical and other literature, such as written works, speeches, and software code, illustrations of scientific or technical nature, such as drawings, plans, maps, sketches, tables and three-dimensional representations;

[3] Details on how to use the ZEF Data Information System and on rules for data management which are detailed in the ZEF Data Management Handbook.
[4] § 4 Section 1 UrhG (Copyright Law).
[5] § 2 Section 2 UrhG.

insofar as the following general conditions are met:[6]

- Personal creation (in the sense of human creative activity, indicating thus that the output must be the result of creativity of a human-being and thus excluding the findings of models or the like);
- Intellectual content;
- Perceptible design;
- Individuality (work as a result of individual intellectual work).

As a consequence, not every type of data is regarded as the intellectual property of the scientist who has collected and/or produced the data (De Cock Buning, M. et al. 2011, p. 12) [7].

In Germany, copyright is non-transferable. But the rights holder can grant to others his/ her rights to use the data for instance in the form of a license. The license can contain various regulations[8].

2.2.1.2 Data Privacy

Personal data implies any information concerning personal or material circumstances of an identified or identifiable natural person. Data associated with an individual (e.g. household data with private/ personal information that can be related back to the individual) can be used and processed in scientific investigations. However, such data must be anonymized as soon as possible without compromising the purpose of the investigation. If this is not possible, the data should be pseudo-anonymized, i.e. by splitting the personal data records in parts and storing them separately, so that an identification of natural persons is impeded. With an assignment function the data can be rejoined if required. This function may not be known to the people involved in the data processing, but only to the persons authorized to custody the sensitive data.

Personal data may be used in publications only if the persons concerned have declared in writing their agreement or if the public interest in the presentation of research results substantially outweighs the legitimate interests of the person(s) affected[9].

2.2.2 Re-use of Data

Only the author of copyright-protected works has the right to utilize them, i.e. to reproduce, distribute or make them accessible. Within the limits of the Copyright Act, copyright-protected data may be reproduced by anyone for private purposes, for individual academic research or for educational purposes[10].

Further use of data by third parties, e.g. processing new data in a professional research framework and providing data for others, requires prior consent from the rights holder. This consent is granted usually in the form of a license.

The right to "re-use" data includes the right to reproduce, distribute and make accessible (e.g. via the internet) the data, in connection with research activities or decision-making processes. “Re-use” is only permitted within the legal limits and/ or the limits of the license.

Possible restrictions on the data can limit the use rights, for example by prohibiting further distribution.

The right to “re-use” does not free the user from the obligation(s) to indicate sources and/ or their obligations towards the author (section 2.3) and, where applicable, the funding agency.

---

[6] § 2 Section 2 UrhG.

[7] For example, when the method used for collecting data is a commonly-used procedure and was not substantially modified by a scientist, the data are not subject to the Copyright Law and are therefore not protected by copyright. This is, for example, the case for most bio-physical raw data generated through measuring devices. However, as soon as these data are processed or interpreted within a specific scientific context, e.g. by creating special diagrams from condensed measurement data, intellectual property comes into place, and the results are thus copyright protected. However, data not subject to copyright protection can be subject to ownership or user rights, for example by the owner of the measuring device.

[8] For example the use of data can be limited to content, time and place.

[9] § 28 Section 4 DSG NRW.

[10] § 53 UrhG.

2.2.3 License

As part of the Copyright Law, the rights holder can grant licenses for the use of his/ her data[11]. A license is a usage agreement stating the scope of the granted use rights (see section 2.2.2). The license must be concluded by a signature of both parties.

In the ZEF-DP, where data under copyright can be downloaded, a license[12] will be included in the metadata as well as built into the data retrieval process, so that during downloading the data and by entering his contact details the user accepts the license conditions.

2.2.4 Data Users

A data user within the scope of the ZEF-DM is defined as anyone who uses data, which is produced or made available by scientists working on projects carried out, or coordinated by ZEF. The data user can re-use the data in accordance with the limitations stated in section 2.2.2. The data user's obligations with regard to the indicated sources are explained in section 2.3. A user, who wishes to download data from the ZEF-DP, must consent to the restrictions before being able to download the data. The right to re-use data is granted to the user as part of a license as defined in section 2.2.3.

Users who wish to access restricted data must have a user account in the ZEF-DIS. User accounts are regulated by the administrators of the ZEF-DIS in collaboration with the holders of the rights to the requested data. Information on how to procure permission to download restricted data is included in the metadata of the requested data.

2.2.5 Data Providers

The legal commitments with regard to the provision of data to the ZEF-DIS depend on the type of employment of the author of the data. Due to the specifics of the German labor legislation, on the one hand PhD, BSc and MSc students with a so-called scholarship student contract at ZEF (hereinafter referred to as "Scholarship Students"), and on the other hand members of the scientific staff at ZEF and project-related partner institutions (hereinafter referred to as "Senior Researchers"), need therefore to be distinguished.

Professors and scientific staff members, who perform their tasks in research and teaching independently, can grant a license for their data to ZEF in form of a data sharing agreement.

2.2.5.1 Scholarship Students

For the provision of data, the following guidelines and options are applicable to Scholarship Students at ZEF, unless otherwise defined by project specific regulations (section 1.3):

a) Scholarship Students hold all copyrights to the data they generate providing the requirements of the Copyright Act are met (section 2.2.1.1); accordingly, a Scholarship Student is free in the decision to share his/ her data with third parties (e.g. project members, external partners) as it deems necessary to him/ her;

b) Scholarship Students can grant a data use license to ZEF in the form of a data sharing agreement between him/ her and ZEF. If the data use rights were transferred by the copyright holding Scholarship Student to ZEF through a license, ZEF can make this data available to the scientific and non-scientific public to re-use as it deems necessary, as long as this is done so in accordance with the Copyright Law[13];

c) Raw data generated and used by Scholarship Students within the context of their study carried out at ZEF, must be handed over to the scientific supervisor[14] in accordance with the guidelines on good academic practice (German Research Foundation, 2013), not later than four (4) months after the

[11] § 31 UrhG.
[12] Based on the CreativeCommons license framework
[13] § 15 UrhG
[14] In German: Doktorvater/-mutter

completion of the study[15] for instance for a verification of results. The ZEF-leaving form, freeing the student of all ZEF-related obligations, required for the issuing of the doctoral degree, is only signed by the supervisors, once this has been taken care of by the student;

d) A study ends when the final reporting according to the contract is due. The data must be accompanied by a separate, clear description of the data (metadata), to enable the supervisor to identify the dataset, understand its content as well as the methods used during the processing with minimum effort. Data properties, to be described by the author while using the obligatory format of the metadata and the handover, are defined by the scientific supervisor in collaboration with the project coordinator(s). The data are stored for at least ten (10) years. This first use of the data is limited thus to the verification of study results, as to review and assess the final work by the supervisor and other authorized persons. A further use of the data by ZEF requires consent and/ or a license from the Scholarship Student (section 2.2.5.1b);

e) The Scholarship Student must be advised to include the necessary metadata in the ZEF-DIS metadata catalogue from the start of the study, to enable publication of the metadata after completing the study, e.g. for purposes of scientific reputation (see point f) in section 2.2.5.1);

f) Scholarship Students are invited and encouraged to archive their data in the ZEF-RDB and to publish the data within the ZEF-DIS at any time during their study. Scholarship Students have the right to determine access and use rights to this data by granting a license to third parties at any time. The following approaches are possible within the context of the ZEF-DIS:

   i. Publishing contact details in the metadata of a dataset, so that potential users may contact the author directly and request access or use rights from him/ her;

   ii. Providing the ZEF data managers with information about users or groups of users (e.g. colleagues in the respective ZEF project), who are permitted to download data without prior individual permissions; the data manager in the ZEF-DMU will thus create accounts in the ZEF-DIS accordingly;

   iii. Furnishing the data with an appropriate license, which permits anyone to access the data and to use it according to the conditions set by the license provider;

   iv. Changing the status of access rights over time, e.g. according to publication activities;

   v. Providing ZEF with a general license for his/ her data (or parts of it). In this case, the options i-iv of this section 2.2.5.1. do not apply.

g) Scholarship Students, who provide the ZEF-DIS with data, hold a personal password-protected account in the system, which allow them to upload data and metadata onto the ZEF-DIS via the ZEF-DP and to revise or change the data and metadata at a later point in time. The access to data and the visibility of the metadata in the ZEF-DP remain completely under the control of this personal account, but must be in accordance with the regulations of the ZEF-DMP. This account remains active as long as the Scholarship Student continues to be working at ZEF. Further control of the account beyond the contract time at ZEF can be granted by the ZEF Directors upon request;

h) If Scholarship Students want to store secondary data from external sources in the ZEF-RDB and to share them by means of the ZEF-DIS, the data rights specified by the external source must be respected and be part of the metadata. With “external source” it is meant either a scientist at ZEF working in a project not taking part in the ZEF-DM, or any other person, project or institution not related to ZEF at that point of time[16]. It is the responsibility of the Scholarship Student to equip in this case the secondary data sets with the corresponding access restrictions in the ZEF-DP.

2.2.5.2 Senior Researchers

[15] See German Research Foundation: Leaflet on Research Fellowships. DFG Form 1.04-9/13, p. 13 (http://www.dfg.de/formulare/1_04/1_04_de.pdf, accessed Sept. 25th, 2014) and University of Bonn: Richtlinien zur Sicherung guter wissenschaftlicher Praxis an der Rheinischen Friedrich-Wilhelms-Universität Bonn, Bonn 01.09.2014 (http://www3.uni-bonn.de/forschung/gute-wissenschaftliche-praxis/amtl.-bek.-1426.pdf, accessed Oct. 10th ,2014)

[16] Restrictions on data gathered in the context of completed ZEF projects have to be respected as well

The following rules and options apply to the integration of data generated by Senior Researchers in the ZEF-DIS, unless otherwise defined by project-specific (section 1.3) or work-contract related specifications. The Senior Researcher provides his data to ZEF for the purposes outlined in the present ZEF-DMP. It contains the rights and commitments regarding the management of his/ her research data as outlined in the following:

a) Senior Researchers hold all copyrights to the data they produce, providing the requirements of the Copyright Act are met (section 2.2.1.1). Accordingly, the Senior Researchers are free in their decision to share his/ her data with third parties as they deem necessary[17];

b) The right to re-use copyright protected data (as defined in section 2.2.2) is granted to ZEF[18] or to any third party by the Senior Researcher in the form of a license; ZEF can make this data available to the scientific and non-scientific public to re-use as it deems necessary, but in accordance with the Copyright Law[19];

c) Data collected by Senior Researchers, which is, however, not subject to copyright must in any case be integrated into the ZEF-DIS

   - within three (3) months after their generation together with adequate metadata records, as specified by the ZEF-DMU. If data require more time to be prepared for exchange, the upload can be postponed in consultation with the respective project managers and/or the ZEF Directorate. In case of such a delay, the metadata record must contain information on the time the data will become part of the ZEF-DP and hence available;
   - and be subject to the data standards as defined by the ZEF-DMU in collaboration with Senior Researchers (e.g. in a project-specific data management plan or in the ZEF Data Management Handbook);

d) Data which is the intellectual property of Senior Researchers and therefore subject to copyright, must be integrated and published in the ZEF-DIS as follows:

   - Data, which will not be part of a publication (yet), must be integrated in the ZEF-DIS within three (3) months after their generation. It must be complemented with adequate metadata records in the ZEF-DP, as specified by the ZEF-DMU. If the data require more time to be prepared for exchange, the upload of the data can be postponed in consultation with the respective project managers and/or the ZEF Directorate. In case of a delay, the metadata record must contain information on the time the data will become part of the ZEF-DIS and hence available;
   - In the event that data is previewed to become part of a scientific publication: A metadata record must be created in the ZEF-DP within three (3) months after the generation of the data. Senior Researchers have the first right to publish his/ her datasets within twenty-four (24) months after uploading the metadata in the ZEF-DIS by using the ZEF-DP. Hence the data does not need to be uploaded to the ZEF-DIS until the publication is done, respectively the access to the data in the ZEF-DIS remains restricted for any other user than the data providing Senior Researcher. After 24 months the data must be uploaded into the ZEF-DIS and the use right must be transferred to ZEF by a license;
   - If data is desired by third parties before being included in a scientific publication[20] or included in the data base, the data must be integrated into the ZEF-DIS within three (3) months after its creation. If required, the deadline can be adjusted, as agreed between the project coordinators and the Senior Researcher;
   - Data must be prepared for exchange by following the standards as defined by the ZEF-DMU or in collaboration with Senior Researchers e.g. in a project specific data management plan (or as laid out in the ZEF Data Management Handbook section 5.1);
   - The intellectual property of the data must be documented in the metadata;

[17] A data sharing agreement form is available at the ZEF-DMU

[18] The decision for which data a license has to be granted to ZEF lies with the ZEF Directors, or the ZEF panels authorised by them (e.g. the project management).

[19] § 15 UrhG

[20] Because, for example, project partners urgently need the data for re-processing as part of project objectives.

e) Senior Researchers, who make data available to the ZEF-DIS, hold a personal password-protected account in the system. This enables them to upload, revise or change data and metadata into the ZEF-RDB via the ZEF-DP. The access to data and the visibility of metadata on the ZEF-DP remain completely under the control of this personal account, but must be in accordance with the regulations of the ZEF-DMP; the account is activated as long as the Senior Researcher is working at ZEF. Upon request, further control of the account beyond the contract time at ZEF can be granted by the ZEF Directorate;

f) If Senior Researchers want to store secondary data from external sources in the ZEF-RDB and make these accessible within the ZEF-DIS, the corresponding copyrights must be respected. With "external source" it is meant either a scientist at ZEF working in a project not taking part in the ZEF-DIS, or any other person, project or institution not related to ZEF at that point of time.[21] It is in the responsibility of the Senior Researcher to complement secondary data sets with the corresponding access restrictions in the ZEF-DP.

2.2.6 Data Privacy Protection

Every ZEF employee, who collects or processes information of a sensitive or confidential nature, is obliged to comply with data protection regulations (section 2.2.1.2). The author of the data that have privacy or disclosure restrictions and confidentiality requirements must protect them against loss, alteration or any other kind of misuse. Sensitive personal data stored in the ZEF-RDB have to meet the legal requirements.

Methods for anonymizing sensitive data may be developed and applied in cooperation with the responsible ZEF Scientist(s) and the ZEF-DMU. Documents describing the methods used to implement data privacy protection must be stored in a secure place.

Additional information on data privacy and protection during field research can be found in the *ZEF Ethics* document.

2.3 Citation and Acknowledgments

a) When data stored in the ZEF-RDB and obtained via the ZEF-DIS is used in a publication, the source of the data must be cited by using the URL address for the metadata found in the ZEF-DP. The link can be included as a footnote, in the bibliography or in an appendix with similar references (see also ZEF Data Management Handbook, section 6);

b) If data obtained from external sources is used, these initial sources must be cited properly which can be achieved in various ways such as:

   - By citing the people who produced the data, by acknowledging them or by offering co-authorship;
   - A citation is sufficient in case the data is available in the ZEF-DIS and section 2.3 article a) is followed;
   - If data is provided by an institution (generally the case with secondary data), the institution must be cited and, if available, the URL from which the data have been downloaded, must be indicated;

c) If two or more people created a single data set, joint publishing (i.e. the naming of all data authors as authors/co-authors[22] in a publication) is to be considered and recommended by ZEF;

d) When gathering or pre-processing of primary data was substantially supported by others (e.g. student assistants, field technicians, etc.), and they are not co-authors of the publication, they must be acknowledged in the metadata of the primary data, which should be referenced following section 2.3 article a);

[21] Restrictions on data gathered in the context of completed ZEF projects have to be respected as well

[22] As a rule, in joint publications, the person that has written the text of the publication (or over 50% of it) has the right to be the first author. If the contribution of two authors is equal or disputed, the project coordinators decide; all others should be listed as co-authors in case they have made a substantial contribution to the work being at the level of study preparation, execution, processing or interpretation

e) Studies conducted within the framework of a project need to acknowledge the funding institutions by the citation of the following sentence (or a similar one) under „Acknowledgements:"

„This study was funded by [names of the institutions] [project number]".

3 Data Management Framework

In data management workflows specific tasks need specific competences, either from the scientists, from the data managers or both. For this reason the tasks, as listed in section 3.1, need to be allocated between the ZEF-DMU and the projects participating in the ZEF-DM.

In this policy ZEF has outlined a general scheme. Details, however, can be negotiated with each project to meet its requirements and resource availability. ZEF recommends to fix data management workflows together with project management staff in a project specific data management plan. Data management plans can be prepared in cooperation with the ZEF-DMU.

Every project is committed to delegate one project member to the ZEF data management working group which meets regularly to discuss and coordinate data management tasks as listed in section 3.1. The committee is convened and lead by the ZEF-DMU and regularly reports to ZEF's Directorate.

3.1 Basics

The regulatory and organizational framework to manage ZEF's data assets is the Data Management Framework, where functional commitments for specific data management tasks are allocated between projects and the ZEF-DMU. Key tasks in data management are:

- The development of data management plans (or comparable documents) containing project specific data management policies and directives on data standards to apply and responsibilities to allocate;
- The preparation of data by following standards in terms of quality control, reformatting, file-naming, metadata collection and controlling, allocation of background information[23];
- Integration of data and metadata in the ZEF-DP and the management of access rights per data set level;
- The integration of web-links to access data when cited in ZEF publications;
- The support of ZEF-DIS data providers in managing their data in the ZEF-DIS through training and technical support.

3.2 Responsibilities

The ZEF-DMU has various responsibilities:

- Technical administration of the ZEF-DIS assuring:
  - 24/7 service availability;
  - Protection of data and personal user information against unauthorized access;
  - Protection of the databases against data loss;
- Management of functional features and thematic content in the ZEF-DP with:
  - Up-dating the ZEF-DP web-interface layout;
  - Up-dating or extension the functional features on demand;
  - Integration of thematic items used for searching of project data such as data categories and keyword-thesauri;

[23] The use of data standards in data exchange environments supports interoperability between applications. Standards are outlined in the ZEF Data Management Handbook or within project-specific data management plans

- Publication and up-dating of a ZEF Data Management Handbook with guidelines how to use the ZEF-DP and basic information on required data standardizations;

- In collaboration with projects, the development of data management plans or guidelines for one or several projects (as appropriate) containing:
  - A Memorandum of Understanding with regard to the ZEF-DMP as far as required (section 1.3);
  - Specification of standards to be applied to the data before sharing them;
  - Specification of licenses[24] defining the re-use of data;
  - Information on data to be integrated into the ZEF-DIS including retention periods;
  - Assignment of staff or position-related responsibilities within projects for:
    - Membership in the data management committee;
    - Implementation respectively control of data standardization;
    - Data quality assurance measures;
    - Creation of metadata templates;
    - Uploading of data and metadata into the ZEF-DIS and reviewing them before publishing;
    - Implementation of data exchange rules by the management of user accounts in the ZEF-DIS.

References:

De Cock Buning, M., van Dinther, B., Jeppersen-de Boer, C.H., Ringnalda, A. (2011): The legal status of research data in Germany. Annex 3 to the Knowledge Exchange report „The legal status of research data in the Knowledge Exchange partner countries". Online: http://www.knowledge-exchange.info/Default.aspx?ID=461.

German Research Foundation (2013): Safeguarding Good Scientific Practice. Recommendation 7: Safeguarding and Storing of Primary Data. WILEY-VCH, Weinheim.

High Level Expert Group on Scientific Data (2010): Riding the wave. How Europe can gain from the rising tide of scientific data. A submission to the European Commission. European Union.

Wissenschaftsrat (2012): Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020. Berlin.

[24] Based on the CreativeCommons license framework