



FDMentor



Deutsche Initiative für
Netzwerkinformation e.V.

nestor

Train-the-Trainer Concept on Research Data Management

Version 3.0

Katarzyna Biernacka
Petra Buchholz
Dr. Dominika Dolzycka
Kerstin Helbig
Dr. Janna Neumann
Dr. Carolin Odebrecht
Cord Wiljes
Dr. Ulrike Wuttke

Translated by Dr. Maik Bierwirth

AUTHORS

Katarzyna Biernacka  0000-0002-6363-0064
Petra Buchholz  0000-0002-2401-1543
Dr. Dominika Dolzycka  0000-0002-6177-8815
Kerstin Helbig  0000-0002-2775-6751
Dr. Janna Neumann  0000-0002-0161-1888
Dr. Carolin Odebrecht  0000-0003-4887-7701
Cord Wiljes  0000-0003-2528-5391
Dr. Ulrike Wuttke  0000-0002-8217-4025

TRANSLATOR

Dr. Maik Bierwirth  0000-0003-1042-6702

This publication was funded by the Federal Ministry of Education and Research as part of the joint project "FDMentor" with the funding numbers 16FDM010 and 16FDM011.

Version 3 has been extended by members of the Sub-Working Group Training/Further Education of the DINI/nestor Working Group Research Data.



IMPRINT

Except otherwise noted, the content of the "Train-the-Trainer Concept on Research Data Management" by Katarzyna Biernacka, Maik Bierwirth, Petra Buchholz, Dr. Dominika Dolzycka, Kerstin Helbig, Dr. Janna Neumann, Dr. Carolin Odebrecht, Cord Wiljes and Dr. Ulrike Wuttke is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>.



The graphic representations are based on the graphics by Patrick Hochstenbach in Bezjak Sonja et al.: *Open Science Training Handbook*. 2018. <https://doi.org/10.5281/zenodo.1212496>

CITE AS

Biernacka, Katarzyna, Maik Bierwirth, Petra Buchholz, Dominika Dolzycka, Kerstin Helbig, Janna Neumann, Carolin Odebrecht, Cord Wiljes and Ulrike Wuttke: *Train-the-Trainer Concept on Research Data Management*. Version 3.0. Berlin, 2020. <https://doi.org/10.5281/zenodo.4071471>

DOI

<https://doi.org/10.5281/zenodo.4071471>



Table of contents

Table of contents	1
Abbreviations	7
Preface	8
How to use	9
Unit structure	10
Training materials	10
Schedule	10
Teaching scripts	10
Level of detail	11
Schedule of the Train-the-Trainer Workshop on Research Data Management	12
Unit 1: Meet and greet	13
Learning objectives	13
Key aspects	13
Contents	13
Didactic methods and exercises	14
Training materials	14
Additional sources	14
Template: Harvest sheet	15
Template: We and I	16
Teaching script: Meet and greet	17
Unit 2: Orientation	18
Learning objectives	18
Key aspects	18
Contents	18
Didactic methods and exercises	19
Training materials	19
Additional sources	19
Template: Workshop map for the Train-the-Trainer Workshop on Research Data Management	20
Thematic map for the first day of the Train-the-Trainer Workshop on RDM	21
Teaching script: Orientation	22
Unit 3: Didactic approach	23
Learning objectives	23
Key aspects	23
Contents	23
Didactic methods and exercises	25
Training materials	25
Additional sources	25
Template: Cards for “Flip and turn” for the process according to Klaus Döring	27
Teaching script: Didactic approach	33

Unit 4: Digital research data	34
Learning objectives	34
Key aspects.....	34
Contents.....	34
Didactic methods and exercises.....	38
Training materials.....	38
Additional sources.....	38
Template: Cards for “Flip and turn” for the research data lifecycle.....	40
Poster: How FAIR are your research data?	43
Teaching script: Digital research data	44
Unit 5: Research data policies.....	46
Learning objectives	46
Key aspects.....	46
Contents.....	46
Didactic methods and exercises.....	49
Training materials.....	49
Additional sources.....	49
Example of an institutional research data policy: Humboldt-Universität zu Berlin RDM Policy	50
Teaching script: Research data policies.....	52
Unit 6: Data management plan.....	53
Learning objectives	53
Key aspects.....	53
Contents.....	53
Didactic methods and exercises.....	57
Training materials.....	57
Additional sources.....	57
Example: Data management plan	58
Teaching script: Data management plan.....	60
Unit 7: Order and structure	61
Learning objectives	61
Key aspects.....	61
Contents.....	61
Didactic methods and exercises.....	64
Training materials.....	64
Additional sources.....	64
Worksheet: Order and structure	65
Checklist: Versioning.....	66
Teaching script: Order and structure	67
Unit 8: Documentation and metadata	68
Learning objectives	68
Key aspects.....	68

Contents.....	68
Didactic methods and exercises.....	71
Training materials.....	71
Additional sources.....	71
In practice: How to start with data documentation?.....	72
Teaching script: Documentation	73
Unit 9: Storage and backup	75
Learning objectives	75
Key aspects.....	75
Contents.....	75
Didactic methods and exercises.....	77
Training materials.....	77
Additional sources.....	77
Worksheet: Storage and backup	78
Proposed solutions for the worksheet: Storage and backup	79
Argumentation aid: Why should I save my data and use a backup?.....	80
Checklist: What do I have to consider when choosing a storage device?	81
Teaching script: Storage and backup.....	82
Unit 10: Long-term archiving	83
Learning objectives	83
Key aspects.....	83
Contents.....	83
Didactic methods and exercises.....	85
Training materials.....	85
Additional sources.....	85
Checklist: What should be considered when deciding on a long-term archive?	86
Teaching script: Long-term archiving	87
Unit 11: Access control.....	88
Learning objectives	88
Key aspects.....	88
Contents.....	88
Didactic methods and exercises.....	90
Training materials.....	90
Additional sources.....	90
Teaching script: Access control.....	91
Unit 12: Formal framework	92
Learning objectives	92
Key aspects.....	92
Contents.....	92
Didactic methods and exercises.....	94
Training materials.....	94

Additional sources.....	94
Template: Mind map on the subject of formal framework.....	95
Checklist: Organisation of workshops	96
Teaching script: Formal framework	98
Unit 13: End of the first day	99
Learning objectives	99
Key aspects.....	99
Contents.....	99
Didactic methods and exercises.....	99
Training materials.....	100
Teaching script: End of the first day	101
Unit 14: Welcome and orientation on day 2	102
Learning objectives	102
Key aspects.....	102
Contents.....	102
Didactic methods and exercises.....	103
Training materials.....	103
Thematic map for the second day of the Train-the-Trainer Workshop on RDM	104
Template: Keyword strips.....	105
Teaching script: Welcome and orientation on day 2.....	106
Unit 15: Data publication	107
Learning objectives	107
Key aspects.....	107
Contents.....	108
Didactic methods and exercises.....	113
Training materials.....	114
Additional sources.....	114
Template: Statement slam	116
Poster: How to publish research data?	118
Teaching script: Data publication	119
Unit 16: Re-use of research data.....	121
Learning objectives	121
Key aspects.....	121
Contents.....	121
Didactic methods and exercises.....	124
Training materials.....	125
Worksheet: Citation.....	126
Teaching script: Re-use of research data	127
Unit 17: Legal aspects	128
Learning objectives	128
Key aspects.....	128

Contents.....	128
Didactic methods and exercises.....	131
Training materials.....	131
Additional sources.....	131
Exercise: De-identification of qualitative data.....	133
Answers: De-identification of qualitative data.....	134
Checklist: Requirements for a consent according to GDPR	135
Teaching script: Legal aspects.....	137
Unit 18: Institutional infrastructure	138
Learning objectives	138
Key aspects.....	138
Contents.....	138
Didactic methods and exercises.....	139
Training materials.....	139
Storage services of the Computer and Media Service of the Humboldt-Universität zu Berlin	140
Teaching script: Institutional infrastructure	141
Unit 19: Training exercise.....	142
Learning objectives	142
Key aspects.....	142
Contents.....	142
Training materials.....	142
Worksheet: Data management plan.....	143
Exemplary solution: Data management plan.....	145
Research data quiz	147
Solution: Research data quiz	148
Teaching script: Training exercise.....	149
Unit 20: 7 steps of concept development.....	150
Learning objectives	150
Key aspects.....	150
Contents.....	150
Didactic methods and exercises.....	152
Training materials.....	152
Additional sources.....	152
Template for a teaching script.....	153
Teaching script: 7 steps of concept development	154
Unit 21: Didactic methods.....	155
Learning objectives	155
Key aspects.....	155
Contents.....	155
Didactic methods and exercises.....	158
Training materials.....	158

Additional sources.....	158
Template: Schema-X	159
Explanation of methods.....	160
Teaching script: Didactic methods	170
Unit 22: Feedback and farewell.....	171
Learning objectives	171
Key aspects.....	171
Contents.....	171
Didactic methods and exercises.....	172
Training materials.....	173
Additional sources.....	173
Example of a feedback form for the Train-the-Trainer workshop on research data management..	174
Example of a certificate of attendance for participation in the Train-the-Trainer workshop on research data management	176
Teaching script: Feedback and farewell.....	177
General resources on the topic	179

Abbreviations

BMBF	Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research)
CC	Creative Commons
DFG	Deutsche Forschungsgemeinschaft (German Research Foundation)
DMP	Data management plan
EC	European commission
L	Long version
M	Medium version
PPTX	PowerPoint slides
RD	Research Data
RDM	Research data management
S	Short version
T	Trainer

Preface

This version is the English translation of the German Train-the-Trainer Concept on Research Data Management in version 3.0¹. The underlying concept was developed within the project FDMentor², which received funding by the German Federal Ministry of Education and Research (BMBF) from 2017 to 2019 (funding code 16FDM010 and 16FDM011). Unfortunately, it was not possible to provide the concept in an English version during the funding phase. Since the Train-the-Trainer Programme has gained a lot of interest both nationally and internationally, employees of the Humboldt-Universität zu Berlin and the Freie Universität Berlin nevertheless dedicated themselves to a translation of the concept and training materials after the completion of the project. This translation has now been finalized with the support of the two participating universities, making the concept accessible and attractive for a broad international readership.

The primary target group for this version of the concept are employees of central facilities as well as data stewards and researchers in Germany who do not speak German or who would like to offer training in research data management in English. As a result, the concept was not adapted to international research data management. For example, in the unit "Legal aspects", it does not refer to international law, but to German and European law. Despite the reference to Germany, the concept can still be useful for an international readership. Due to the free licensing of the concept and all training materials (Creative Commons Attribution CC BY 4.0 International), it is easy to supplement or adapt the concept to local requirements. This is even encouraged by the authors.

Since all authors involved are non-native English speakers, the concept and materials were translated to the best of our knowledge. We ask for your understanding for any non-idiomatic formulations or inaccuracies that may occur as a result. We would be pleased to receive comments and suggestions of any kind in order to improve the translation of the concept. The Train-the-Trainer Concept is actively being further developed within the DINI/nestor-AG Forschungsdaten³ (Working Group Research Data) – with DINI and nestor being national non-profit associations of infrastructure facilities such as data centers, libraries and archives. In addition, the follow-up project of FDMentor called "FDNext"⁴, which is funded by the German Research Foundation (DFG), will develop a discipline-specific train-the-trainer module within the next three years.

We hope the Train-the-Trainer Concept on Research Data Management will be used extensively in both German and English and serves as an inspiration for many more courses.

We now wish you a pleasant reading!

Katarzyna Biernacka, Maik Bierwirth, Petra Buchholz, and Kerstin Helbig

Berlin, November 2020

¹ Biernacka, Katarzyna, Petra Buchholz, Dominika Dolzycka, Kerstin Helbig, Janna Neumann, Carolin Odebrecht, Cord Wiljes and Ulrike Wuttke: *Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement* (version 3.0). Zenodo, 2020. <https://doi.org/10.5281/zenodo.3938533>.

² Helbig, Kerstin, Katarzyna Biernacka, Petra Buchholz, Dominika Dolzycka, Niklas Hartmann, Thomas Hartmann, Bea M. Hiemenz, Boris Jacob, Monika Kuberek, Nadin Weiß, and Malte Dreyer: Lösungen und Leitfäden für das institutionelle Forschungsdatenmanagement. *o-bib* 6(3), 2019, pp. 21-39. <https://doi.org/10.5282/o-bib/2019H3S21-39>.

³ DINI e.V. „DINI/nestor-AG Forschungsdaten.“ Last accessed 25.09.2020, <https://dini.de/ag/dininestor-ag-forschungsdaten/>.

⁴ Forschungsdaten.org. „FDNext.“ Last accessed 25.09.2020, <http://www.forschungsdaten.org/index.php/FDNext>.

How to use

This document is primarily intended for trainers who want to conduct a Train-the-Trainer workshop on research data management. It contains background knowledge on the PowerPoint slides and teaching scripts as well as further information on the individual subject areas required for reuse and implementation of a two-day workshop of seven and a half hours a day. The contents are spread over six hours per day and are suspended by one lunch break (one hour) and two shorter breaks (fifteen minutes each). It is also possible to reuse only individual parts of the concept. In general, the contents of this concept can also be used by trainers who design and conduct RDM workshops for researchers. The trainers themselves may already have advanced or even in-depth previous knowledge, so that they often do not learn much new content. For this group of people, however, the added value lies in the units, methods and templates which can be reused directly.

In order to prepare for a training, it is useful to work through the respective unit in the concept, to follow the teaching scripts and to check the PowerPoint slides. The concept contains units that deal with four different components (see Figure 1⁵).

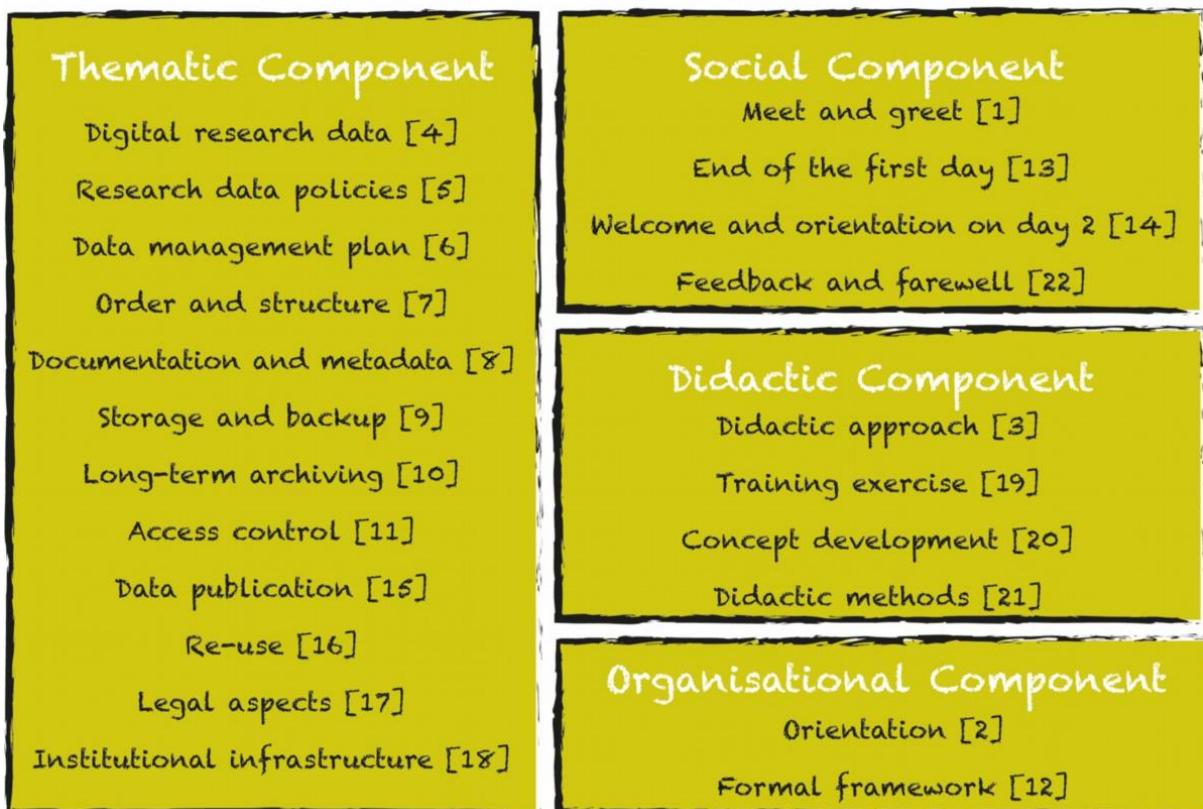


Figure 1: Four components of the Train-the-Trainer concept. The indicated numbers correspond to the unit numbers in the document.

Next, it is important to familiarize yourself with the methods and look them up in unit 21. Trainers can always seek and acquire more background knowledge from other sources as well. The required documents can be printed from the ZIP document "WorkingMaterials_FDMentor_V3" and duplicated, if necessary.

The sequence of the units is identical to the workshop schedule (see p. 12). There is one unit for each aspect of research data management and two are focused on didactic concepts and workshop design.

⁵ Based on Biernacka, Katarzyna, Kerstin Helbig, Petra Buchholz and Dominika Dolzycka: Forschungsdatenmanagement schulen – ein Train-the-Trainer Programm zur Kompetenzvermittlung. *Information - Wissenschaft & Praxis* 70 (2019), p. 266. <https://doi.org/10.1515/iwp-2019-2040>.

The workshop starts with unit 1: 'Meet and greet' (p. 13) and ends with unit 22: 'Feedback and farewell' (p. 171).

Unit structure

All units follow the same structure. At the beginning of each unit the learning objectives are defined, i.e. what should be grasped when reading the unit. The next step is specifying the main topics and elaborating on the contents of these topics. These contents should inform the trainers about the respective topic and enable them to reproduce this unit of the workshop with the help of the PowerPoint slides.

Furthermore, the didactic methods and exercises used in the teaching script are listed as well as the training materials needed for it. More detailed descriptions required to guide the exercises can be found in unit 21: 'Didactic methods' (p. 155). It lists goals, time requirements and materials necessary and it describes the procedure. Readers are encouraged to also use the exercises for other units, to modify them or even to design their own exercises and materials.

In addition, there are further sources listed. This facilitates a more in-depth exploration of the topics. The didactic methods and exercises used in the unit, the training materials and the training scripts conclude each unit.

Training materials

The training materials are displayed for each unit. These may be worksheets, checklists, templates for flip chart sheets or materials for individual or group exercises. In unit 1: 'Meet and greet' (p. 13), for example, an illustration of the 'Harvest sheet' can be found. These illustrations are intended to help visualize the method when studying the unit during preparation. However, they are not suitable as print templates for subsequent use as they contain headings. The print templates are included in a second download document "WorkingMaterials_FDMentor_V3". These can be used directly and without further changes.

We have also included templates for working materials that refer specifically to the courses we organise. They serve as an orientation for the creation of your own materials. The presentation of the feedback questionnaire from FDMentor was thus used to evaluate the entire workshop as part of the piloting process, as well as the corresponding teaching materials. At the same time, the feedback should provide information on whether the focus and the chosen methodological formats meet the approval of the target group. This questionnaire is therefore particularly comprehensive. For your own activities, the template should be shortened and adapted. A certificate of attendance must also be drawn up individually for the event carried out.

Schedule

The schedule on page 12 shows the sequence of topics. It also includes the required time for each unit, the pauses and activations at one glance. Activations are methods to re-energize the group in case of fatigue (see unit 21 point 3, p. 156). The remaining time is called *buffer time* and is mostly used in case exercises or discussions last longer than expected.

Teaching scripts

At the end of each unit we have included the training scripts, which are described in more detail in unit 20 point 4 (p. 151). Using these, the readers learn the exact sequence of the unit comparable to a 'production script'. Each training script is structured according to subtopics, called 'modules'. In the first unit the modules include "Introduction", "Welcome" and "Getting to Know". Each module contains one or more steps for the workshop, shown as a table. The headings are structured as follows:

- module,
- the learning goal to be achieved within this module,
- required time in minutes,

- description of the content of this step,
- specification of the work form, for example, individual or partner work,
- list of required materials,
- assignment to ‘inhalation’ and ‘exhalation’ (see unit 3 point 2, p. 24),
- making voices sound, indicating whether the participants speak or remain silent in this step (see unit 3 point 3, p. 25),
- alternatives: for some steps you will find other suggestions here, and
- remarks.

In addition, there is also the ZIP document “WorkingMaterials_FDMentor_V3”, which contains all further documents required:

- schedule of the two-day workshop as PDF and DOCX,
- PowerPoint slides for both workshop days,
- training scripts as PDF and DOCX files,
- print templates of working materials for exercises, partly with solution sheets, as PDF and DOCX files,
- templates for flipchart sheets,
- checklists,
- templates for your own training scripts and attendance reports, and
- feedback form

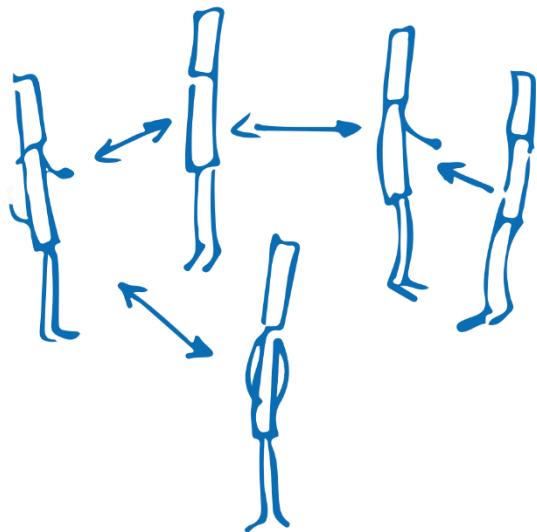
PDF documents may be printed out and re-used. If adjustments or changes are required, the Word files are also included and can be easily modified.

Level of detail

This Train-the-Trainer programme is designed on the assumption that the researchers to be trained have little or no previous knowledge of RDM. Since the concept was developed as an introduction for this target group, it does not go into too much detail in specific subject areas, but provides an overview as complete as possible and points out to further information sources.

Schedule of the Train-the-Trainer Workshop on Research Data Management

Unit No.	Topic	Duration	Page No.
First day			
1	Meet and greet	25 min	p. 13
2	Orientation	20 min	p. 18
3	Didactic approach	20 min	p. 23
4	Digital research data	45 min	p. 34
<i>Coffee break</i>		15 min	
Activation		5 min	
5	Research data policies	15 min	p. 46
6	Data management plan	20 min	p. 53
7	Order and structure	20 min	p. 61
<i>Lunch break</i>		60 min	
Activation		5 min	
8	Documentation and metadata	50 min	p. 68
9	Storage and backup	10 min	p. 75
10	Long-term archiving	20 min	p. 83
<i>Coffee break</i>		15 min	
Activation		5 min	
11	Access control	15 min	p. 88
12	Formal framework	30 min	p. 92
13	End of the first day	20 min	p. 99
Second day			
14	Welcome and orientation on day 2	15 min	p. 102
15	Data publication (Part 1)	55 min	p. 107
<i>Coffee break</i>		15 min	
15	Data publication (Part 2)	35 min	see above
16	Re-use	30 min	p. 121
17	Legal aspects	25 min	p. 128
<i>Lunch break</i>		60 min	
Activation		5 min	
18	Institutional infrastructure	10 min	p. 138
19	Training exercise	20 min	p. 142
<i>Coffee break</i>		15 min	
Activation		5 min	
20	7 steps of concept development	15 min	p. 150
21	Didactic methods	55 min	p. 155
22	Feedback and farewell	50 min	p. 171



Unit 1: Meet and greet

Learning objectives

The participants get to know each other and the workshop instructors.

Key aspects

1. Welcome
2. Getting to know each other and socializing

Contents

1. Welcome

The welcome is an important aspect in every workshop. The length of the welcoming phase itself is not important. However, it is necessary to offer the participants a structured introduction and start of the event. It is therefore advisable to announce the title of the workshop in order to attract their attention and to resolve any room confusion right at the beginning. The participants thus arrive at the workshop not just physically, but also mentally. The beginning of the workshop is also a good time to recall or briefly present the rules for the workshop, the so-called Code of Conduct (see unit 12: Formal framework, p. 92), or to jointly develop them in case of a longer workshop.

With the first words of the trainer, the participants pay attention to him/her. Whether one manages to arouse interest or irritation right from the start, the participants' attention is ensured.

2. Getting to know each other and socializing

People work with more openness and initiative when they know their counterparts.⁶ Especially for workshops of several days it is recommended to give the group time to get to know each other.⁷ In terms of content, it also pays off if a relaxed, open and trusting working atmosphere is created. Discussions

⁶ Szepansky, Wolf-Peter: *Souverän Seminare leiten*. Bielefeld: Bertelsmann, 2006, p. 16-18.

⁷ Cohn, Ruth: *Von der Psychoanalyse zur themenzentrierten Interaktion: von der Behandlung einzelner zu einer Pädagogik für alle*. Stuttgart: Klett, 1975, p. 112.

become more profound and it is more likely that participants will address problems or ambiguities and actively contribute.

The most important advantage, however, is that the participants are included in the activities at a very early stage. The earlier they speak, the more likely it is that they will speak during the workshop and participate actively, as the first inhibition threshold has already been crossed (see unit 3 point 3, p. 25).

Didactic methods and exercises

Methods:

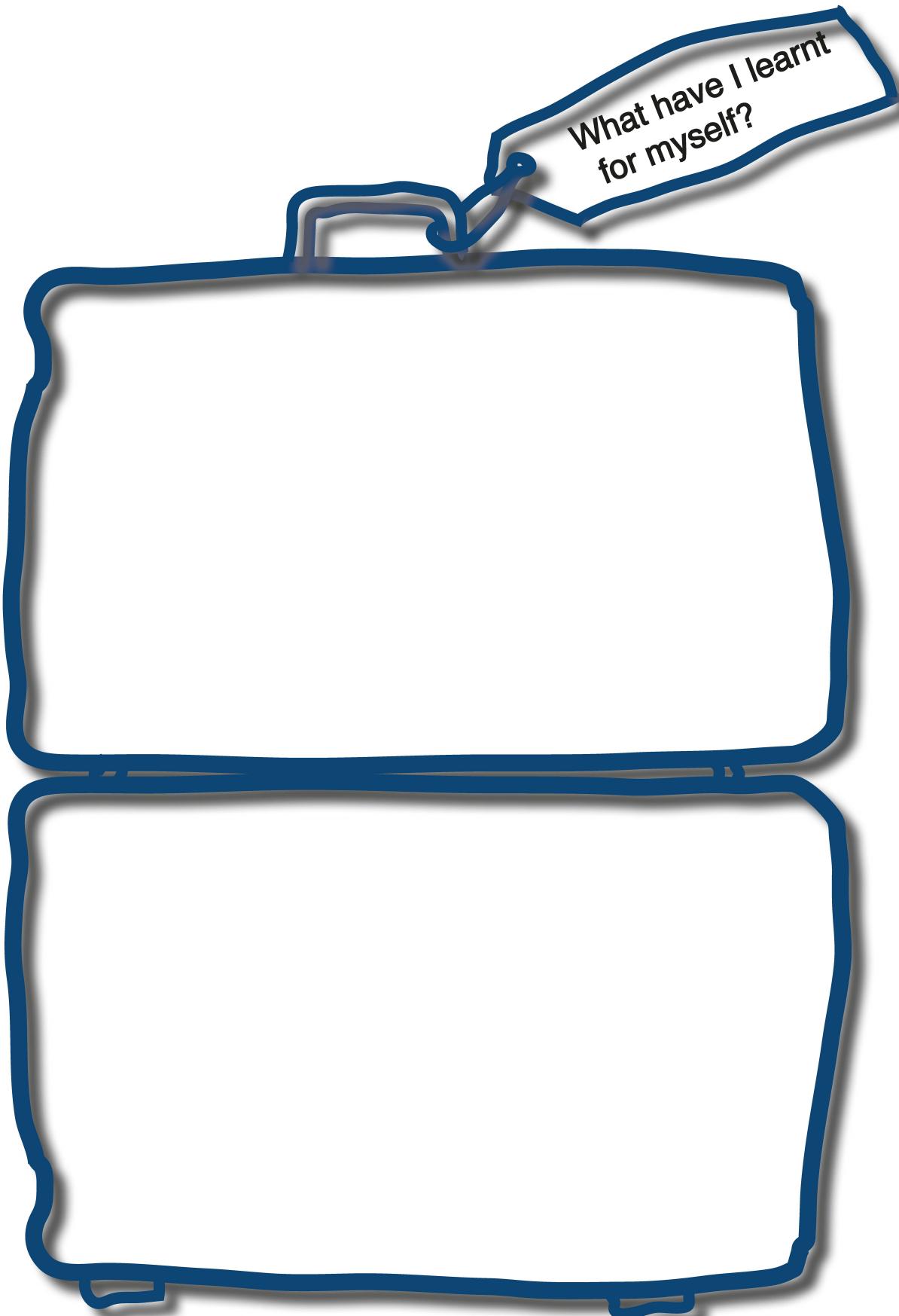
- Landscape setting
- We and I

Training materials

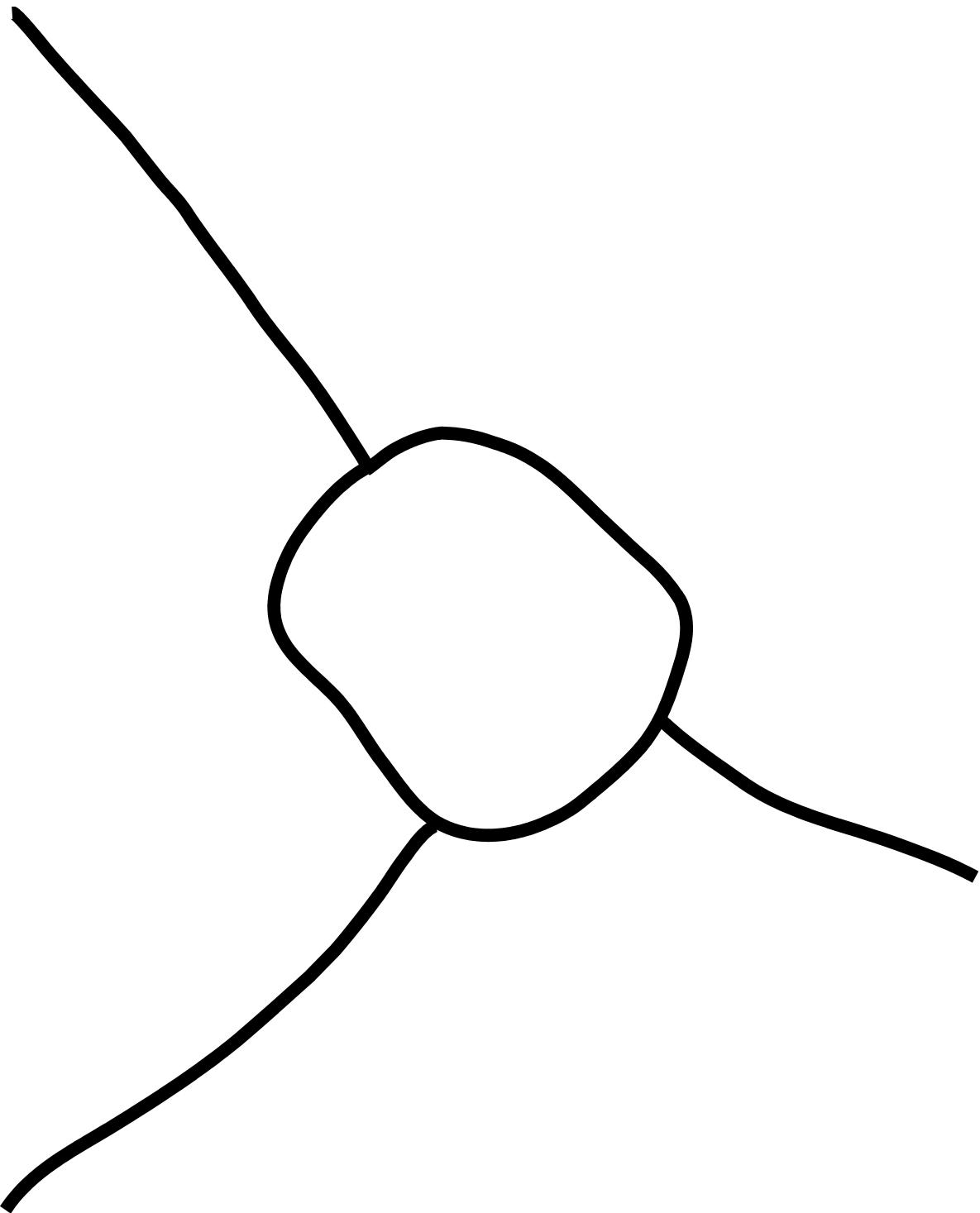
- Template: Harvest sheet
- Template: We and I
- Teaching script: Meet and greet

Additional sources

- Szepansky, Wolf-Peter: *Souverän Seminare leiten*. Bielefeld: Bertelsmann, 2006.
- Cohn, Ruth: Von der Psychoanalyse zur themenzentrierten Interaktion: von der Behandlung einzelner zu einer Pädagogik für alle. Stuttgart: Klett, 1975.



Template: We and I



1. Teaching script: Meet and greet									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternative	Notes
Presentation	The participants overcome the inhibition threshold to speak	1	Participants give their names	Method: Ideas out loud	-	Ex	Yes		
Welcome	T give their names	1	Greet, T give their names	Presentation	-	In			
	The participants get to know each other and are activated by movement	4	The participants line up in the room and welcome each other, naming 1. name, 2. institution, 3. area of responsibility Time per appointment: 20 s	Exercise: Introduction round	Bell & Stopwatch/ Timer	Ex	Yes	Speed dating (but not all participants get to know each other)	Participants tend to speak longer, therefore very clear announcement important: 1. name, 2. institution, 3. area of responsibility
Get to know	The participants get to know each other a little better	8	The participants will meet in groups of 3, the persons should not know each other beforehand. Each group receives one sheet of paper. They note 3 similarities and 3 differences on it	Method: We and I	Flip chart sheets, pens	Ex	Yes		
	Participants speak in front of the group	8	The groups present their sheet to the other participants	Exercise	Pins and pin boards	Ex	Yes		
	The participants get to know the seminar instructors	2	The T introduce themselves	Lecture	-	In		With two trainers, they introduce each other with We and I	
	Participants create name badges	1	Create name badges	Exercise	Empty name badges	-	-	Prepared name badges	

Duration of the unit: 25 minutes



Unit 2: Orientation

Learning objectives

The participants will receive an overview of the contents of the entire workshop by means of a workshop map.

Participants become aware of their expectations of the workshop.

The participants learn the schedule of the day.

Key aspects

1. O! O! O! - Orientation
2. Workshop map
3. Schedule of the day

Contents

1. O! O! O! - Orientation

The orientation section at the beginning of the workshop is intended to give participants an overview of the topics to be dealt with as well as the course of events. The number of breaks and their approximate timing should also be mentioned here. The participants get a sense for the topic and the process and can mentally adjust to both. The thematic orientation helps to activate previous knowledge. As the participants learn what to expect, uncertainty is reduced. Learning is easier in a relaxed mindset. The following methods facilitate the orientation of the participants and give structure to the process.

Orientation is a good time to clarify organizational issues. A 'code of conduct' for the event can be agreed on with the participants. The announcement of feedback and the question whether there will be certificates of attendance should be included here. The workshop instructors can explain their expectation to the participants: Do they wish the participants to actively participate, for example, and would they like questions in between, or would they prefer them to be asked at the end of a unit?

2. Workshop map

A workshop map is a visual aid for orientation. It is an overview of all topics covered in the workshop. In addition to the title, it contains keywords that should be short and concise. Pictograms, shapes and colours will loosen up the illustration. This map provides participants with a more global, broad overview. Details are omitted in order to avoid overburdening the participants when introducing the topics. The participants are told briefly what each item on the map is all about.

In the course of the workshop, the map can be used again for orientation when topics change in order to discuss: what have we already done, what comes next?

3. Schedule of the day

The schedule of the day draws from those parts of the workshop map that will be covered on that day. It also provides information on how the breaks will be distributed throughout the day. The exact times can be omitted here in order to allow more flexibility for the workshop leader when dealing with individual subject areas, for example:

Start (10 a.m.)

Meet and Greet

Orientation, O, O, O

Coffee break (15 minutes)

Topic 1

Lunch break (1 hour)

Topic 2

Topic 3

Coffee break (15 minutes)

Topic 4

Feedback and farewell

End (5.30 p.m.)

Didactic methods and exercises

Methods:

- Expectation query
- Workshop map

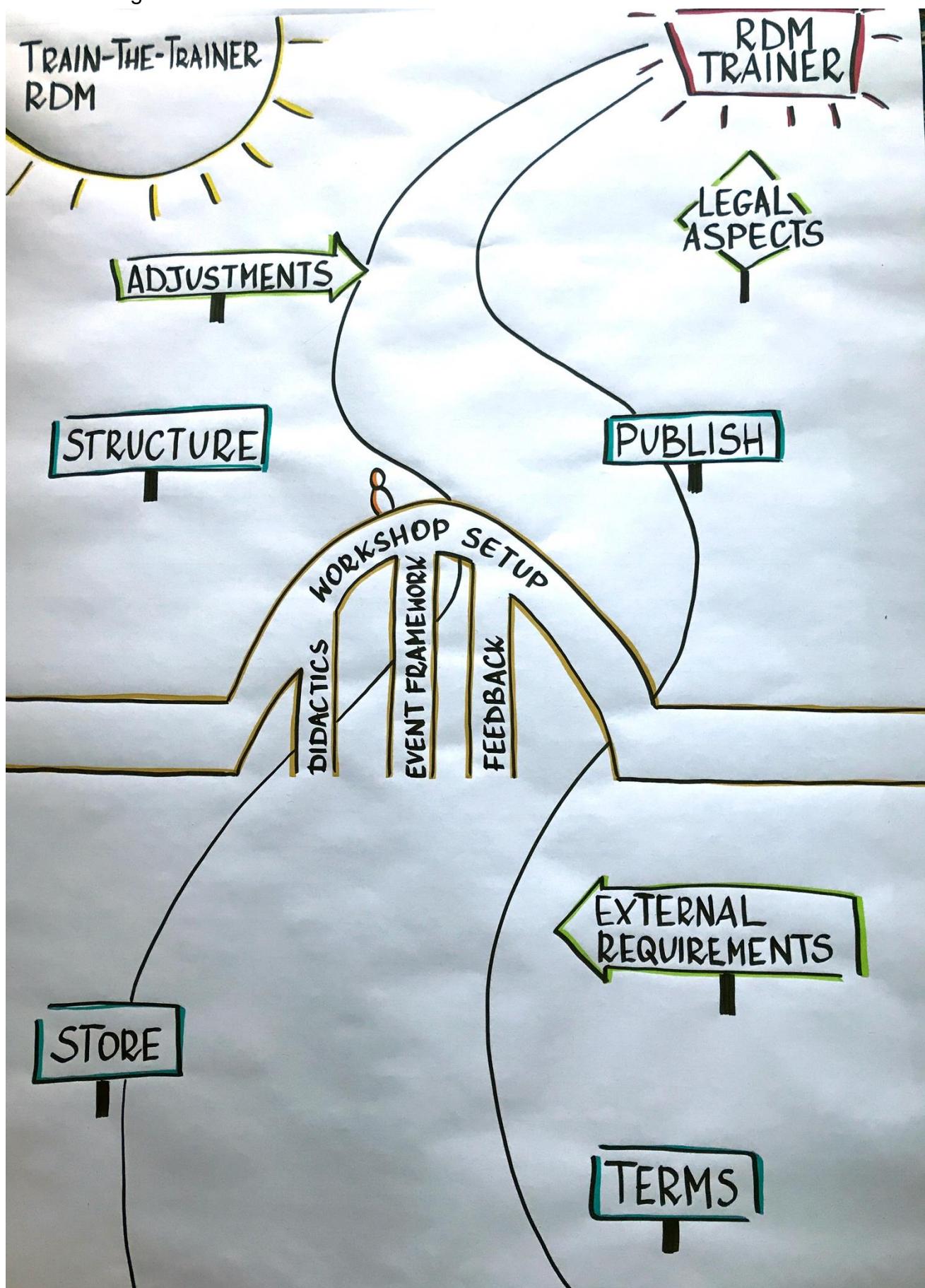
Training materials

- Workshop map for the Train-the-Trainer workshop on Research Data Management
- Thematic map for the first day of the Train-the-Trainer Workshop on RDM
- Teaching script: Orientation

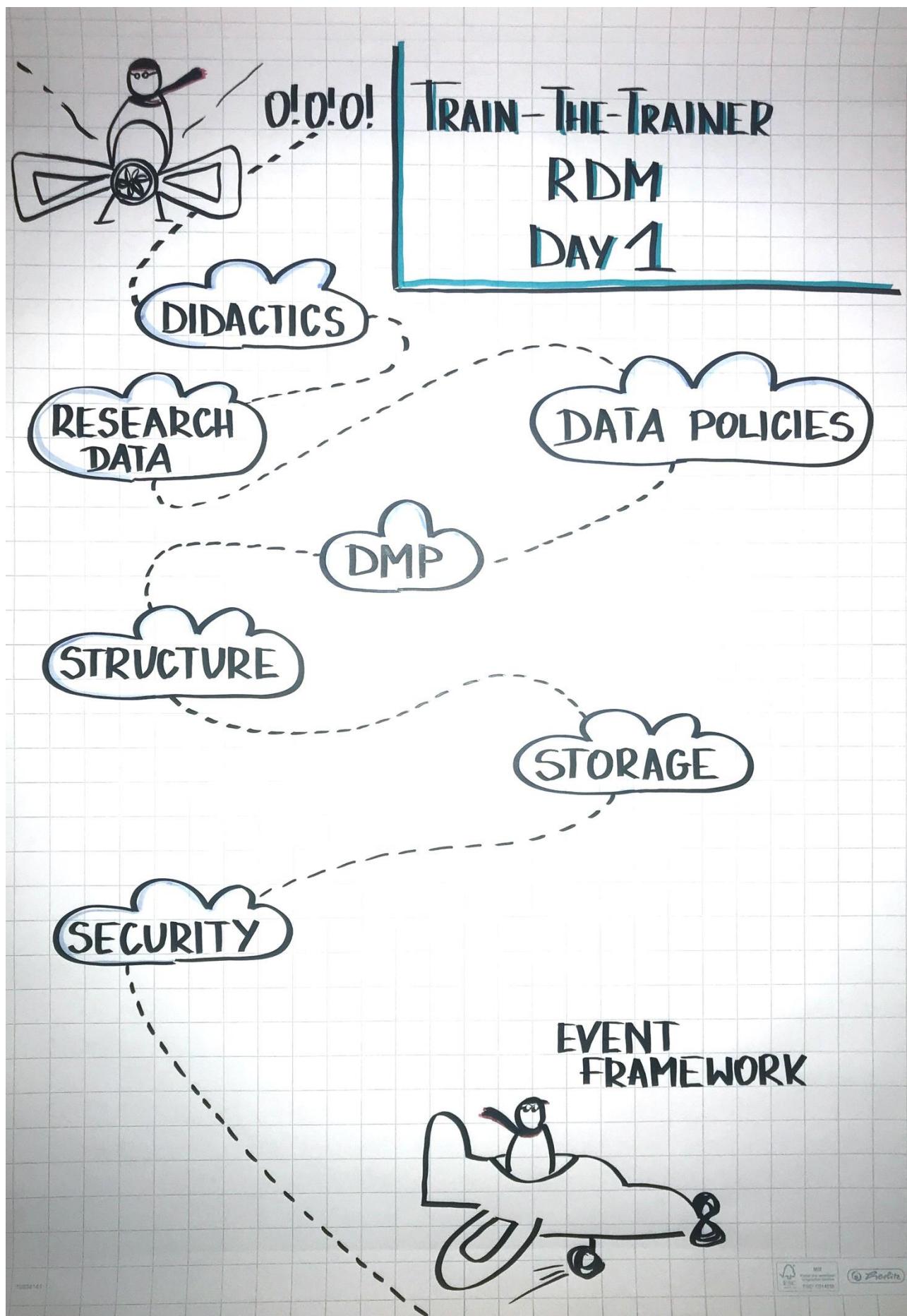
Additional sources

- Haussmann, Martin and Holger Scholz: *bikablo. Das Trainerwörterbuch der Bildsprache*. 2nd ed. Eichenzell: Neuland, 2007.
- Haussmann, Martin and Holger Scholz: *bikablo 2.0: neue Bilder für Meeting, Training & Learning*. Eichenzell: Neuland, 2009.

Template: Workshop map for the Train-the-Trainer Workshop on Research Data Management

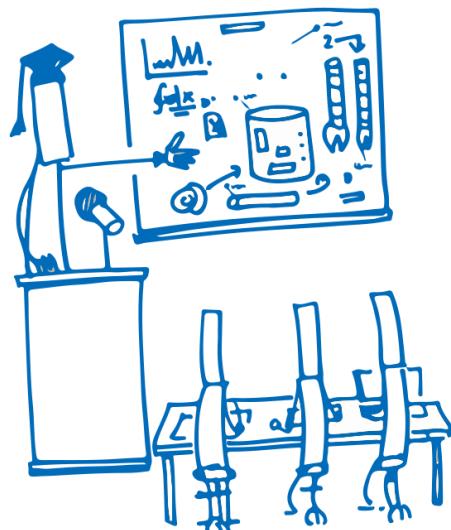


Thematic map for the first day of the Train-the-Trainer Workshop on RDM



2. Teaching script: Orientation						
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation
					Voices sound	Alternative
Workshop map	Participants receive an overview of the contents of the entire workshop in the form of a workshop map	3	Explaining the workshop map	Presentation	Workshop map on sheet; pinboard	In -
Workshop map	Participants clarify & express their own expectations / wishes	5	There are cards in front of the participants. They should reflect: "What is most important to me in this workshop?" and write it down on the card	Method: Expectation query	Moderation cards, thick pens	Ex -
Workshop map	Participants locate request on workshop map	10	The participants present their wish and pin it on the map, where it fits approximately	Participants explain their wish for the workshop & locate it on the workshop map	Pins	Ex Yes
Know the day's schedule	Participants get to know the day's schedule	2	T presents day's schedule, with topics and breaks - Swabian savings plan	Prepared day's schedule	Prepared day's schedule	In -

Duration of the unit: 20 minutes



Unit 3: Didactic approach

Learning objectives

Participants will learn practical tips for inspiring teaching.

The participants are introduced to the learning process according to Döring.

Key aspects

1. Stimulate learning
2. The learning process according to Klaus Döring
3. A supportive learning atmosphere

Contents

1. Stimulate learning

In order to convey knowledge well, it is necessary to find a suitable method to attract the attention of the participants and to stimulate their interest in the topic. Reciting profound knowledge in a short amount of time is not sufficient. In order for the participants to actually learn something, the learning process must first be initiated in them.

There are several ways which have proven helpful in practice to enable participants to learn well. Many of them seem to cost valuable learning time at first, but in the end they pay off. According to Harald Groß⁸, these include:

- stimulate the appetite for learning
- direct learners' attention
- provide processing aids
- slow down the learning process (thinking time, silence, test duration)
- provide for "refreshing minutes" early on and frequently
- have the participants explain in their own words what they have learned

⁸ Groß, Harald: „Arbeitsheft Nr. 1: Das Lernen auslösen.“ In Harald Groß (Ed.): *Didaktik*. Orbium Seminare, internal seminar material.

- enable transfer options
- frequent exercises

This allows participants to repeat what they have learned, transfer it to their daily work routine, think about it or take notes.

2. The learning process according to Klaus Döring

Different learning models have been created over the years. In addition to Klaus Döring's model⁹ presented and used in this concept, there are, for example, Rolf Arnold's LENA model¹⁰ and Horst Siebert's constructivist methods¹¹. Furthermore, Andreas Gold offers a sound overview of empirical teaching and learning research.¹²

According to Döring, learning is a process consisting of two phases: "inhalation" and "exhalation" (cf. Figure 2). Alternately, the learner absorbs knowledge and reproduces it. In the teaching scripts of this concept, activating pre-existing knowledge is also seen as exhaling. Activating pre-existing knowledge facilitates learning because the new knowledge is linked to the old knowledge. The receptive phase (inhalation) should not be longer than 20 minutes.

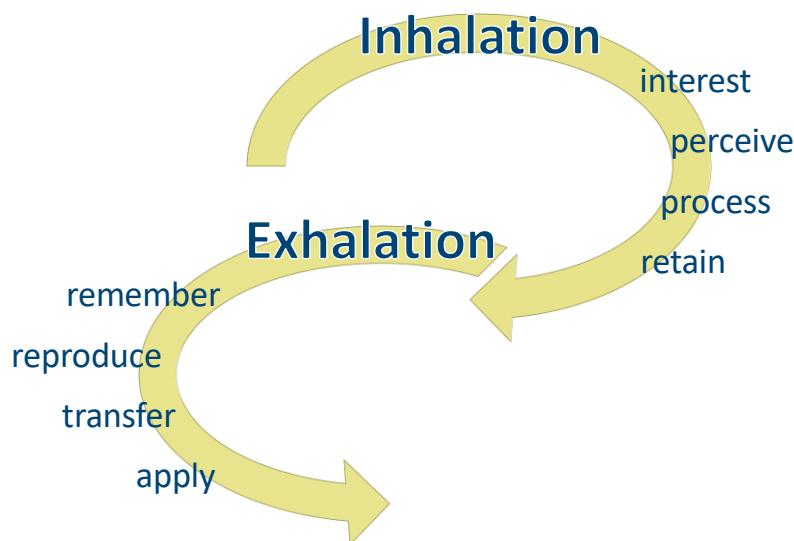


Figure 2: Learning process according to Klaus Döring.

In order to stimulate learning at all, Döring states that the learners' interest in the subject-matter must be awakened to ensure that they perceive it. New knowledge is processed and firmly anchored. The learning model implies one can only speak of learning if one is also able to remember what has been learned, to reproduce it, to transfer it to another context, and finally to apply it.

⁹ Döring, Klaus W.: *Handbuch Lehren und Trainieren in der Weiterbildung*. Weinheim: Beltz, 2008, pp. 57–58.

¹⁰ Arnold, Rolf: *Wie man lehrt, ohne zu belehren: 29 Regeln für eine kluge Lehre*, 4th ed. Heidelberg: Carl-Auer, 2018.

¹¹ Siebert, Horst: *Didaktisches Handeln in der Erwachsenenbildung: Didaktik aus konstruktivistischer Sicht*. 7th ed. Augsburg: ZIEL, 2012.

¹² Gold, Andreas: *Guter Unterricht: Was wir wirklich darüber wissen*. Göttingen: Vandenhoeck & Ruprecht, 2015.

3. A supportive learning atmosphere

In order to facilitate learning, it is important to create a stimulating atmosphere. Therefore, the workshop leader is in charge of creating such an atmosphere.¹³ This can be achieved in particular by using two approaches.

First of all, at the beginning of the event participants should be given the opportunity to get to know each other well.^{14,15} People who are shy to speak dare to do so more often when they are among people they know. This can be achieved particularly well in a multi-stage approach during the getting to know phase. First, everyone only mentions their name. Secondly, in a relaxed situation, all participants express themselves initially in pairs, then in small groups and finally in front of everyone.

In the second phase, in order to receive contributions from as many participants as possible, it is helpful to enable work in smaller groups. The frequent work in changing groups requires participation and constructive cooperation.¹⁶ Furthermore, in smaller groups more people can have their say at the same time than in one larger group. This kind of active cooperation also contributes to a good learning atmosphere.¹⁷

We summarize both approaches under "Making voices sound" and check whether our concept adequately addresses this issue by providing a column in the teaching scripts specifically for this purpose. This column quickly indicates whether voices resonate and whether we are thus creating a supportive learning atmosphere.

Didactic methods and exercises

Methods:

- Flip and turn
 - Learning process according to Klaus Döring
- Chatter
 - Think about criteria by which you would notice this is not a research data management workshop, but a Train-the-Trainer workshop on research data management.

Training materials

- Template: Cards to "Flip and turn" for the process according to Klaus Döring
- Teaching script: Didactic approach

Additional sources

- Arnold, Rolf: *Wie man lehrt, ohne zu belehren: 29 Regeln für eine kluge Lehre*, 4. Aufl. Heidelberg: Carl-Auer, 2018.
- Clement, Ute and Klaus Kräft: *Lernen organisieren*. Berlin: Springer, 2002.
- Döring, Klaus W.: *Handbuch Lehren und Trainieren in der Weiterbildung*. Weinheim: Beltz, 2008.
- Gold, Andreas: *Guter Unterricht: Was wir wirklich darüber wissen*. Göttingen: Vandenhoeck & Ruprecht, 2015.

¹³ Szepansky, Wolf-Peter: *Souverän Seminare leiten*. Bielefeld: Bertelsmann, 2006, p. 65.

¹⁴ Lucas, Jenifer: Communication apprehension in the ESL classroom: Getting our students to talk. *Foreign Language Annals* 17(6), 1984: S. 593. <https://doi.org/10.1111/j.1944-9720.1984.tb01748.x>.

¹⁵ Will, Hermann: *Mini-Handbuch Training und Seminar*. Weinheim: Beltz, 2016, p. 63–70.

¹⁶ Szepansky, *Souverän Seminare leiten*, p. 18.

¹⁷ Clement, Ute and Klaus Kräft: *Lernen organisieren*. Berlin: Springer, 2002, p. 34.

- Lucas, Jenifer: Communication apprehension in the ESL classroom: Getting our students to talk. *Foreign Language Annals* 17(6), 1984. <https://doi.org/10.1111/j.1944-9720.1984.tb01748.x>.
- Siebert, Horst: Didaktisches Handeln in der Erwachsenenbildung: Didaktik aus konstruktivistischer Sicht. 7th ed. Augsburg: ZIEL, 2012.
- Szepansky, Wolf-Peter: *Souverän Seminare leiten*. Bielefeld: Bertelsmann, 2006.
- Will, Hermann: *Mini-Handbuch Training und Seminar*. Weinheim: Beltz, 2016.

Template: Cards for “Flip and turn” for the process according to Klaus Döring

interest

perceive

process

retain

remember

reproduce

transfer

apply

inhalation

exhalation

**receptive
phase**

**expressive
phase**

3. Teaching script: Didactic approach						
Component	Target	Time	Content	Working modes	Material	Inhalation / exhalation
					Voces sound	Alternative
Introduction	Participants activate their knowledge on special features of training	13	The participants consider criteria by which they would notice that this workshop is not an FDM Workshop but a Train-the-Trainer Workshop on the topic of FDM. Afterwards the considerations are discussed in pairs of two with the person on the left. At the end they report their results.	Method: Chatter	PPTX with question, paper, pens	Ex Yes
	Model according to Klaus Döring	5	Participants get to know the model of Klaus Döring	S: Learning with the two phases of inhalation and exhalation and their further components is presented	PPTX, flip chart with inhalation and exhalation; For L: cards for Flip and turn	In - L: Method: Flip and turn: Match terms from inhalation and exhalation to their definitions/descriptions - L: Method: Flip and turn: Match terms from inhalation and exhalation to their definitions/descriptions

Duration of the unit: 18 minutes



Unit 4: Digital research data

Learning objectives

The participants know the definition of research data and their meaning.

Participants know the research data life cycle and can distinguish between the individual steps.

Participants know what research data management entails and why it is important. They know which aspects are involved.

Key aspects

1. Research data
2. Research data lifecycle
3. Research data workflow
4. Research data management
 - a) What is research data management?
 - b) Tasks of research data management
 - c) The FAIR principles
 - d) Motivation

Contents

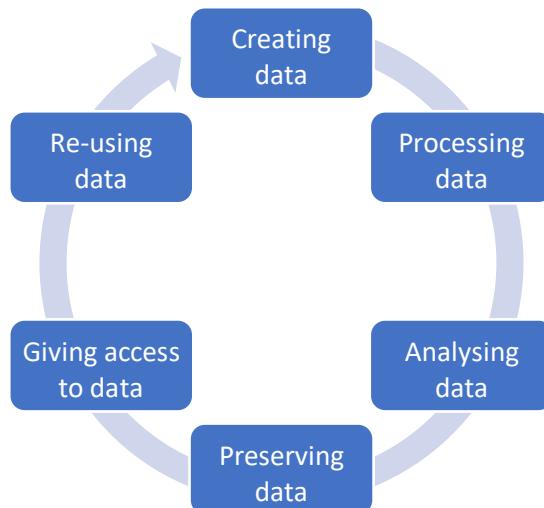
1. Research data

This concept considers only digital research data. There is no fixed definition of research data. Kindling and Schirmbacher provided one of the first definitions in 2013: "By digital research data we mean [...] all digitally available data that are generated during or as a result of the research process".¹⁸ In 2015, the German Research Foundation (DFG) adopted the "Guidelines on the Handling of Research Data", in which research data is defined as follows: "Research data might include measurement data, laboratory values, audiovisual information, texts, survey data, objects from collections, or samples that were created, developed or evaluated during scientific work. Methodical forms of testing such as questionnaires, software and simulations may also produce important results for scientific research and

¹⁸ Kindling, Maxi and Peter Schirmbacher. "Die digitale Forschungswelt' als Gegenstand der Forschung." *Information - Wissenschaft & Praxis* 64 (2013): pp. 127-136, p. 130.
<https://doi.org/10.1515/iwp-2013-0017>. (transl. M.B.)

should therefore also be categorised as research data.”¹⁹ The DARIAH project²⁰ has developed a research data definition for the humanities and cultural studies that includes: “all those sources/materials and results collected, written, described and/or evaluated in the context of a research and research question in the field of human and cultural sciences, and in machine-readable form for the purpose of archiving, citation and for further processing.”²¹ Depending on the subject area, research data may vary significantly. The characteristics of research data strongly depend on the context (conditions of generation, methods used, perspective). Since they can be very heterogeneous, a further subdivision is not reasonable, therefore one usually speaks merely of “research data”.

2. Research data lifecycle



The research data lifecycle presents the steps necessary to map the process of a research project with regard to the research data. In particular, according to the UK Data Archive²², the lifecycle includes the collection, processing, analysis, archiving, access and re-use of research data (see Figure 3).

Figure 3. Source: Research data lifecycle of the UK Data Archive (Accessed: June 2018). In the meantime, a newer version of the life cycle is presented there.

3. Research data workflow

Based on the research data lifecycle, a research data workflow describes in more detail the individual processing steps of research data depending on the selected software and the required infrastructures and services. A process-oriented perspective is adopted, which allows the data creators to depict the data transfer and conversion necessary between data processing and analysis, for example. Responsibilities (roles and stakeholders) are also defined in a workflow.

¹⁹ DFG Guidelines on the Handling of Research Data, https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_research_data.pdf, last accessed 26.08.2020.

²⁰ DARIAH-DE. „Digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften.“ <https://de.dariah.eu/en/web/guest/home>, last accessed 21.09.2020.

²¹ DARIAH-DE, “Forschungsdaten” (research data), <https://de.dariah.eu/en/forschungsdaten>, last accessed 26.08.2020.

²² UK Data Service. „Research data lifecycle.“ Accessed on 18.05.2018. <https://www.ukdataservice.ac.uk/manage-data/lifecycle> In the meantime, a newer version of the life cycle is presented there. In this concept, however, we stick to the version shown above for didactic reasons, as we believe it is better suited to inspire reflection and discussion.

4. Research data management

a) What is research data management?

Research data are among the most important resources of science. Accordingly, it is essential to deal with them systematically and responsibly. Within the framework of research data management, "our own work processes concerning the generation and handling of research data" are organized and continuously controlled "as efficiently and goal-oriented as possible".²³ Research data management thus complements research from the initial planning stage all the way to archiving, reuse or deletion of the data.

As part of research data management, researchers develop methods and guidelines which they apply to their research activities associated with research data. This results in a strategy for handling data. This strategy helps to manage, control and standardize the handling of data in the subsequent research process.

By documenting the strategy with the planned methods and guidelines, an initial data management plan is created. It includes technical, organizational, structural, legal and ethical aspects of data handling for the duration of a project. But even more far-reaching aspects, such as the sustainability of the data, can be considered right from the start.

b) Tasks of research data management

Research data management is included in all steps of the research process. The central tasks of research data management are²⁴:

- Planning the handling of research data at the beginning of a research project and, if necessary, presenting the planned measures in funding proposals.
- Defining a folder structure and file naming conventions
- Documentation of research data and tagging with metadata
- Backup and long-term archiving of research data
- IT security and access control for research data
- Long-term archiving of research data
- Publication of research data
- Discovery and reuse of existing research data
- Consideration of data protection and copyright law when handling research data

c) The FAIR principles

In 2016, FORCE11 - a group of researchers as well as employees of libraries, archives, publishers and funders - established principles for the handling of research data.²⁵ The so-called FAIR principles comprise four goals²⁶: the **findability**, **accessibility**, **interoperability** and **re-usability** of data. With the achievement of these goals, the sustainable re-usability of research data is meant to be guaranteed. The data and its metadata should be findable by humans as well as by machines, with basic machine-readable descriptive metadata enabling the retrieval of relevant data sets. In addition, the data and its metadata should be archived in such a way that they can easily be accessed, downloaded or used

²³ Meyermann, Alexia: „Datenmanagement – Eine zentrale Voraussetzung für den Erfolg der Data Sharing-Idee in den Sozialwissenschaften.“ *DSZ-BO Working Paper Series* 2,7 (2012). Last Accessed 27.10.2020 http://www.uni-bielefeld.de/dsz-bo/pdf/2012-07-23_WP2_Datenmanagement.pdf. (transl. M.B.)

²⁴ Louise Corti, Veerle Van den Eynden, Libby Bishop and Matthew Woppard: *Managing and Sharing Research Data: A Guide to Good Practice*. Los Angeles, CA: SAGE, 2014.

²⁵ Wilkinson, Mark D., Michel Dumontier, IJsbrand J. Aalberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al.: „The FAIR Guiding Principles for scientific data management and stewardship.“ *Scientific Data* 3 (2016). <https://doi.org/10.1038/sdata.2016.18>.

²⁶ Go Fair Initiative. „FAIR Principles“ Last accessed 10.06.2020. <https://www.go-fair.org/fair-principles/>.

locally by humans and machines using standard communication protocols. The data should be available in a form in which it can be exchanged, interpreted and combined with other data sets by humans and machines in a (partially) automated manner. A good description of the data and its metadata is to ensure the re-use of the data for future research and the comparison with other compatible data sources. Proper citation of the data must be facilitated and the conditions for re-use need to be presented in a way that is comprehensible to humans and machines.

The main objective of the FAIR principles is to prepare the research data for humans and machines in the best possible way. This does not mean the data must be accessible without restrictions. Limited accessibility, for example due to the protection of data privacy, does not contradict the FAIR principles.

d) Motivation

A proper strategy for research data management simplifies working with the data during the research project as well as afterwards. It serves as a compass for all participants managing research processes and their results. In the planning phase of the research it takes time to develop guidelines and methods. Ultimately, this effort will pay off on several levels. Retrieving the data and reviewing the processes is much easier if the analyses and results can be reproduced. The prospects of reusing the data will increase.

Most importantly, research data management makes research more comprehensible, more reproducible and facilitates the validation of results in terms of good scientific practice. For researchers, this can contribute to additional scientific recognition and reputation.²⁷ Research funding agencies and publishers are placing increasing relevance on professional research data management for researchers. They demand a systematic and well-planned handling of generated data during the project period as well as access to research data after completion of the research project, i.e. the proactive management of research data.

Research data management facilitates the following aspects:

- better findability of data, e.g. by meaningful naming
- clarity, e.g. no scattered storage of data in different versions on different computers
- knowledge preservation – data are accessible independently of individual people, projects or institutions
- transfer of data to future projects
- facilitation of cooperation
- long-term reproducibility of results instead of new generation of data (retention of primary and secondary data)
- prevents data loss, e.g. due to defective hardware or software or to the loss of initial versions of files
- (semi-)automatic processing enabled by metadata
- transfer and reuse of data by using informed consents formulated accordingly, e.g. without stating, that data will be deleted after the project has expired
- optimised use of resources, e.g. cost savings through reuse instead of new data collection
- compliance with requirements imposed by funding agencies
- citation of research data
- referencing²⁸
- increasing the relevance of one's own work by improved visibility

²⁷ Piwowar, Heather, Roger S. Day and Douglas B. Fridsma: Sharing detailed research data is associated with increased citation rate. *PLoS one*, 2,3, e308 (2007).

<https://doi.org/10.1371/journal.pone.0000308>.

²⁸ This refers to persistent identifiers in particular, more about this in unit 15 point 8, p. 113.

Didactic methods and exercises

Methods:

- Question ball
 - What research data do you work with?
- Flip and turn
 - Research data lifecycle
- Truth or lie

Which of the following scenarios comply with the FAIR criteria?

- Within a project, extensive Excel tables with different data have been created. For the purpose of long-term archiving, these tables are published as PDF files in a research data repository.
- The data are provided with an individually specified license.
- After completion of the project, the data is published in a repository with a free and open protocol, e.g. http.
- If different versions of a data set exist, links are provided to the other versions.

Training materials

- Template: Cards for “Flip and turn” for the research data lifecycle
- Poster: How FAIR are your research data?
- Teaching script: Digital Research Data

Additional sources

- Aust, Pamela, Kerstin Helbig, Ulrike Schenk, Dennis Zielke, Anja Rosenbaum and Jörg Schulze: *Was sind Forschungsdaten?* Video. Berlin: Humboldt-Universität zu Berlin – Medien-Repositorium, 2016. <https://doi.org/10.18450/dataman/90>.
- Bertelmann, Roland, Petra Gebauer, Tim Hasler, Ingo Kirchner, Wolfgang Peters-Kottig, Matthias Razum, Astrid Recker, Damian Ulbricht and Stephan van Gasselt: *Einstieg ins Forschungsdatenmanagement in den Geowissenschaften*. Potsdam, GFZ – GFZpublic, 2014. <http://doi.org/10.2312/lis.14.01>.
- Corti, Louise, Veerle Van den Eynden, Libby Bishop and Matthew Woppard: *Managing and Sharing Research Data: A Guide to Good Practice*. Los Angeles, CA: SAGE, 2014.
- Dumontier, Michel: *Are we FAIR yet? Webinar*. RDA, 2018. Last accessed 29.07.2020, <https://rd-alliance.org/webinar-are-we-fair-yet>.
- Helbig, Kerstin, Katja Krause, Florian Rehak, Carolin Kruse, Anja Rosenbaum and Gianpiero Tari: *Forschungsdaten in der Chemie*. Video. Berlin: Humboldt-Universität zu Berlin - Medien-Repositorium, 2018. <https://doi.org/10.18450/dataman/96>.
- Helbig, Kerstin and Pamela Aust: *Forschungsdatenmanagement für Agrarwissenschaftler und Biologen*. Präsentation. Zenodo, 2016. <http://doi.org/10.5281/zenodo.53196>
- IANUS - Forschungsdatenzentrum für Archäologie & Altertumswissenschaften: „IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften“. Version: 1.0.1.0. Berlin: IANUS - Forschungsdatenzentrum für Archäologie & Altertumswissenschaften, 2017, <http://doi.org/10.13149/000.111000-a>.
- Kindling, Maxi und Peter Schirmbacher: "Die digitale Forschungswelt" als Gegenstand der Forschung. *Information - Wissenschaft & Praxis* 64 (2013), p. 130. <https://doi.org/10.1515/iwp-2013-0017>.
- Meyermann, Alexia: „Datenmanagement – Eine zentrale Voraussetzung für den Erfolg der Data Sharing-Idee in den Sozialwissenschaften.“ *DSZ-BO Working Paper Series* Nr. 2 (2012), p. 7.

Last accessed 28.07.2020. http://www.uni-bielefeld.de/dsz-bo/pdf/2012-07-23_WP2_Datenmanagement.pdf.

- PARTHENOS. "Manage, improve and open up your research and data." Last accessed 12.02.2019. <https://training.parthenos-project.eu/sample-page/manage-improve-and-open-up-your-research-and-data/>.
- Piwowar, Heather A., Roger S. Day and Douglas B. Fridsma: Sharing detailed research data is associated with increased citation rate. *PLoS ONE* 2 (2007), p. e308. <https://doi.org/10.1371/journal.pone.0000308>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand J. Aalberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al.: "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3 (2016). <https://doi.org/10.1038/sdata.2016.18>.
- ZBW, GESIS, RatSWD: *Auffinden, Zitieren, Dokumentieren: Forschungsdaten in den Sozial- und Wirtschaftswissenschaften*. Version 2.0. Hamburg: ZBW, GESIS, RatSWD, 2015. <https://doi.org/10.4232/10.fisuzida2015.2>.

Template: Cards for “Flip and turn” for the research data lifecycle

creating
data

processing
data

analysing
data

preserving
data

giving
access to
data

re-using
data

How FAIR are your data?

Findable

It should be possible for others to discover your data. Rich metadata should be available online in a searchable resource, and the data should be assigned a persistent identifier.

A persistent identifier is assigned to your data

There are rich metadata, describing your data

The metadata are online in a searchable resource e.g. a catalogue or data repository

The metadata record specifies the persistent identifier

Accessible

It should be possible for humans and machines to gain access to your data, under specific conditions or restrictions where appropriate. FAIR does not mean that data need to be open! There should be metadata, even if the data aren't accessible.

Following the persistent ID will take you to the data or associated metadata

The protocol by which data can be retrieved follows recognised standards e.g. http

The access procedure includes authentication and authorisation steps, if necessary

Metadata are accessible, wherever possible, even if the data aren't

Interoperable

Data and metadata should conform to recognised formats and standards to allow them to be combined and exchanged.

Data is provided in commonly understood and preferably open formats

The metadata provided follows relevant standards

Controlled vocabularies, keywords, thesauri or ontologies are used where possible

Qualified references and links are provided to other related data

Reusable

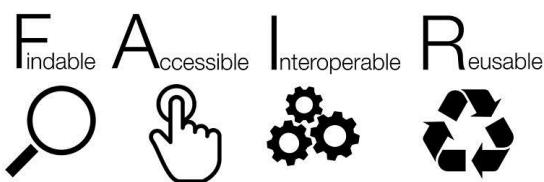
Lots of documentation is needed to support data interpretation and reuse. The data should conform to community norms and be clearly licensed so others know what kinds of reuse are permitted.

The data are accurate and well described with many relevant attributes

The data have a clear and accessible data usage license

It is clear how, why and by whom the data have been created and processed

The data and metadata meet relevant domain standards



'How FAIR are your data?' checklist, CC-BY by Sarah Jones & Marjan Grootveld, [EUDAT](#) Image CC-BY-SA by [SangyaPundir](#)

4. Teaching script: Digital research data									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternative	Notes
Data	The participants orient themselves in the topic	1	Introduction to the topic and explanation of the definition of data	Presentation	PPTX	In	-		
	The participants activate their previous knowledge	4	The participants answer the question: "What kind of research data do you work with?"	Method: Question ball	Small, soft ball	Ex	Yes	Brainstorming	
Research data	The participants understand the term research data.	1	Explanation of the term digital research data	Presentation	PPTX	In	-		
	The participants develop a research data life cycle	6	The participants are divided into groups and receive moderation cards with the components of the research data life cycle and should arrange the cards in a meaningful order. Additions are possible	Method: Flip and turn	Prepared moderation cards, blank cards for additions, pin board, pins, crepe tape	Ex	Yes	S: Participants work alone, receive an envelope with paper strips with the components of the research data life cycle to arrange 2 groups	
Research data life cycle	The participants thus anchor the newly acquired knowledge	8	The participants present their results and justify the chosen order	Group work	Pin board, pins	Ex	Yes		

	The participants learn what research data management is	4	The concept of research data management is explained	Presentation	PPTX	In	-	
	The participants work out the benefits of research data management	7	The participants answer the question: What are the benefits of RDM?	Method: Chatter	Moderation cards, pens	Ex	Yes	Discussion in groups of 2 from different areas of expertise, so that the differences are as big as possible and the exchange is livelier
	The participants exchange arguments	5	Participants present their arguments to the group	Group work	Pins, pin board	Ex	Yes	
	The participants learn about the FAIR principles	5	The principles: Data should be Findable, Accessible, Interoperable and Reusable	Presentation	PPTX	In	-	If experts in the group: let these participants explain
	The participants know the main aspects of RDM	2	The aspects of RDM are explained	Presentation	PPTX	In	-	

Duration of the unit: 43 minutes



Unit 5: Research data policies

Learning objectives

The participants are familiar with different research data policies.

The participants will learn for which subject areas guidelines already exist and where they can obtain information on them.

Key aspects

1. What is a research data policy?
2. Journal and publisher policies
3. Institutional policies
4. Discipline-specific requirements
5. Policies of research funding agencies

Contents

1. What is a research data policy?

A research data policy describes the guidelines for handling research data. There are different kinds of research data policies, for example:

- Journal and publisher policies
- Institutional policies
- Project-specific policies
- Discipline-specific policies
- Policies of research funding organisations

2. Journal and publisher policies

Since 2016 publishers such as Springer Nature,²⁹ Elsevier³⁰ and Wiley³¹ have set new guidelines for the handling of research data and apply them in their journals. They are guided by the Transparency and Openness Promotion (TOP) Guidelines³² published in 2015 by the Center for Open Science (COS). Publishers usually distinguish between three to four types of research data policies. The following are examples of the policy types and, in parentheses, examples of journals published at Springer Nature that use these types:

- The sharing and citation of data is encouraged ('Photosynthesis Research')
- The sharing of data and proof of data accessibility is encouraged ('Plant and Soil')
- Data sharing is encouraged and data availability statements are mandatory ('Palgrave Communications')
- It is a prerequisite to share data, to verify this and to allow peer review of the data ('Scientific Data')

Depending on the journal, it is therefore necessary to check carefully which of the publisher policies apply.

3. Institutional policies

Institutional research data policies are increasingly being adopted by universities and other educational institutions to regulate the management of research data and to clarify the basic legal aspects.³³ In addition to the regulation of open access of research data, they also assess and decide on the allocation of the research institution's personnel, organisational and technical capacities for RDM, i.e. cost and resource management.³⁴ In half of the German university RD policies, statements on the costs of RDM are considered important in relation to a data management plan.

Figure 4 shows the elements of an institutional research data policy and lists the content-related issues that are regulated in these categories³⁵. One example is the research data management policy of Humboldt-Universität zu Berlin³⁶, which was introduced in 2014 (see Training materials).

In addition to institutional policies, which often have an overarching character and are meant to represent the entire institution, project-specific policies can also be created. These define project-internal regulations and standards and can be adapted specifically to the respective project.

²⁹ Springer Nature. „Research Data Policies.“ Last accessed 16.01.2019, <https://www.springernature.com/gp/authors/research-data-policy/journal-policies/15369670>.

³⁰ Elsevier. „Sharing research data.“ Last accessed 03.11.2020, <https://www.elsevier.com/authors/author-services/research-data>.

³¹ Wiley. „Sharing and Citing your Research Data.“ Last accessed 03.11.2020, <https://authorservices.wiley.com/author-resources/Journal-Authors/licensing-open-access/open-access/data-sharing.html>.

³² Nosek, Brian A., George Alter, George C. Banks, Denny Borsboom, Sara D. Bowman, Steven J. Breckler, Stuart Buck et al.: „Promoting an open research culture.“ *Science* 348, 6242 (2015): pp. 1422–1425. <https://doi.org/10.1126/science.aab2374>.

³³ List of research data policies by German universities and colleges: forschungsdaten.org: „Data Policies“. Last accessed 28.07.2020.

³⁴ Hiemenz, Bea M. and Monika Kuberek: „Leitlinie? Grundsätze? Policy? Richtlinie? – Forschungsdaten-Policies an deutschen Universitäten.“ *o-bib* 5,2 (2018): pp. 1-13. <https://doi.org/10.5282/o-bib/2018H2S1-13>.

³⁵ Adapted from Hiemenz, Bea and Monika Kuberek: *Empfehlungen zur Erstellung institutioneller Forschungsdaten-Policies. Das Forschungsdaten-Policy-Kit als generischer Baukasten mit Leitfragen und Textbausteinen für Hochschulen in Deutschland*. Berlin: Technische Universität - DepositOnce, 2018. <https://depositonce.tu-berlin.de/handle/11303/8372>.

³⁶ Humboldt-Universität zu Berlin. „Humboldt-Universität zu Berlin Research Data Management Policy.“ Berlin, 2014. Last accessed 03.11.2020. <https://hu.berlin/RDM-policy>.

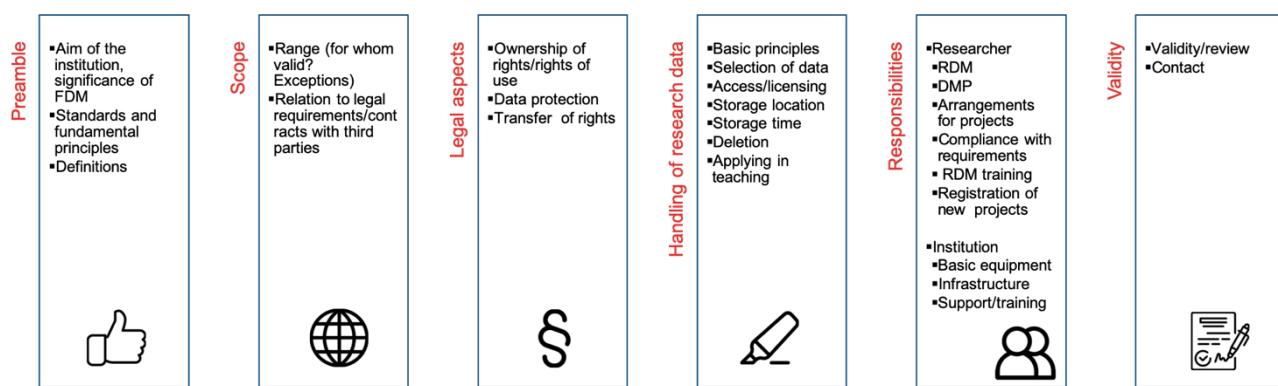


Figure 4: Categories and contents of institutional research data policies. Source: Hiemenz, Bea and Monika Kuberek: Empfehlungen zur Erstellung institutioneller Forschungsdaten-Policies. Das Forschungsdaten-Policy-Kit als generischer Baukasten mit Leitfragen und Textbausteinen für Hochschulen in Deutschland. Berlin: Technische Universität - DepositOnce, 2018. <http://dx.doi.org/10.14279/depositonce-7521>. Image source: Freepik: <https://www.flaticon.com/>.

4. Discipline-specific requirements

For some subject areas specific guidelines for the handling of research data already exist (e.g. psychology, genetics, biodiversity, linguistics, educational research, social and economic sciences). In the social sciences, for example, there is an agreement on the cooperation of European data archives, which was drawn up by CESSDA³⁷ (Consortium of European Social Science Data Archives). In the life sciences, the policy on "Good Clinical Practice (GCP)" and the "Principles of Good Laboratory Practice (GLP)" determine how the data are managed. Both principles are anchored in German legal code³⁸.

Such discipline-specific requirements are necessary since research data are very heterogeneous and the way they are handled can vary greatly. At the same time, the definition of these standards also serves to establish comparability and interoperability between disciplines. Thus, the discipline-specific data formats (e.g. in archaeology: 3D data) and the different ways of dealing with research data and their indexing (e.g. in the humanities: data can change permanently in the context of their indexing) need to be considered.

5. Policies of research funding agencies

Research funders are increasingly creating their own guidelines for handling research data. Examples of this are the European Commission³⁹ and the German Research Foundation⁴⁰. Applicants are requested to provide information on RDM, to draw up a data management plan (see in more detail in unit 6, p. 53) or to make research data generated in the project available under open licenses. Since

³⁷ CESSDA. Last accessed 03.11.2020, <https://www.cessda.eu>.

³⁸ BMJV. „Verordnung über die Anwendung der Guten Klinischen Praxis bei der Durchführung von klinischen Prüfungen mit Arzneimitteln zur Anwendung an Menschen.“ Last accessed 03.11.2020, <https://www.gesetze-im-internet.de/gcp-v/index.html>.

³⁹ European Commission, H2020 Programme. „Guidelines on FAIR Data Management in Horizon 2020.“ Version 3.0. Last accessed 07.10.2020, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

⁴⁰ Deutsche Forschungsgemeinschaft. „Leitlinien zum Umgang mit Forschungsdaten.“ Last accessed 28.07.2020, http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf.

the existence and content of the data management plan can be included in the peer review, it is advisable for applicants to follow the guidelines closely and provide information as precisely as possible. Potential sanctions for non-compliance with the policy or the data management plan may include withholding the project budget or reduced chances for a follow-up grant proposal. However, audits after the end of the project are currently not being carried out across the board.

Didactic methods and exercises

Methods:

- Chatter
 - Does a research data policy exist at your institution?
 - What is the scope of the research data policy at your institution?
 - Does the research data policy cover everything you would expect?
 - Alternatively, if there is no policy at your institution, would you prefer to have one and what should it contain?

Training materials

- Example of an institutional research data policy: Humboldt-Universität zu Berlin RDM Policy⁴¹
- Teaching script: Research data policies

Additional sources

- Forschungsdaten.org: „Data Policies“. Last accessed 28.07.2020, https://www.forschungsdaten.org/index.php/Data_Policies.
- Hiemenz, Bea and Monika Kuberek: *Empfehlungen zur Erstellung institutioneller Forschungsdaten-Policies. Das Forschungsdaten-Policy-Kit als generischer Baukasten mit Leitfragen und Textbausteinen für Hochschulen in Deutschland*. Berlin: Technische Universität - DepositOnce, 2018. <http://dx.doi.org/10.14279/depositonce-7521>
- Nosek, Brian A., George Alter, George C. Banks, Denny Borsboom, Sara D. Bowman, Steven J. Breckler, Stuart Buck, et al.: Promoting an open research culture. *Science* 348 (2015). <https://doi.org/10.1126/science.aab2374>.

⁴¹ Humboldt-Universität zu Berlin. „Humboldt-Universität zu Berlin Research Data Management Policy.“ Berlin, 2014. Last accessed 28.09.2020. <https://hu.berlin/RDM-policy>.



Humboldt-Universität zu Berlin Research Data Management Policy

Preamble

Handling of research data responsibly is critical for confirming research results, for the advancement of science and for the dissemination of knowledge. This policy is addressed to all researchers at Humboldt-Universität zu Berlin both in their role as independent researchers, and as teachers and supervisors. Their obligation includes instructing students and doctoral candidates about handling of research data properly, as well as imparting subject-specific skills and standards.

What is research data?

Research data means all data that was created in the course of research or results from it. Since the generation or collection of research data depends on a variety of research questions and methods, followed by processing, analysis and publishing and/or archiving, the data may involve a range of media, aggregation levels and formats. Sharing and reusing research data requires documentation about the original context and about tools that were used.

Policy

To further the [Proposals for Safeguarding Good Scientific Practice of the German Research Foundation](#), and the Rules of the Humboldt-Universität zu Berlin for Safeguarding Good Scientific Practice and Investigating Allegations of Scholarly Misconduct ([Satzung der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens](#), in German) the University adopts following policy:

1. HU researchers are committed to the secure storage, processing, documentation and preservation of their research data for the long term. Responsibility for ensuring these processes lies with HU staff in charge of research projects or programmes.
2. All HU researchers are encouraged to process research data resulting from their research activities according to the conventions and standards of their respective scientific community. They should document the complete research lifecycle including tools and procedures that they used.
3. HU researchers should take responsibility for deciding at what time and on what legal terms research data may be accessed. The Open Access Declaration of the HU ([Open-Access-Erklärung der Humboldt-Universität zu Berlin](#), in German) recommends making research data as well as scholarly publications publicly available in a timely manner. Protection of personal data, copyrights and the legitimate interests of third parties must also be ensured.
4. Research data underlying scholarly publications should be archived for the long-term and/or published in an appropriate trustworthy data archive or repository. They count as scholarly research output at Humboldt-Universität zu Berlin.

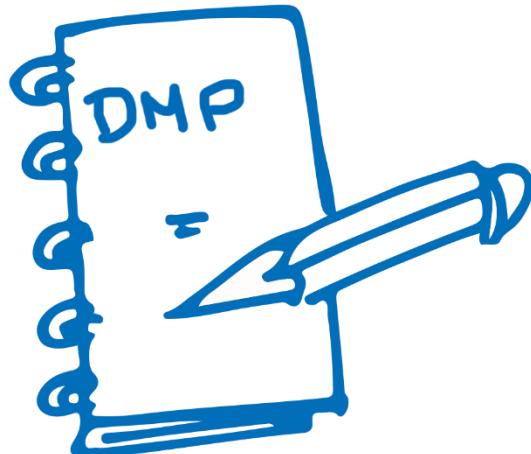
Humboldt-Universität zu Berlin has committed itself to establishing the means to enable this policy to be fulfilled.

This research data management policy was approved by the Academic Senate on 8th of July 2014.

Please note that this translation is made available in addition to the original version of the “Grundsätze zum Umgang mit Forschungsdaten an der Humboldt-Universität zu Berlin”. Only the German version is authoritative.

5. Teaching script: Research data policies						
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation
					Voices sound	Alternative Notes
Research data policies	Participants learn about the different policy types	4	Introduction to the topic. The different policy types are presented	Presentation	PPTX	In -
Institutional policies	Participants get to know the requirements for handling research data	5	Example of an institutional research data policy (Humboldt-Universität zu Berlin)	Presentation	PPTX; Printout of the HU Research Data Policy	In -
		5	The participants discuss whether their institution has a research data policy, what's the scope of it, where it is located and whether it covers everything they would expect	Method: Chatter	Ex Yes	

Duration of the unit: 14 minutes



Unit 6: Data management plan

Learning objectives

The participants know the components of a data management plan.

Participants will learn about the requirements of the funders regarding data management plans.

Participants get to know tools for creating data management plans.

Key aspects

1. What is a data management plan?
2. Motivation
3. Components of a data management plan
4. Funder requirements
5. Tools and templates

Contents

1. What is a data management plan?

A data management plan (DMP) is a document that describes the intended management of research data. This includes activities during the research process as well as after completion of the project. The plan contains all information necessary to describe and document the collection, processing, storage, archiving and publication of research data. The scope of a DMP can vary from a few paragraphs to several pages.

2. Motivation

A data management plan consumes resources in its creation and offers many advantages at the same time.

A data management plan:

- creates a firm basis for a uniform handling of data in the research process
- facilitates the understanding of your own data
- facilitates the coordination between project partners
- helps to identify potential problems at an early stage and to outline solutions for them
- defines responsibilities

- regulates access rights
- helps to avoid data duplication, data loss and security loopholes
- can be (sometimes obligatory) part of a grant application

A useful initial consideration and argument for a data management plan is to think backwards, i.e. where and how should the data be archived or published? These considerations make it necessary to set the course in the data management workflow early on, e.g. with regard to formats, standards, metadata, licenses, etc.

3. Components of a data management plan

Depending on the size of the project and the variety of data, data management plans vary greatly. It is important to consider recommendations and specifications of third parties, e.g. funding agencies or employers, when preparing them. The most frequently used components of DMPs are:

- Project title, project duration and research question(s)
- Name of person responsible for the data management
- Reuse of existing data
- Data to be collected:
 - Description of the data to be collected, data types and formats
 - expected memory capacity requirements
 - Data collection methods, hardware and software used
- Data organization:
 - Data storage
 - Backup
 - Folder structure
 - File naming conventions
 - Documentation and metadata
- Legal aspects, for example:
 - Data protection
 - Copyright
- Data exchange and access:
 - In the project
 - with external partners and service providers
- Long-term preservation and archiving
- Data Publication
- Costs of data management
- Quality assurance
- Access and re-use

The variety of research data and of the way they are handled determines the length of a data management plan. It should be short, specific and approved by all project participants. An incomplete DMP is better than none at all. Changes to the plan are not unusual and updates therefore necessary. Ideally, a data management plan evolves in a dynamic way: i.e. it is continuously updated and expanded during the course of the project and thus advances from a sketch to a detailed documentation of the

data management process (active Data Management Plan⁴²) and thus contributes to the reusability of the data.

4. Funder requirements

In Germany, data management plans are already required by many research funders when submitting proposals. Research funders such as the European Commission (EC), the German Research Foundation (DFG), the Volkswagen Stiftung (VW Foundation) and the Federal Ministry of Education and Research (BMBF) are increasingly demanding: a data management plan to be provided at the start of funding (EC),⁴³ information on how to handle collected research data (DFG)⁴⁴ and - depending on the funding guideline - a plan for the use of project results or a detailed data management plan with the funding proposal (BMBF⁴⁵, VW Foundation⁴⁶) (see Table 1).

5. Tools and templates

There are many ways to support researchers in creating a data management plan. Templates, support, examples and online tools are available. Some examples:

- Checklists:

https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdbinfo_2.pdf

www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf

http://opus.bath.ac.uk/36009/4/DMP_Guidance_for_PGRs_v1.3.pdf

http://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf

Table 1: Comparison of the requirements of the sponsors regarding the data management plan (status: 28.09.2020).

Funder	Plan demanded?	Submission on application?	Content	Updates?
European Commission Horizon 2020	Data management plan	No, first plan within the first 6 months of the project	Contents of the Horizon 2020 Template	Update, if significant changes occur and at the end of the project
German Research Foundation (DFG)	Information on the handling of research data	Yes	Contents of the DFG Guidelines on the Handling of Research Data	No

⁴² RDA Interest Group Active Data Management Plans. Last accessed 12.06.2020. <https://www.rd-alliance.org/node/9136/outputs>.

⁴³ European Commission, H2020 Programme. "Guidelines on FAIR Data Management in Horizon 2020." Version 3.0. Last accessed 03.11.2020, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

⁴⁴ Deutsche Forschungsgemeinschaft. „DFG Guidelines on the Handling of Research Data.“ DFG, 2015. Last accessed 03.11.2020.

https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_research_data.pdf.

⁴⁵ Verband Forschungsdaten Bildung. „Checklist zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung.“ Version 1.1. *fdbinfo* Nr. 2 (2015). Last accessed 26.02.2019. https://www.forschungsdaten-bildung.de/files/fdbinfo_2.pdf

⁴⁶ VolkswagenStiftung. „Information Open Access – Open Data – Open Source.“ Last accessed 09.07.2020,

https://www.volkswarenstiftung.de/sites/default/files/downloads/OpenAccessOpenDataOpenSource-Hinweise_10_2018.pdf.

German Federal Ministry of Education and Research (BMBF)	Plan sometimes required, depending on the programme	If required, yes	Content depends on the respective programme; Educational research: Checklist (in German)	Depends on the programme
VolkswagenStiftung (VW Foundation)	Data management plan	Yes	Contents of the Science Europe Template	No

- Templates (see attachment):

<https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-h2020-v3>

<https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-dfg>

<https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-bmbf>

<https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-vwstiftung-pdf>

<http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/dmp/framework.html>

- Tools:

DMPTool (<https://dmptool.org/>): US-American tool for creating a DMP. Target group are researchers who want to submit an application to American research funding agencies.

RDMO (<http://rdmorganiser.github.io>): German tool developed in a DFG project. The target group are researchers who want to better organize their RDM. The tool is intended to accompany the research process. Not primarily designed for the creation of a DMP, but a creation is possible with the tool and especially suitable for German research funding agencies. There are also discipline-specific templates (so-called question catalogs) available. On forschungsdaten.info RDMO can be used without an installation via Shibboleth.

DMPonline (<https://dmponline.dcc.ac.uk/>): A British tool for creating a DMP. The target group is researchers who would like to submit an application to British research funding agencies or the European Commission. DMPTool and DMPonline are now based on the same software code (DMPRoadmap) and cooperate in further development.

Data Stewardship Wizard (<https://ds-wizard.org/>): A Czech tool developed by Elixir, TU Prague and other partners. The target group of this tool are data stewards who support researchers in creating a DMP.

OpenDMP/Argos (<https://argos.openaire.eu>): A Greek tool developed in the context of a project of OpenAire and EUDAT. Target group are researchers who want to create a DMP. The tool supports the RDA maDMP template allowing the re-use of machine-readable information.

- Discipline-Specific Tools:

GFBio Data Management Plan Tool (<https://www.gfbio.org/de/plan>): A tool for creating a DMP. Target group are researchers from biology and related disciplines.

DataWiz (<https://datawiz.leibniz-psychology.org/DataWiz/>): The tool supports data management in psychological projects. Target group are researchers from the field of psychology.

Wizzard by CLARIN-D (<https://www.clarin-d.net/de/aufbereiten/datenmanagementplan-entwickeln>): A short questionnaire for researchers from the humanities.

- Video tutorials:

<https://doi.org/10.18450/dataman/91>

<https://www.youtube.com/watch?v=ukVHHKp6sck&feature=c4-overview&list=UULTOHF6qQrYhEvQzbu03tTg>

Didactic methods and exercises

Methods:

- Ideas out loud
 - What benefits does a data management plan provide?
 - What assistance do you have for creating data management plans?
- Chatter
 - What are your comments on the data management plan?

Exercises:

- Creating a data management plan using a (fictitious) example in a DMP tool

Training materials

- Example: Data management plan
- Teaching script: Data management plan

Additional sources

- CESSDA ERIC: „Adapt your Data Management Plan.” Last accessed 09.07.2020. https://www.cessda.eu/content/download/4302/48656/file/TTT_DO_DMPExpertGuide_v1.3.pdf.
- Helbig, Kerstin, Katja Krause, Carolin Kruse, Florian Rehak and Gianpiero Tari: *Was sind Datenmanagementpläne?* Video. Berlin: Humboldt-Universität zu Berlin - Medien-Repositorium, 2017. <https://doi.org/10.18450/dataman/91>.
- Helbig, Kerstin and Pamela Aust: *Datenmanagementpläne für EU, DFG und BMBF. Präsentation.* Zenodo, 2015. <https://doi.org/10.5281/zenodo.33482>.
- Helbig, Kerstin: *Hinweise und Checklist zur Erstellung eines Datenmanagementplans.* Zenodo, 2015. <https://doi.org/10.5281/zenodo.33485>.
- Minn, Gisela and Marina Lemaire: "Forschungsdatenmanagement in den Geisteswissenschaften. Eine Planungshilfe für die Erarbeitung eines digitalen Forschungskonzepts und die Erstellung eines Datenmanagementplans." Universität Trier eSciences Working Papers 3. Trier: Universität Trier - Opus, 2017. <https://ubt.opus.hbz-nrw.de/frontdoor/index/index/docId/799>. Last accessed 26.02.2019.
- Neuroth, Heike, Claudia Engelhardt, Jochen Klar, Jens Ludwig and Harry Enke: Aktives Forschungsdatenmanagement. *ABI Technik* 38 (2018), pp. 55–64. <https://doi.org/10.1515/abitech-2018-0008>.
- Verband Forschungsdaten Bildung. „Checklist zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung.“ Version 1.1. *fdbinfo* Nr. 2 (2015). Last accessed 26.02.2019. https://www.forschungsdaten-bildung.de/files/fdbinfo_2.pdf.

POLARSTERN cruise ANT-X/6

Plan ID ANT-X/6

Principal Investigator / Researcher Kerstin Helbig

Plan Data Contact kerstin.helbig@hu-berlin.de

Plan Description Ice and animal observations during POLARSTERN cruise ANT-X/6.

Your ORCID 0000-0002-2775-6751

Data Collection

What data will you collect or create?

The numeric data will be saved as TAB-delimited TXT files (ASCII) as well as ZIP-archives and XLSX files. The chosen formats and software are suitable for longterm archiving. The area has not yet been researched, therefore no data for reuse exists.

How will the data be collected or created?

The data will be collected via underway cruise track measurements (CT). The files will be structured and named via cruise and project name. Version control - if necessary - will be realised with three digits (1.0.0).

Documentation and Metadata

What documentation and metadata will accompany the data?

The classification used will be discipline-specific: Aquatic Sciences and Fisheries and Oceanic Abstracts Classification Codes

- 1362 Ornithology – Geographical distribution
- 1372 Mammalogy – Geographical distribution
- 2144 Regional studies, expeditions and data reports
- 2150 Ice

The applied metadata standard will be repository-specific (according to Pangaea). In addition Darwin Core and ISO 19115 as discipline-specific standards will be used.

The documentation and metadata will be created during the data collection by the data curator of the project.

Ethics and Legal Compliance

How will you manage any ethical issues?

Does not apply.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Does not apply.

Storage and Backup

How will the data be stored and backed up during the research?

The data will be stored locally on the ship as well as directly backed up in the repository.

How will you manage access and security?

The data will be immediately saved and made available in Pangaea data repository. The raw data as well as the processed data will have no embargo period and are not sensitive.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

All data generated during the cruise is of long-term value and should be preserved. The data is shared with the public.

What is the long-term preservation plan for the dataset?

The data will be long-term preserved in the discipline-specific repository Pangaea. The fee for preservation will be 300,00 Euro for each data upload.

Data Sharing

How will you share the data?

The data will be made available immediately after collection via Pangaea. The data will get a Digital Object Identifier (DOI) through data publication in Pangaea.

Are any restrictions on data sharing required?

The data will not be used exclusively. There will be no restrictions. A data sharing agreement will not be required.

Responsibilities and Resources

Who will be responsible for data management?

The data curator in close cooperation with the principal investigators will be responsible for implementing the DMP and data management. Data ownership will be contractually clarified.

What resources will you require to deliver your plan?

The project needs a data curator for data description, metadata creation and upload within Pangaea. The data repository charges 300,00 Euro for the data upload. No hardware or software is required.

6. Teaching script: Data management plan									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternative	Notes
Definition and motivation	The participants get to know the term data management plan and learn a use case	2	Explanation of the term RDM plan	Presentation	PPTX	In	-	-	
Requirements of the research funders	The participants learn about the requirements of external research funders.	3	Tabular comparison of the most important funders	Presentation	PPTX	In	-	-	
Components of an RDM plan	The participants learn about the most frequently used components of RDM plans	2	The scope and components of an RDM plan are presented and discussed	Presentation	PPTX	In	-	Match single aspects of an RDM plan to the components	S: If most participants have not yet dealt with this: in the plenum
RDM plan tools	The participants activate their previous knowledge on support tools	5	The participants describe which support contacts or tools they know	Method: Gotcha and note results on flip chart	Flip chart	Ex	Yes	-	
	The participants get to know different tools and assistance and become familiar with templates.	10	Naming of the different tools and example presentation of a tool (RDMO, DMPonline, DMPTool)	Presentation and Live-Demo	PPTX, Internet access	In	-		

Duration of the unit: 22 minutes



Unit 7: Order and structure

Learning objectives

The participants understand the purpose of the structured approach and recognize the benefits of the initial preparation effort.

Participants know tools to structure and organise research data well.

Key aspects

1. Motivation
2. Folder structure
3. Naming
4. Rename multiple files at once
5. Version control

Contents

1. Motivation

You need a structured approach:

- to ensure that what, how and why things were done remains comprehensible even years later
- in order to make the naming conventions known to other researchers, but also to yourself, and to facilitate collaboration
- in order to allow other researchers to work with the data
- to be able to search for data more easily and to find it more quickly
- to avoid any duplication of work
- to prevent data loss due to overwriting or accidental deletion
- to identify the current status without effort
- to ensure machine readability

Overall, a structured approach will result in more efficient work.

2. Folder structure

A folder structure (also called directory tree) is the hierarchical arrangement in which folders are created. Hierarchical structures make it easier to find data (see Figure 5). The folder structure should be self-evident and thus also comprehensible for other researchers. The more carefully you plan it, the easier it will be to find your way around later. Ideally, folder structures follow the workflow in the respective project and thus support the step-by-step creation, analysis and publication of the data.

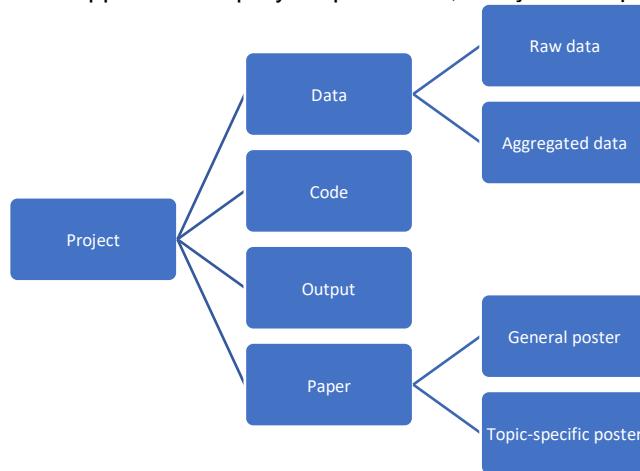


Figure 5: Example of a folder structure with subfolders.

For a clear overview, folder structures on servers should be identical to those on the local devices.

3. Naming

The file name should be objective and intuitive, and should be comprehensible to anyone. The naming and identification can be done according to the following three criteria:

- The system: For later access and retrieval of the data, it is crucial to consider the system under which the file is stored.
- The context: The file name contains content-specific or descriptive information, so that regardless of the storage location, it remains clear which context the file belongs to, e.g. "Schedule.pdf" or "ScheduleFDMentor.pdf".
- Consistency: The naming convention should be chosen in advance to ensure that it can be adhered to systematically and contains the same information (such as date and time) in the same order (e.g. YYYYMMDD).

File names should be as long as necessary and as short as possible to keep them concise and readable under any operating system. Name components that are already included in the folder names do not need to be repeated in the file names. To ensure uniform naming, the following name components can be used:

- Content
- Creator
- Date of creation
- Date of processing
- Name of the working group
- Date of publication
- Project number
- Version number

Spaces, periods and special characters (like { } [] < > () * % # ' ; " , : ? ! & @ \$ ~) should be avoided as they are differently interpreted under different systems and this may lead to errors. On most operating systems you can replace spaces with underscores or you can use capitalization of the first letters of words.

To ensure chronological order, it is recommended to start the name with a date, for example YYYYMMDD_Name or YYMMDDName etc.

Examples for uniform naming:

- 20160512_Climate Measurement1_original.jpg
- 20160522_Climate Measurement1_MHU_excerpt.jpg
- 20160523_Climate Measurement1_MHU_excerpt_edited_color.jpg

You should avoid automatically generated names (e.g. from digital cameras), as they may lead to conflicting names due to repetition. When deciding on a naming convention, scalability should be taken into account: e.g. when choosing a two-digit file number, the number of files will be limited to 99.

Whether it is a large project or a small research project, it is worthwhile to keep a written record of the chosen naming conventions. In particular, chosen abbreviations should be explained in a data management plan (see unit 6: Data Management Plan, p. 53) or a readme file. Without such a file the reconstruction of these conventions can be difficult after some years have passed.

4. Rename multiple files at once

Renaming multiple files at once is useful in many situations, such as

- to change the automatically generated names from your digital camera or other software in one step;
- to remove or replace spaces or other special characters from multiple file names in one operation.

Software for renaming multiple files at once exists for most operating systems.

Windows:

- Ant Renamer (<http://www.antp.be/software/renamer> " \t "_blank)
- Rename-IT (sourceforge.net/projects/renameit)
- Bulk Rename Utility (www.bulkrenameutility.co.uk/)

Mac:

- Renamer 5 (for Mac) (renamer.com/)
- Name Changer (mrrsoftware.com/namechanger/)

Linux:

- GNOME Commander (<https://gcmd.github.io/>)
- GPRename (<http://gprename.sourceforge.net/>)

Unix:

- On Unix the command "rename" can be helpful to find and rename files with regular expressions.

5. Version control

Versioning is used for different purposes. Versions and their history help, for example, to keep an overview of the steps performed and to make them traceable. They also allow you to easily go back one step. Versions made available to the public can also support debugging. The inclusion of new data and/or changes in a file structure – especially in the case of software as research data – may lead to new versions of the same file or even to new results.

The most common way of labelling versions is to assign integral numbers for major version changes and numbers subsequent to an underscore for minor changes (e.g. v1, v2, v1_01, v2_03 etc.). It is not recommended to use names like final, final2, revision, definitely_final.

Version control software (e.g. Git or Subversion) is very helpful in managing versions. For collaborative documents and storage locations such as Wiki, Google Docs or cloud, versioning and change tracking is available.

Examples of file labelling with version control:

- [document name][version number]
- Doe_interview_July2010_V1
- Lipid_analysis_rate_V2
- 2017_01_28_MR_CS3_V6_03

Didactic methods and exercises

Methods:

- Ideas out loud
 - Why do we need a structured organisation of data?

Exercises:

- Design a naming convention for your files and give some examples.
- Design a structure for your data storage as a directory tree.

Training materials

- Worksheet: Order and structure
- Checklist: Versioning
- Teaching script: Order and structure

Additional sources

- CESSDA Training Working Group: „CESSDA Data Management Expert Guide. Bergen: CESSDA ERIC, 2017-2018. Last accessed 28.07.2020. <https://www.cessda.eu/DMGuide>.
- Corti, Louise, Veerle Van den Eynden, Libby Bishop and Matthew Woppard: *Managing and Sharing Research Data: A Guide to Good Practice*. Los Angeles, CA: SAGE, 2014.
- Haenel, Valentin and Julius Plenz: *Git: verteilte Versionsverwaltung für Code und Dokumente*. 2nd ed. München: Open Source Press, 2014.
- Krejčí, Jindřich: *Introduction to the Management of Social Survey Data*. Prague: Institute of Sociology, 2014. Last accessed 26.02.2019. http://archiv.soc.cas.cz/sites/default/files/introdatamanagement_web.pdf
- Pilato, C.M., Ben Collins-Sussman and Brian W. Fitzpatrick: *Versionskontrolle mit Subversion*. 3rd ed. Köln: O'Reilly, 2009.
- Software Carpentry. *Data - Episode 1 - Data Management*. Video. Youtube, 2011. Last accessed 28.07.2020. <https://youtu.be/3MEJ38BO6Mo>.
- Software Carpentry. "Data Management: Data Management." Last accessed 28.07.2020, <https://v4.software-carpentry.org/data/mgmt.html>.
- Technology, Information and Learning Support (TILS). *TILS Document Naming Convention*. Queensland University of Technology, 2009. Last accessed 28.07.2020. https://www.data.cam.ac.uk/files/gdl_tilsdocnaming_v1_20090612.pdf.



Worksheet: Order and structure

Which of these examples follow a good naming convention?

Olga_170413_probe17k
Naturepaper karl britta james finished!
Vm4520132Schmidt.pdf
647749157.pdf
170413_sample17k_olga
Naturepaper+karl+britta+james &nal
Olga170413sample17k
Krst_765_spct_1203
Naturepaper+karl+britta+james finished! reworked
Cristal_765_spectr_20161203
Nature_karlbrittajames_endendversion
28q8QGIHKwrRw.pdf
Conference_Digital_Science.pdf

Please design a naming convention for your files and give some examples:

Please design a structure for your storage in form of a directory tree:



Checklist: Versioning

- The conditions for data use have been defined and communicated to team members and other users.
- A “master file” has been created and measures taken to maintain its authenticity, i.e. access rights and responsibilities are defined - who is entitled to make what kind of changes?
- Differences between shared versions of researchers and working versions of individuals were clarified.
- It was determined how many versions of a file should be kept, which versions should be kept (e.g. major versions instead of minor versions (version 2.0 but not 2.1)), how long and how versions should be organized.
- A clear and systematic naming of file versions and editions has been introduced.
- Relationships between elements, e.g. between code and the data file needed for execution, between data file and associated documentation or metadata, or between multiple files have been documented as needed.
- Changes in any version are documented.
- Original versions of files or documentation thereof are retained so that original files can be reconstructed.
- The files are synchronized regularly at different locations.

Sources:

- CESSDA Training Working Group. CESSDA Data Management Expert Guide. Bergen, Norway: CESSDA ERIC, 2017-2018, <https://www.cessda.eu/DMGuide>. This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).
- Krejčí, Jindřich. Introduction to the Management of Social Survey Data. Praha: Sociologický ústav AV ČR, v.v.i.64 s. 2014. ISBN 978-80-7330-252-8
- Corti, Louise, Veerle Van den Eynden, Libby Bishop und Matthew Woppard. Managing and Sharing Research Data: A Guide to Good Practice. Los Angeles, CA: SAGE, 2014.

7. Teaching script: Order and structure									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternatives	Notes
Folder structure	The participants recognize a neat folder structure	15	1. Participants write an example on a moderation card: "When have I ever missed the order and/or structure of my data? – 2 min 2. Participants stand up, walk and form pairs for 2 minutes: one person briefly reports their example, the other considers the most original solutions with order and structure. Ring to form new pairs and repeat task (5x).	Exercise	Moderation cards, pens, timer	Ex	Yes	S: Lecture in which the additional effort on good structure and order and its positive effects for the overall work are discussed	
Naming and renaming of files	The participants recognise a proper file naming	2	Hints for proper naming of files are given	Presentation	PPTX	In	-	Worksheet: Order and structure	
Version control	The participants get to know the tools for simultaneous file renaming.	2	Tools for the simultaneous renaming of multiple files are presented.	Presentation	PPTX	In	-		
	The participants learn about the possibilities of versioning.	1	Explanation of the importance of versioning and presentation on the possibilities of versioning and version control	Lecture and presentation of tool (Screenshot)	PPTX	In	-	Checklist: Versioning	

Duration of the unit: 20 minutes



Unit 8: Documentation and metadata

Learning objectives

The participants know the benefits and contents of data documentation.

Participants know the metadata standards relevant to them and understand why standards should be applied.

The participants are familiar with the terms: Metadata, controlled vocabulary and authority files.

Participants know where to find metadata standards.

Key aspects

1. Data documentation
2. What are metadata?
3. Metadata standards
4. Discipline-specific metadata standards
5. Thesaurus, authority files and controlled vocabulary
6. Electronic Lab Notebooks (ELN)

Contents

1. Data documentation

In order to make research data more findable and traceable, documentation of data is essential. It considerably facilitates the further use of the data and enables reproducibility. Well-documented data will be used and cited more often, which will increase the reputation of the creator. Documentation is also important with regard to the subsequent usability and traceability of the data for your own work. Over time, details might be forgotten, so it is recommended to document the data while working on it.

Basic contents of a documentation include:

- description of the research project

- project goals
- hypotheses
- detailed information on data collection (methods, units, time periods, locations, technology used)
- measures for data cleansing
- structure of data and their relationships to each other
- explanation of variables, labels and codes
- differences between versions
- information on access and terms of use

2. What are metadata?

Metadata refers to structured data that contains information on other data – "data about data". They are stored either independently of or in combination with the data they describe. There is a distinction between content-related and technical metadata. They form a specific subset of the documentation data and serve primarily to make the data findable, including in library reference systems. In order to make them machine-readable, for example in Semantic Web⁴⁷ applications, they are often stored in XML⁴⁸ format.

3. Metadata standards

Standardisation of metadata vocabulary is necessary to improve findability of the data and to provide interoperability. The linking of the metadata will ensure this. Furthermore, standards allow a uniform description of similar data sets in terms of content and structure.

Metadata standards contain a defined selection of information which is necessary to find and identify these data. A reusability of the data is not necessarily guaranteed thereby (compare section on documentation). Among the most common bibliographic interdisciplinary metadata standards are Dublin Core, DataCite Metadata Schema and MARC21.

4. Discipline-specific metadata standards

Since each scientific community has its own requirements, different discipline-specific metadata standards are also being developed. For example, in the social and economic sciences the Data Documentation Initiative (DDI) standard is frequently used, while in the natural sciences the ICAT scheme or the Crystallographic Information Framework is used.

An overview on discipline-specific metadata standards is available, for example, on the pages of the British Digital Curation Centre⁴⁹ and in an overview of the Research Data Alliance.⁵⁰

5. Thesaurus, authority files and controlled vocabulary

For persons, institutions, research funders, locations and much more, so-called authority files are assigned to enable a clear attribution. This facilitates, for example, the finding of persons in case of identical names and allows search engines to interpret this data unambiguously. Among the most important authority files are:

- GND (dnb.de/gnd): the Integrated Authority File is mainly used for cataloguing literature in libraries, but is also increasingly used for other purposes.

⁴⁷ Semantic Web facilitates the exchange of data between computers and serves as an extension of the Web with a machine-readable layer.

⁴⁸ XML stands for "Extensible Markup Language". It allows the representation of hierarchically structured data in the format of a text file.

⁴⁹ Digital Curation Centre: "Disciplinary Metadata." Last accessed 03.11.2020. <http://www.dcc.ac.uk/resources/metadata-standards>.

⁵⁰ Research Data Alliance: "Metadata." Last accessed 03.11.2020. <http://rd-alliance.github.io/metadata-directory/subjects/>.

- ISNI (isni.org): the International Standard Name Identifier serves to uniquely identify the public identity of persons involved in a publication. The ISNI is a standard of the International Organization for Standardization (ISO) and is comparable to the ORCID.
- VIAF (viaf.org): the Virtual International Authority File is an international authority file for personal data and is maintained by the Online Computer Library Center (OCLC). The GND and ISNI authority files are part of VIAF.
- The Open Funder Registry (<https://support.crossref.org/hc/en-us/articles/214360886-The-Open-Funder-Registry>) is used by research funding agencies for identification purposes.

Controlled vocabularies are also necessary to enable structured documentation of data. Thesauri and classifications are documentation languages used to describe the content of research data. Classifications are used to assign objects to (mostly hierarchically structured) classes. These classes are characterized by certain attributes. A thesaurus, on the other hand, is a natural-language, structured collection of terms and their relationships to one another.

Thesauri and controlled vocabulary significantly enhance metadata and increase the findability of the data. Many disciplines already provide their own specialised classifications and thesauri.

Examples of discipline-specific classifications:

- Environmental Classification, <https://sns.uba.de/umthes/de/collections/UK.html>
- Classification of Social Sciences, <http://www.gesis.org/unser-angebot/recherchieren/tools-zur-recherche/klassifikation-sozialwissenschaften/>
- Physics and Astronomy Classification Scheme (PACS), <https://journals.aps.org/PACS>
- Mathematics Subject Classification (MSC), <https://mathscinet.ams.org/msc/msc2010.html>

Examples of discipline-specific thesauri:

- Agricultural Sciences: AGROVOC Multilingual agricultural thesaurus, <http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus>
- Humanities: A Thesaurus of Old English, <http://oldenglishthesaurus.arts.gla.ac.uk/>
- Information Sciences: INFODATA Thesaurus, http://www.infodata-edepot.de/thesaurus/T_SM.HTM
- Art and Architecture: Art and Architecture Thesaurus (AAT), <http://www.aat-deutsch.de/>
- Life Sciences: Environmental Thesaurus (UMTHES), <https://sns.uba.de/umthes/de.html>
- Medicine and Life Sciences: Thesaurus Medical Subject Headings (MeSH), <https://www.nlm.nih.gov/mesh/>
- Psychology: Psychology Thesaurus (PSYNDEX), <http://www.zpid.de/index.php?wahl=products&uwahl=printed&uuwahl=psyndexterms>
- Economics: Standard Thesaurus for Economics (STW), <http://zbw.eu/stw/version/latest/about>

Discipline-specific classifications and thesauri can be searched via the Basel Register of Thesauri, Ontologies & Classifications (BARTOC).⁵¹

6. Electronic Lab Notebooks (ELN)

Electronic Lab Notebooks (ELN, or ELB for Electronic Lab Books) are designed to document the conception, execution and evaluation of scientific experiments, observations or other studies and the research data generated in this context. They are the digital versions of paper lab notebooks, which have so far been an essential part of the scientific work process in natural and life science disciplines. With increasing digitization, especially in the collection of data, ELN are also experiencing a growing acceptance and use. Meanwhile, there are a number of ELN software products available - from commercial providers to open source solutions. To name just a few examples:

- Chemotion (Open Source) <https://chemotion.net>
- eLabFTW (Open Source) <https://www.elabftw.net>

⁵¹ BARTOC.org. Last accessed 03.11.2020. <http://www.bartoc.org>.

- Labfolder (commercial) <https://www.labfolder.com>
- openBIS (Open Source) <https://openbis.ch>
- Rspace ELN (commercial) <https://www.researchspace.com>

In general, different disciplines have very different requirements regarding the features an ELN software should provide, so there is no "one-fits-all solution". In a manual prepared by the ZBMed⁵², practical advice is given that may help in the selection and implementation of an ELN.

Didactic methods and exercises

Methods:

- User manual
 - Write an instruction manual for your provided item.
- Ideas out loud
 - Why do you need data documentation?
 - Which metadata standards may be used in your field?
- LEGO® SERIOUS PLAY®
 - LEGO® metadata for reproducibility⁵³

Exercises:

- Create a list of metadata that (might) occur in your subject area. Exchange ideas with your right-hand neighbour.
- Search for discipline-specific metadata standards that might be relevant to you.

Training materials

- In practice: How to start with data documentation?
- Teaching script: Documentation

Additional sources

- CESSDA Training Working Group. „CESSDA Data Management Expert Guide. Bergen: CESSDA ERIC., 2017-2018. Last accessed 26.02.2019. <https://www.cessda.eu/DMGuide>
- Donaldson, Mary and Matt Mahon. *LEGO® Metadata for Reproducibility game pack*. University of Glasgow, 2019. <http://dx.doi.org/10.36399/gla.pubs.196477>
- ZB MED (Ed.). Elektronische Laborbücher im Kontext von Forschungsdatenmanagement und guter wissenschaftlicher Praxis – ein Wegweiser für die Lebenswissenschaften. Köln, 2019. <https://doi.org/10.4126/FRL01-006415715>.

⁵² ZB MED (Ed.). *Elektronische Laborbücher im Kontext von Forschungsdatenmanagement und guter wissenschaftlicher Praxis – ein Wegweiser für die Lebenswissenschaften*. Köln, 2019. <https://doi.org/10.4126/FRL01-006415715>.

⁵³ Donaldson, Mary and Matt Mahon. *LEGO® Metadata for Reproducibility game pack*. University of Glasgow, 2019. <http://dx.doi.org/10.36399/gla.pubs.196477>.



In practice: How to start with data documentation?

1. Do not panic. Much documentation is simply good research practice, so you are probably already doing much of it.
2. Start early! Careful planning of your documentation at the beginning of your project helps you to save time and effort. Do not wait till the end of the project to start the documentation. Remember to include procedures for documentation in your data management planning.
3. Think about the information that is needed in order to understand the data. What will other researchers and re-users be needing in order to understand your data?
4. Create a separate documentation file for the data that includes the basic information about the research data. You can also create similar files for each data set. Remember to organize your files so that there is a connection between the documentation file and the data sets.
5. Plan where to deposit the data after the completion of the project. The repository probably follows a specific metadata standard that you can adopt.
6. Document continuously throughout the project. Data documentation provides contextual information about your records. It records the objectives of the project and contains explanatory material, including the data source, methodology and data collection process, data set structure, and technical information. Comprehensive and structured information helps you to identify a data set and make decisions about its content and usability.

TIP: Use English language for the documentation. It increases the likelihood that your data will be understood and reused.

Source:

CESSDA Training Team. (2020). CESSDA Data Management Expert Guide. Bergen, Norway: CESSDA ERIC. <https://doi.org/10.5281/zenodo.3820472>. This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).

8. Teaching script: Documentation and metadata									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternative	Notes
Data documentation	The participants write a documentation	27	Task: Write an instruction manual for your machine; one person from the other group will test the description in the operating manual. Introduction – 2 min; Writing: – 15 min; Test & discussion – 10 min	Group work w/ two groups Method: Instruction manual	Scooter and skipping rope; flip chart, markers	Ex -	-	-	L: the advantages are derived from the answers and noted on a flip chart
	The participants become aware of the benefits of data documentation	2	The participants answer the question: When have you ever needed documentation?	Ideas out loud -	-	Ex	Yes	-	Fig.: Argumentation aid: How to start with data documentation
Metadata and metadata schemas	The participants get to know the contents of the data documentation	1	The basic contents of a documentation are presented	Presentation	PPTX	In	-	-	
Metadata standards	The participants get to know the term metadata	2	Explanation of the concept of metadata. Explanation of the differences between content and technical metadata	Presentation	PPTX	In	-	-	
	The participants understand the benefits of standards. They learn about the terms thesaurus, authority control and controlled vocabulary	5	Explanation of the terms thesaurus, authority control, controlled vocabulary	Presentation	PPTX	In	-	-	

	The participants deal with their discipline-specific metadata standards	The participants brainstorm and compile a list of which metadata (could) occur in their subject area, participants from central institutions can work on general metadata. Participants exchange information with their right-hand neighbours. Introduction: 1 min Brainstorming: 3 min Exchange: 6 min	Individual work	Paper, pencils	Ex	L: Exercise: Fill in a metadata schema using a photo example (3 minutes). Needed: Template and photo	
Subject-specific metadata standards	10	The participants learn where to search for cross-disciplinary and discipline-specific metadata standards	3	Examples of cross-disciplinary and discipline-specific standards are presented	PPT	In	L: Final exercise: Search for subject-specific metadata standards

Duration of the unit: 50 minutes



Unit 9: Storage and backup

Learning objectives

Participants are aware of the risks of careless handling of data.

The participants know the strategies for a secure backup.

The participants know the local backup services.

Key aspects

1. Data storage
2. Backup options
3. Strategies for a secure backup

Contents

1. Data storage

Research data can be stored on different media with different advantages and disadvantages. Depending on the storage device, there are serious differences in the protection against data loss and unauthorized access. The following is an overview of the features, advantages and risks of the most common storage devices and locations:

Personal Computer

- + self-responsible for security and backup; own control
- everything that happens with the computer happens with the backup
- possible lack of resources and know-how to configure and check the quality of backup copies
- individual solutions are complex, costly and inefficient in the overall picture of a working group

Mobile storage device

- + easy to transport
- + can be stored in a lockable cabinet or safe
- particularly easy to lose and can be easily stolen, therefore particularly unsafe

- contents are unprotected in case of loss, if they were not encrypted before (see also unit 11 point 2, p. 89)
- susceptible to shock and wear (external hard drive)

Institutional storage location

- + backup of data is ensured
- + professional implementation and maintenance
- + storage in accordance with the institution's data protection policy
- + data protection regulated by access rights
- + usable for remote office work
- speed may depend on network conditions
- access to backups may be delayed due to the chain of command
- possibly unclear which security criteria and strategies are used
- may involve higher costs

External storage location

- + easy to use and manage
- + are professionally maintained
- + usable for remote office work
- depending on the provider, connection may also be insecure
- depending on access to internet
- upload and download may take some time
- access to backups may be delayed
- data protection: unclear which security criteria and strategies are used and whether they are appropriate for sensitive data and meet the requirements of your institution
- many institutions have issued special regulations for the use of such services, e.g. Freie Universität Berlin⁵⁴

Free cloud storage services are not appropriate locations for data that needs to be protected. It is questionable how secure the data is in a cloud and who has control over what happens to the data. Although it is the task of cloud service providers to prevent misuse of data, it is unclear whether they are able to do so comprehensively.

In addition, researchers can also contribute physically to securing their (sensitive) data (see also unit 11: access control, p. 88). For example, they can store their storage devices in a separate lockable room or cabinet. Notebooks can be secured against theft by a lock. It is important, however, that at least two people have access to the data in order to guarantee the accessibility of the data in case of illness or absence.

2. Backup options

Backup means the creation of a copy of the data on another storage device. A backup should be carried out in a planned and structured manner in order to facilitate data reconstruction if necessary.

⁵⁴ Freie Universität Berlin. „Richtlinie zur Auslagerung von Daten in die Cloud“. Freie Universität Berlin, 2011. Last accessed 26.02.2019. https://www.fu-berlin.de/sites/itsicherheit/downloads/Richtlinie_Cloud-Datenablage_-_1_0.pdf

Only when the hard drive or the external storage is no longer functioning or gets lost, does one notice how important a regular backup is. If this has been set up and carried out correctly beforehand, the lost data can be easily restored. To avoid file loss due to defective hardware or software, you should therefore think about backup strategies beforehand.

Many operating systems are already equipped with their own backup programs and do not require any installation effort (e.g. Mac provides Time Machine).

If you are not satisfied with these integrated solutions, you can choose from a whole range of programs (some of which require payment), e.g. Duplicati, Cobian Backup, Areca Backup or Aoemei Backupper Standard. These offer two different backup types: incremental or differential. With both types, a full backup is created first. After that, the

- incremental backup only saves files or parts of files that have changed or been added since the last incremental backup,
- differential backup saves all data that has changed or been newly added since the last full backup.

3. Strategies for a secure backup

Backup devices should be separate from the infrastructure usually used. It is recommended to save data at least once a day and to perform a weekly overall backup. These '3-2-1' principles have proven useful for backups:

- at least three copies of your data,
- on at least two different storage devices and
- one of them should be stored remotely, at a different location.

Data recovery should be checked at the outset and at regular intervals.

Most institutions offer an automatic solution where all data is stored exclusively on drives safeguarded by the IT department. This professionalisation ensures the backups will not be forgotten and the configuration of the backup system does not have to be done individually.

Didactic methods and exercises

Methods:

- Ideas out loud
 - Who of you knows if your institution provides a backup service? And who can explain how it works?

Exercises:

- Work out in groups the advantages and disadvantages of different storage devices.

Training materials

- Worksheet: Storage and backup
- Proposed solutions for the worksheet: Storage and backup
- Argumentation aid: Why should I secure my data and use a backup?
- Checklist: What do I have to consider when choosing a storage device?
- Teaching script: Storage and backup

Additional sources

- Hanson, Karen, Alisa Surkis and Karen Yacobucci: Data Sharing and Management Snafu in 3 Short Acts. NYU Health Sciences Libraries, 2012. Video, Youtube. Last accessed 26.02.2019. https://www.youtube.com/watch?v=66oNv_DJuPc



Worksheet: Storage and backup

Advantages	Disadvantages
Own PC	
Mobile storage (USB stick, external hard drive)	
Institutional storage (cloud, virtual drives)	
External storage (cloud of a free or commercial provider)	



Proposed solutions for the worksheet: Storage and backup

Advantages	Disadvantages
Own PC <ul style="list-style-type: none"> • self-responsible for security and backup • maximum control 	<ul style="list-style-type: none"> • what happens with the PC happens with the backup • possible lack of resources and know-how • individual solutions are complex
Mobile storage (USB stick, external hard drive) <ul style="list-style-type: none"> • easy to transport • can be stored in a lockable cabinet or safe 	<ul style="list-style-type: none"> • loss, theft (very unsafe) • in case of loss: contents are unprotected, if not encrypted • external hard drive is susceptible to shock and wear
Institutional storage (cloud, virtual drives) <ul style="list-style-type: none"> • backup of the data is guaranteed • professional implementation and maintenance • storage according to data protection guidelines of the institution 	<ul style="list-style-type: none"> • speed may depend on network • access to backups possibly delayed by service path • possibly unclear which safety criteria and strategies are used
External storage (cloud of a free or commercial provider) <ul style="list-style-type: none"> • easy to use and manage • backup of the data is guaranteed • can be used for mobile work • professional execution and maintenance 	<ul style="list-style-type: none"> • depending on the provider the connection can also be insecure • depending on internet access (upload and download might be slow) • access to backups may be delayed • issues with data protection



Argumentation aid: Why should I save my data and use a backup?

Below you will find three different scenarios to illustrate the importance of backups and highlight some important aspects when planning a backup strategy. What could have been done to prevent data loss?

1. “Fire destroys leading research institute”: On 30 October 2005 a fire broke out at the University of Southampton in the early hours of the morning. A leading computer science research unit was destroyed. It is estimated that the fire caused 50 million pounds of damage. A spokesman said it was where some of the world's most advanced research was conducted and its loss was devastating.

Source: BBC News. Online available:

http://news.bbc.co.uk/2/hi/uk_news/england/hampshire/4390048.stm

2. “Encryption Trojan”: In 2016, more and more encryption Trojans (also known as ransomware) were sent by e-mail (as office attachments, JavaScript files, packed in zip files). These Trojans encrypt the data of the affected device and demand a ransom for decryption or release. Approximately 5,000 computers were infected per hour with variants of the Ransomware Locky. Among the victims was also the Fraunhofer Institute in Bayreuth. Source: heise Online. Online available:
<https://www.heise.de/security/meldung/Krypto-Trojaner-Locky-wuetet-in-Deutschland-Ueber-5000-Infektionen-pro-Stunde-3111774.html>

3. “Lost bag”:

A screenshot of a Twitter post. The profile picture is a small portrait of a man. The name is "Andrew Penson" and the handle is "@ADPenson". There is a "Folgen" (Follow) button. The tweet text is: "****Please SHARE****Lost bag containing all my research notes on laptop and paper. I can't complete my research masters degree without this - eighteen months worth of daily research. Left it on a train arriving at Charing Cross Saturday 8th July 08:50 ish which became the 09:06 to Sevenoaks. Black and grey with red "Swiss" features. I'm ruined if I don't get it back, please share as widely as possible! Thank you. andrewpenson@hotmail.co.uk". At the bottom left of the tweet card, it says "03:41 - 8. Juli 2017".

Source: Twitter. Online available:

<https://twitter.com/ADPenson/status/883637257323896832>



Checklist: What do I have to consider when choosing a storage device?

- How much storage space do I need?
.....
- Which file types do I have and how often will I replace them?
.....
- Who needs access?
.....
- Is it necessary to have remote access to the data?
.....
- How important is fast access?
.....
- Is simultaneous and synchronised access required?
.....
- What precautions should I take to protect my data from loss (password, encryption, physical protection, etc.)?
.....
- Which storage solutions are suitable for personal data (if applicable)?
.....
- How often will I make a backup and where will it be stored?
.....
- How much money do I have available to store the data?
.....

Source:

Based on: CESSDA Training Working Group. CESSDA Data Management Expert Guide. Bergen, Norway: CESSDA ERIC, 2017-2018, <https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/4.-Store/Adapt-your-DMP-part-4>. This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

9. Teaching script: Storage and backup							
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound
						Alternative	Notes
Data storage	The participants work out the advantages and disadvantages of different storage media and service offers	6	Different storage media are compared and their advantages and disadvantages are highlighted. The participants fill out the worksheet	Group work (group of 3): complete the worksheet together	Worksheet, pens	Ex	Yes
Backup	The participants learn the strategies for a secure backup	4	The criteria of a secure backup are presented. The institutional service offers are pointed out	PPTX (for alternative: internet access)	Presentation	In	Video: https://www.youtube.com/watch?v=66oNv_DJuPc (ca. 5 min); Question to the group: What services are available at your institution?

Duration of the unit: 10 minutes



Unit 10: Long-term archiving

Learning objectives

The participants know the requirements for long-term archiving and critically examine questions on the sustainability of file formats.

Participants are aware that special precautions must be taken to ensure that data can be made accessible for a longer period of time and to preserve its authenticity and integrity.

Key aspects

1. Motivation
2. The term "long-term archiving"
3. Sustainable file formats
4. Requirements for long-term archives

Contents

1. Motivation

In order to keep data searchable, accessible and readable in the future, it is archived. The requirements of the DFG's Code of Conduct on "Good Research Practice"⁵⁵ demand that relevant research data be made available for 10 years. Many academic institutions require their researchers to ensure the long-term preservation of their data (e.g. within the framework of a research data policy).

2. The term "long-term archiving"

"Long-term" is an auxiliary word to describe an unspecified period of time during which technological and socio-cultural changes may occur that may affect the preservation, access, search and reuse of digital research data. Accordingly, digital long-term archiving comprises a series of measures that must be planned, controlled and carried out.

⁵⁵ Deutsche Forschungsgemeinschaft: *Guidelines for Safeguarding Good Research Practice. Code of Conduct*. Zenodo, 2019. <http://doi.org/10.5281/zenodo.3923602>.

3. Sustainable file formats

Not every file format is suitable for long-term archiving. Table 2 gives an overview of the most common format recommendations.⁵⁶⁵⁷ In this context, a distinction is made between proprietary and open formats. Proprietary formats are those that require commercial software (e.g. Microsoft Office, AutoCAD, SPSS, MaxQDA). The files to be archived should be unencrypted, uncompressed, patent-free and created in an open, documented standard. These formats require less frequent migration⁵⁸ and are characterized by a longer life span and higher distribution.

Sometimes proprietary file formats are indispensable for your own work. For long-term digital preservation, however, they should be converted into recommended formats. It is important to check whether the conversion was successful and the format is valid, as software may also produce errors. Both the original file and the file in the converted format should be saved.

Table 2: Format recommendations for long-term archiving.

File format	Recommendation	Avoid
Tables	CSV, TSV, SPSS portable, XLSX	XLS, SPSS
Text	TXT, HTML, RTF, PDF/A, DOCX	DOC, PDF
Multimedia	Container: MPEG4, MKV Codec: Theora, Dirac, FLAC	QuickTime, Flash
Images	TIFF, JPEG2000, PNG	GIF, JPG

4. Requirements for long-term archives

The following aspects should be considered when selecting a suitable storage for long-term archiving:

- Technical requirements: the service provider should have a strategy for data conversion and migration. In addition, a check of the readability of the files and a virus check should be carried out at regular intervals. All steps should be documented.
- Seal for trustworthy long-term archives: "A digital long-term archive is considered trustworthy if it operates in accordance with its goals and specifications for information preservation over long periods of time and if its users, producers, operators, partners rely on it".⁵⁹ For an external assessment as to whether or to what extent a long-term archive is trustworthy, various seals with different focuses are developed, which do not address every type and operating model of repositories equally (e.g. Nestor seal, DIN 31644 or CoreTrustSeal).
- Cost of services: Always check whether service providers charge for data storage. The costs can depend, for example, on the amount of data, the implementation of technical standards or the affiliation of the data providers.
- Making the data accessible: Before choosing the storage location, you should ask yourself whether the data should be accessible or just stored.
- Longevity of the service provider: Economic and political factors influence the longevity of service providers.

⁵⁶ IANUS – Forschungsdatenzentrum für Archäologie & Altertumswissenschaften: „IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften. Dateiformate“. Version: 1.0.1.0. Berlin: IANUS – Forschungsdatenzentrum für Archäologie &

Altertumswissenschaften, 2017. <http://doi.org/10.13149/000.111000-a>.

⁵⁷ UK Data Service. „Recommended formats.“ Last accessed 29.07.2020.

<https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats>.

⁵⁸ Data migration is the process of moving data from one system to another.

⁵⁹ Keitel, Christian: *Vertrauenswürdige digitale Archive: DIN Norm 31644*. Workshop Archivierung sozial- und wirtschaftswissenschaftlicher Datenbestände. Frankfurt am Main: DNB, 2011. Last accessed 03.11.2020. https://www.ratswd.de/ver/docs_Archivierung_2011/keitel.pdf (transl. M.B.).

Didactic methods and exercises

Methods:

- Ideas out loud
 - Which formats do you use?
 - Which formats could you convert into?
 - Which problems could occur?
 - Who can do that? Who could help?
 - What do you expect from a long-term archive for your data?
 - How does archiving differ from backup?

Exercises:

- Which format to choose? (Advantages and disadvantages of .docx, .txt, PDF/A)
- Work in groups to develop criteria for selecting a long-term archive.

Training materials

- Checklist: What should be considered when deciding on a long-term archive?
- Teaching script: Long-term archiving

Additional sources

- Digital Curation Centre. "Five steps to decide what to keep: a checklist for appraising research data" Version 1. Edinburgh: Digital Curation Centre, 2014. Last accessed 03.11.2020. <http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep>.
- IANUS – Forschungsdatenzentrum für Archäologie & Altertumswissenschaften: „IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften. Dateiformate“. Version: 1.0.1.0. Berlin: IANUS – Forschungsdatenzentrum für Archäologie & Altertumswissenschaften, 2017. <http://doi.org/10.13149/000.111000-a>.
- KFM Wissenschaftliches Kompetenzzentrum für Mehrsprachigkeit: "Leitfaden zur Datenspeicherung." Last accessed 03.11.2020. https://campus.hesge.ch/researchdatamanagement/wp-content/uploads/2016/06/LOOK_kfm_leitfaden_datenspeicherung.pdf.



Checklist: What should be considered when deciding on a long-term archive?

- How long the data shall be kept?
.....
- How much storage space do I need?
.....
- Which file formats do I have? Do they need to be converted into sustainable formats?
.....
- Who needs access?
.....
- Where is the data and its documentation stored after the end of the project?
.....
- Does the service provider have a strategy for data conversion and migration?
.....
- Is the integrity of the data regularly checked?
.....
- Is the long-term archive trustworthy? Does it have a seal?
.....
- How durable is the service provider?
.....
- How often a backup is made and where is it stored?
.....

Source:

Based on: CESSDA Training Working Group. CESSDA Data Management Expert Guide. Bergen, Norway: CESSDA ERIC, 2017-2018, <https://www.cessda.eu/DMGuide>. This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).



10. Teaching script: Long-term archiving									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternative	Notes
Basics	The participants activate their previous knowledge	3	How does archiving differ from backup?	Ideas out loud	-	Ex	Yes		
	The participants understand the term long-term archiving	1	Explanation of the term and differentiation from backup. Challenges of archiving are explained	Presentation	PPTX	In	-		
Sustainable file formats	The participants learn about suitable file formats for archiving	3	Explanation of the difference between open and proprietary formats and the reasons for using standardized formats for long-term archiving	Presentation	PPTX	In	-		
Requirements for a long-term archive	The participants develop criteria for the selection of a suitable long-term archive	6	Development of criteria for the selection of a long-term archive. The participants answer the question: What should be considered when making the selection?	Group work	Moderation cards (alternatively flip chart sheets), pens	Ex	Yes		2-3 groups in total
	The participants discuss the criteria for selecting a suitable long-term archive and get to know others	7	The groups present their criteria	Presentations by the participants	-	In and Ex	Yes		Checklist: What should be considered when deciding on a long-term archive?

Duration of the unit: 20 minutes



Unit 11: Access control

Learning objectives

Participants are aware that data security and the distribution and maintenance of access rights are important aspects of data management.

Participants will learn methods for increasing the protection of their data and setting up access rights as well as useful criteria for their allocation.

Key aspects

1. Motivation
2. Encryption
3. Checking the usability
4. Password protection and access rights

Contents

1. Motivation

Research data are among the most valuable resources in science, so great importance is attached to its protection. The aspects of data security and access rights of data management should consist of measures to protect against data loss on the one hand and measures to prevent misuse of the data on the other. Sometimes there are sensitive data that must be appropriately protected. These may be personal data as well as data whose protection has been contractually guaranteed (e.g. company secrets, commissioned research). Research results that have not yet been published also require protection.

Data may be protected by means of encryption, backups, storage on trustworthy storage devices and/or by specific regulation of access rights. The usage becomes traceable through creation and retention of log files. With all security measures, it is vital to ensure all persons who need the data for their work have access to the data. The regulation of these aspects is particularly relevant in case of inter-institutional cooperation.

It is useful to set up and document the technical and organisational measures suitable for data security and to check regularly whether they still meet current requirements.

2. Encryption

Physical access to a computer may allow unauthorized access to data, therefore it might be necessary to encrypt the data. However, this measure only makes sense if all affected data is encrypted, including, in particular, copies and backups. To do this, all parties involved must be aware of the need for encryption and of all storage locations.

It is possible to encrypt selected storage locations or even entire data carriers. Automatic encryption solutions – using programs such as FileVault, Bitlocker or dm-crypt – are another conceivable option. These programs must be set for all intended data storage locations.

Furthermore, file encryption is an additional protective measure when uploading data into a cloud environment.

If data is passed on by e-mail, encryption of the e-mail is recommended as well. This prevents unauthorized reading and modification of messages and attachments on their way to the recipient (end-to-end encryption).

3. Checking the usability

A data backup is only helpful if data recovery is guaranteed. Sometimes files are corrupted and become defective from then on. Sometimes copying files causes errors by itself. It's a good idea to test your data recovery at the outset of the backup and at regular intervals to prevent data loss.

In addition to checking for readability, a virus check should also be performed. Otherwise, faulty files may overwrite undamaged backup files.

4. Password protection and access rights

When dealing with data worthy of protection, secure passwords should be assigned and access should be restricted to the circle of persons directly involved.

- A secure password bears the following characteristics⁶⁰:
- The longer the better. It should consist of at least 8 characters
- contains lower- and upper-case letters as well as special characters and numbers
- used characters should not be next to each other on the keyboard
- the password should not appear in dictionaries.

Avoid simple passwords such as 123456, password, 111111, qwerty, abc123 or admin. Names, birthdays, terms from dictionaries, movie characters or license plates shouldn't be used as passwords either, even if they are written backwards. Similarly, adding numbers or special characters to a word does not make a secure password.

The assignment of authorizations determines which persons or groups of persons are allowed to access certain directories and files and with which specific rights. It is possible to assign gradual read and write rights as well as execution rights. Thus, some users only have access to view the data, while others are granted full access to the data. It is important that the assignment is well thought out so as not to interfere with the workflow.

If access rights are too lax, people may access sensitive data who should not be able to do so for privacy reasons. If, on the other hand, access rights are granted too restrictively, the FAIR principles are violated and subsequent use is made more difficult or even prevented.

⁶⁰ Bundesamt für Sicherheit in der Informationstechnik: „Passwörter“. Last accessed 08.07.2020, https://www.bsi-fuer-buerger.de/BSIFB/DE/Empfehlungen/Passwoerter/passwoerter_node.html.

Didactic methods and exercises

Methods:

- Ideas out loud
 - Why is data security important?
 - How do you handle your data?
 - What changes should you make?
 - How is this change to be achieved?
- Estimation question
 - Participants develop their own recommendations for secure passwords and their memorization aids or places where a password list can be deposited, and then discuss them in pairs or threes.
- Chatter
 - What do I want to change/improve after this learning unit?
- Blind decision
 - Are my data safe?
 - Is a personal safe useful?
 - Is it wise to work in a cloud?

Training materials

- Teaching script: Access control

Additional sources

- Bundesamt für Sicherheit in der Informationstechnik: „Passwörter“. Last accessed 08.07.2020, https://www.bsi-fuer-buerger.de/BSIFB/DE/Empfehlungen/Passwoerter/passwoerter_node.html.
- Leibniz Universität Hannover. „Datenverschlüsselung.“ Last accessed 03.11.2020. https://www.luis.uni-hannover.de/its_encryption.html.

11. Teaching script: Access control									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voice sound	Alternative	Notes
	The participants get to know reasons for a secure handling of data	3	Reasons are presented	Presentation	PPTX	In	-		
Introduction	The participants add their reasons for secure data handling	3	Participants supplement the lecture with their own reasons	Ideas out loud	-	Ex	Yes		May be left out if no other reasons are given
Encryption and physical protection	The participants learn about the options for encryption and physical protection of data	6	The importance of physical protection and encryption is presented and tools are given as examples (e.g. FileFault, Bitlocker, dm-crypt)	Presentation	PPTX	In	-		
Password protection	The participants can create secure passwords	2	The criteria for creating secure passwords are explained	Presentation	PPTX	In	-		
Access rights	The participants learn about the criteria for granting access rights	1	Criteria for granting access rights are presented	Presentation	PPTX	In	-		

Duration of the unit: 15 minutes



Unit 12: Formal framework

Learning objectives

The participants get to know the 3T formula.

Participants critically examine the aspects of content and organisation in setting up the training workshop.

Key aspects

1. Clarify the terms and conditions
 - a) (objective) target
 - b) time
 - c) target group
2. Content aspects of the workshop structure
3. Organisational aspects of the workshop structure

Contents

1. Clarify the terms and conditions

Before you start organising a workshop, training or class, you should clarify the terms and conditions for the event. The 3T formula of Martin Lehner⁶¹ is suitable for this purpose. (Objective) target, time and target group. Whether the event proves to be useful depends largely on this.

⁶¹ Lehner, Martin: *Viel Stoff – wenig Zeit. Wege aus der Vollständigkeitsfalle*. 4th ed. Bern: Haupt, 2013.

a) (objective) target

The first and most important aspect is the objective target. The teacher must know exactly the goal of each learning unit and he/she must be able to specify it. When defining the objective target, one must consider what the main focus of the workshop will be.

b) time

The time available is an important factor that determines whether all desired learning goals can be achieved. Often the time frame is fixed from the beginning and so the learning content must be adapted. Sometimes you will have the opportunity to determine the time frame yourself when designing the workshop.

c) target group

The more you know or can presume about the participants, the better you can tailor the contents to them: Do they have little or much previous knowledge, are they a beginner, advanced or at a professional level? How old are they? What is the gender balance? Do they know each other? How many similarities do they have? What are their expectations? What learning methods are they used to?

Only if you know your target group, you can adapt the contents and the teaching methods to their needs.

2. Content aspects of the workshop structure

As soon as the framework conditions for a workshop are determined on the basis of the 3T formula, the subject-related contents must be aligned.

The first step is to define the indicative objectives – these have already been outlined briefly in the conditions.

The second step is to define the learning objectives: what should the participants be capable of in the end? What are they supposed to know by then? Based on these learning objectives, the individual topic areas can be highlighted.

In the last step of the content aspects for the workshop setup, the sequence of the units is deliberated. Teaching scripts with exact times, learning objectives and topics covered, as well as type of work and materials required, provide support. At this point you should also consider the methods you want to use in teaching.

3. Organisational aspects of the workshop structure

When organising a workshop, besides the content aspects, the organisational aspects must also be carefully considered. The date of the event should be determined as early as possible, thus enabling you to book an appropriate room. The equipment depends on the needs of the trainer. Depending on financial means, catering could also be provided (coffee, lunch break, pastries). In order to create a friendly, inclusive and respectful environment for the workshop, it is a good idea to set up rules for the workshop (Code of Conduct) and communicate these to the participants in advance of the workshop. The rules should be briefly presented no later than at the beginning of the workshop or, in the case of longer formats, they can be worked out together.⁶²

You should also think about registration and about certificates of attendance (see p. 176) for all participants to receive at the end of the event.

In order to make the workshop a success and to avoid confusion, it is recommended that the participants are informed in advance about the most important organizational cornerstones. The workshop website (if available) or an e-mail are suitable for this purpose. In addition to the basic information about the

⁶² See ‘Code of Conduct’ in Sonja Bezjak, April Clyburne-Sherin, Philipp Conzett, Pedro Fernandes, Edit Görögh, Kerstin Helbig, Bianca Kramer et al.: Open Science Training Handbook. Zenodo, 2018, pp. 99-100. <http://doi.org/10.5281/zenodo.1212496>. Further links with examples for Code of Conducts can be found there.

workshop (topic, trainers, place, time, directions, etc.), important aspects are (if applicable) the code of conduct for the workshop, which equipment or software will be used, or advice for preparatory reading.

Didactic methods and exercises

Methods:

- Mind map
 - What do I need to consider when planning a workshop?

Training materials

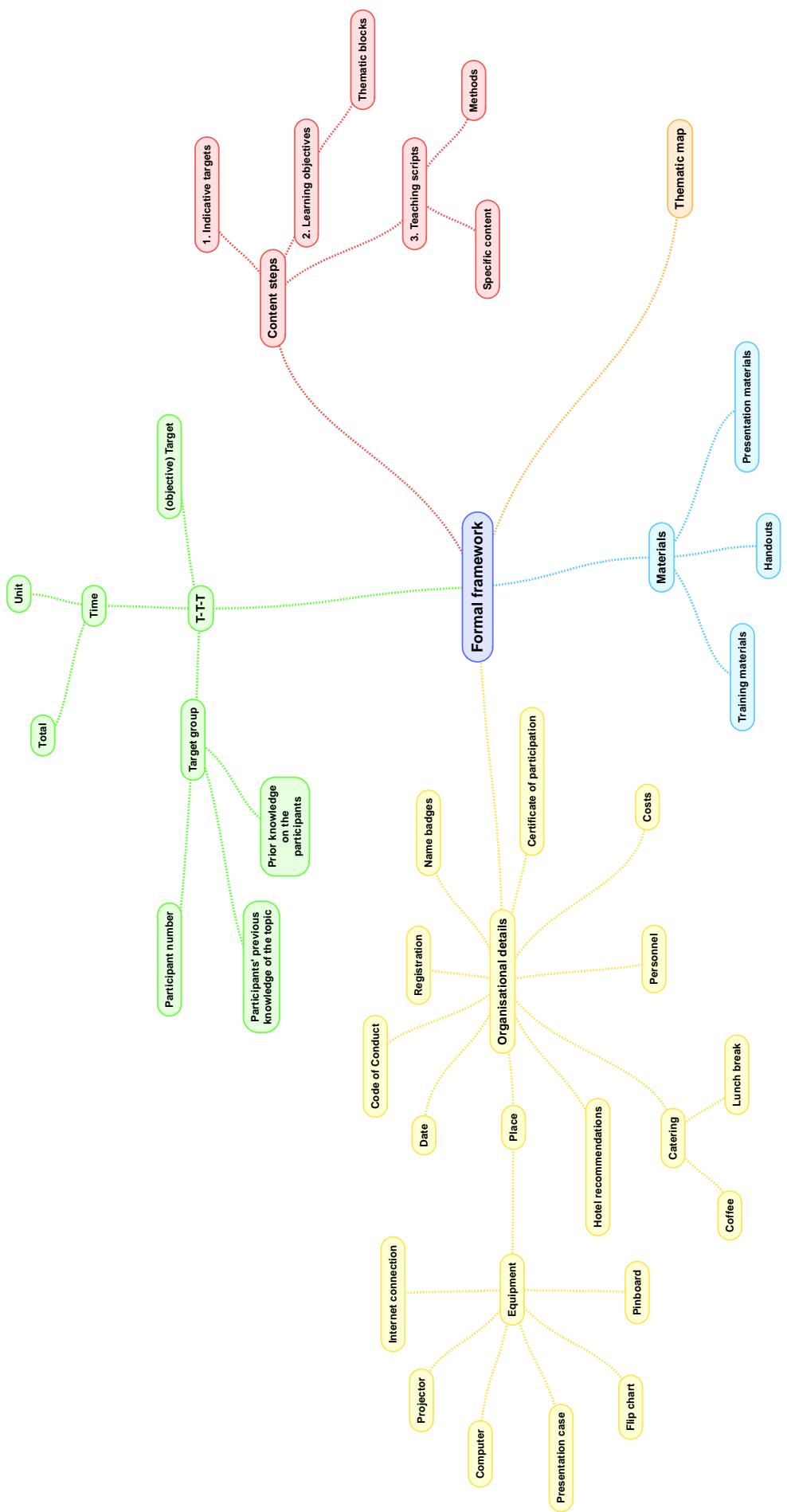
- Template: Mind map on the subject of formal framework
- Checklist: Organisation of workshops⁶³
- Teaching Script: Formal framework

Additional sources

- Bezjak, Sonja, April Clyburne-Sherin, Philipp Conzett, Pedro Fernandes, Edit Görögh, Kerstin Helbig, Bianca Kramer et al.: Open Science Training Handbook. Zenodo, 2018. <http://doi.org/10.5281/zenodo.1212496>.
- Lehner, Martin. *Viel Stoff – wenig Zeit. Wege aus der Vollständigkeitsfalle*. 4. Aufl. Bern: Haupt, 2013.

⁶³ Based on Sonja Bezjak, April Clyburne-Sherin, Philipp Conzett, Pedro Fernandes, Edit Görögh, Kerstin Helbig, Bianca Kramer et al. *Open Science Training Handbook*. Zenodo, 2018, pp. 142–143. <http://doi.org/10.5281/zenodo.1212496>.

Template: Mind map on the subject of formal framework





Checklist: Organisation of workshops

What?	When and where?	Done?
Book a venue		
Determine which technical equipment is required		
Check the equipment of the event location		
Check whether there are enough power sockets available		
Set up WLAN for guests		
Order video or photo shoots		
Test the equipment a few days before the event		
Print out handouts, worksheets, feedback forms and materials or make them available online		
Check if there are enough flipcharts and pin boards available and prepare them		
Check access to the room (handicapped accessible)		
Place signs to the event room in the building		
Check parking possibilities		
Find and inform helpers before the event		
Locate gender-neutral toilets, prayer room and maternity room		
Identify communication channels		
Advertising (e.g. via social media)		
Prepare online presence		



What?	When and where?	Done?
Prepare registration procedure		
Send event information to the known mailing lists		
Send registration confirmations including the key event data to the participants		
Send a reminder to the participants 1-2 days before the event		
Prepare name tags		
Print participant list		
Organize catering (e.g. coffee and biscuits)		
Inform participants about emergency exits, catering and toilets		
Obtain consent for video, photo, or live streaming		
Distribute feedback forms to fill out or provide the link for the online version		
Take photos of the flipcharts and other non-digital materials and results (photo protocol)		
Distribute digital training materials and photo protocol after the event		
Hand over or send participation certificates		

Based on Bejak et al.: *Open Science Training Handbook*. 2018. DOI: 10.5281/zenodo.1212496

12. Teaching script: Formal framework							
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound
						Alternative	Notes
Introduction	Activate previous knowledge	10	The participants work in three groups on a mind map on the topic "Planning a workshop?"	Group work, Method: Mind map	Paper (A3), pens	Ex Yes	The group is divided by counting from 1-3; alternatively: distributing playing cards and assigning them according to card motif (Ace, King, Queen). Required: Playing cards GR1: Terms and conditions GR2: Contents GR3: Organisation
First step	The participants learn the terms and conditions (3T formula: target, time, target group)	2	Group 1 answered: What should you start with?	Ideas out loud	Mind map worked out in groups; pens	In and ex	Yes
		3	Discussion of the prepared mind map	Presentation	Section of a prepared mind map (T-T-T)	In -	The mind map should be visible for the rest of the workshop
Content	The participants learn to create the contents of a WS (indicative goals, learning objectives, schedules, etc.)	2	Group 2 answered: How do you develop the contents?	Ideas out loud	Prepared mind map, pens	Ex Yes	
		7	Discussion and emphasis on the importance of the order of indicative goals, topic blocks, learning goals, schedules, contents and methods	Presentation	Section from prepared mind map (content steps)	In -	
Organisational information	The participants know what belongs to the organisational preparation of a workshop	3	Group 3 answered: Which organisational steps are part of the planning?	Ideas out loud	Prepared mind map, pens	Ex Yes	
		2	Addressing the individual points for organisational preparation	Presentation	Section from prepared mind map (organisational)	In -	
Summary	Participants get a general overview	1	Presentation of the entire mind map	Presentation	Prepared mind map	In -	Checklist for the organisation of workshops

Duration of the unit: 30 minutes



Unit 13: End of the first day

Learning objectives

Participants will repeat and review what they have learned.

Key aspects

1. Recap
2. Obtain feedback
3. Farewell

Contents

1. Recap

It is recommended to end the first day of the workshop with a short summary and review. During this step, participants must remember what they have just learned and repeat it in their own words. This will help them to acquire and consolidate the new material.

2. Obtain feedback

It is important for the trainer to gain feedback from the participants at all times. This is the only way to adjust the teaching sequence and learn from it yourself. It is just as interesting to hear what the participants take home with them from this training day.

3. Farewell

A formal and proper farewell from the teacher completes the first working day and releases the participants from work.

Didactic methods and exercises

Methods:

- Inventory
 - Recall what you have learned today.
- Ideas out loud

- What will you take home with you from today's session?
- Feedback
 - What else would you like to tell us?

Training materials

- Teaching Script: End of the first day

Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternative	Notes
Recap	The participants remember the learning contents	10	The participants remember what they have learned today	Method: Inventory	Paper A3, pens	Ex	Yes	S: Ideas out loud	
	The participants transfer what they have learned to their everyday work	5	The participants answer the question: What can you use for yourself?	Ideas out loud	-	Ex	Yes		
	The participants get the opportunity to communicate, to add or ask, if something is important to them	3	The participants address questions and feedback to the trainer	"What else would you like to tell us?"	-	Ex	Yes	Summary of the day as a short message to the colleagues	
Farewell		2	The trainer gives a short outlook on the 2nd part, thanks the participants and bids them farewell until the second workshop day	Announcement	-	In			

Duration of the unit: 20 minutes



Unit 14: Welcome and orientation on day 2

Learning objectives

The participants remember what they have learned on the first day and reproduce it.

The participants know the broad outline of the 2nd workshop day.

Key aspects

1. Welcome
2. Repeat
3. Orientation

Contents

1. Welcome

As on day one of the workshop, the welcome on day two is used to introduce the event and to draw the participants' attention to the 'here and now'.

It might be helpful for the teacher to ask about the participants' mood or energy in order to better assess the overall atmosphere in the classroom.

2. Repeat

Depending on the interval between the first and second day of the workshop, the new knowledge may still be very fresh in the minds of the participants, but it may also be forgotten. In order to bring the participants quickly back to the subject matter, it is recommended to repeat what they have learned. In this way you activate the participants and summon up their knowledge.

3. Orientation

Orientation is also essential on the second day of the workshop. The teacher will give a broad overview of the topics and the course of the day, which the participants can use for orientation.

Didactic methods and exercises

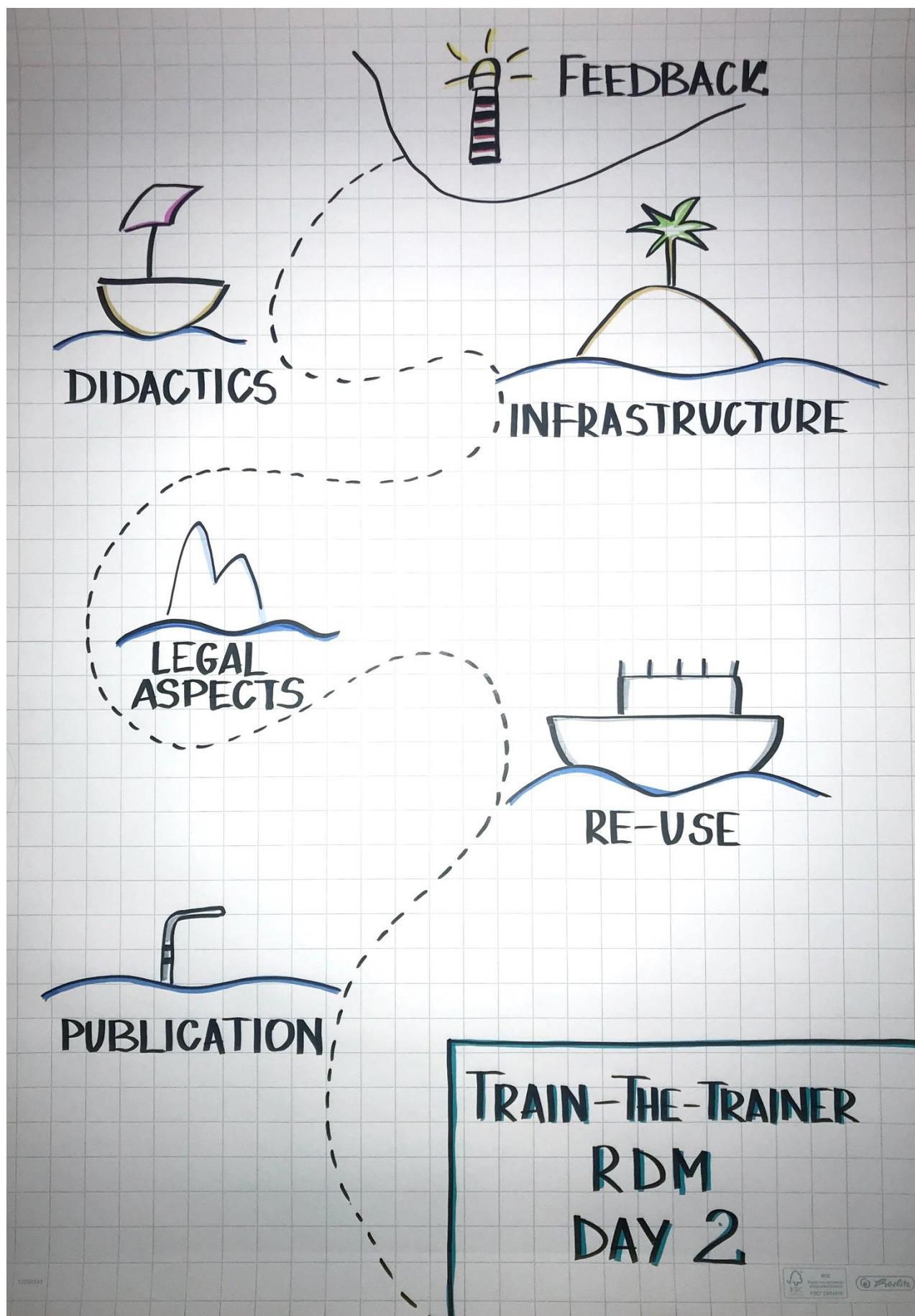
Methods:

- Energy level
- Keyword strips
 - 3T formula
 - Backup
 - Research data lifecycle
 - Long-term archiving
 - DMP
 - Dublin Core
 - Research Data Policy
 - Metadata
 - Inhale and exhale
 - Components of a documentation
- Thematic Map for the second day

Training materials

- Thematic map for the second day of the Train-the-Trainer Workshop on RDM
- Template: Keyword strips
- Teaching script: Welcome and orientation on day 2

Thematic map for the second day of the Train-the-Trainer Workshop on RDM



Template: Keyword strips	
3T formula	Backup
Research data lifecycle	Long-term archiving
DMP	DublinCore
Research data policy	Metadata
Inhalation and exhalation	Elements of a documentation

14. Teaching script: Welcome and orientation on day 2						
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation
					Voices sound	Alternative
Welcome	The attention of the participant is attracted	2	T welcome the participants and ask how much energy they have at the moment (0-10). Participants say number	Energy query	-	In and ex
Remember and repeat	The participants remember what they have learned and can consolidate knowledge	10	The participants remember what they learned on the 1st day and can ask questions or answer each other	Method: Keyword strips	Prepared keywords in 3 envelopes	Yes
Orientation	Participants get an overview of the contents of the 2nd day	5	T explains, which contents are covered	Presentation	Flip chart	In -

Duration of the unit: 17 minutes



Unit 15: Data publication

Learning objectives

The participants are aware of the benefits of and barriers to data publication.

Participants know what a repository is and learn about different types of repositories.

Participants learn about different ways to publish their data.

Participants know that the rights of third parties must be respected when publishing research data and are introduced to some examples.

Key aspects

1. Benefits of and barriers to data publication
2. Key questions for the selection of data
3. Methods for publishing data
4. Repositories
5. Types of repositories
6. Selection of a repository
7. Licences
8. Persistent Identifiers
 - a) Digital Object Identifier (DOI)
 - b) Open Researcher and Contributor ID (ORCID)

Contents

1. Benefits of and barriers to data publication

To allow data to be re-used for research beyond the initial research question, it should be published. It is essential that these data are findable, accessible, interoperable and reusable.⁶⁴ There are certain objections to the publication of data, but there are also good reasons to do so. Both are discussed below.

Researchers increasingly have to deal with competition when applying for public funding and publishing their results. Data can be regarded as a public investment. Their publication enables researchers to gain further recognition, namely for their data as a scientific asset that stands for itself and that may be brought to the table in future applications for funding.⁶⁵

The publication of data contributes to academic integrity. It makes research replicable and transparent. Replication of the results by third parties verifies one's own work and has a positive effect on the reputation. There are indications that articles whose data have been published are cited more often.⁶⁶

Some researchers are concerned about their data being misinterpreted, edited or misused. But these concerns are void if the data constitute a significant part of the study in which they are reused and if this results in new citation, collaboration or co-authoring for the data originators.

Sharing data within one's own discipline helps to advance the current state of knowledge. Researchers want to maintain their interest in publishing the findings from their data comprehensively and as the primary source. They fear other researchers might publish results based on their research data which could overlap with their own planned publications and thus pre-empt and complicate their publication. However, it is the authors of the data who decide on an embargo period, i.e., they decide whether and when they publish their data and from what time on the data can be used by others.

By reusing the data, duplicate surveys and thus unnecessary costs are avoided, which enables a more efficient allocation of resources. In addition, published data are excellent resources for training and teaching.

The preparation of data for publication as well as the processing of requests for data can be very time-consuming for the data producers. Dealing with the publication of data at an early stage in the research process will lead to a better and more consistent documentation and quality of the data, which in turn may facilitate the publication of the research results and the long-term archiving of the data.

Last but not least, the publication of data is increasingly required by publishers, institutions and funding agencies.

2. Key questions for the selection of data

For each publication, it must be decided under which terms it will be published, for example whether it will be made freely accessible (open access) or archived in an access-controlled manner. Competitive pressure within the scientific community can make a limited or time-delayed publication worthwhile: if

⁶⁴ Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al.: „The FAIR Guiding Principles for scientific data management and stewardship.” *Scientific Data* 3 (2016). <https://doi.org/10.1038/sdata.2016.18>. See also unit 4: Digital research data.

⁶⁵ Poisot, Timothée E., Ross Mounce and Dominique Gravel: Moving toward a sustainable ecological science : don't let data go to waste! *Ideas in Ecology and Evolution* 6,2 (2013): pp. 11-19. <https://doi.org/10.4033/iee.2013.6b.14.f>

⁶⁶ Pienta, Amy M., George C. Alter and Jared A. Lyle. *The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data*. Ann Arbor, Mi: University of Michigan – Deep Blue, 2010. <http://hdl.handle.net/2027.42/78307>; Piwowar, Heather A and Todd J. Vision. Data reuse and the open data citation advantage. *PeerJ* 1:e175 (2013) <https://doi.org/10.7717/peerj.175>; Henneken, Edwin A. and Alberto Accomazzi, A.: *Linking to data-effect on citation rates in astronomy*. arXiv 1111.3618, 2011. <https://arxiv.org/pdf/1111.3618.pdf>; Piwowar, Heather A., Roger S. Day and Douglas B. Fridsma: Sharing detailed research data is associated with increased citation rate. *PLoS ONE* 2,3,e308 (2007). <https://doi.org/10.1371/journal.pone.0000308>

the data collected is to be used for further publications, the time of publication of the data and the choice of publication model are crucial aspects.

Below you will find a selection of key questions⁶⁷ to be answered anew before each publication of data:

- Is it a completed data collection or a cumulative data set still growing?
- At what point in the research process are the data to be published?
- What is the motivation for publishing the data?
- Are raw data or edited data going to be published?
- Should the data be subject to a peer review process?
- Is it sufficient to publish a single data set to meet the various requirements of: own publication, long-term archiving, requirements of funding agencies, own institution, ...?

It is not always clear in advance which data will be particularly valuable for subsequent use. Later studies may examine data sets or evaluate metadata⁶⁸ under completely unanticipated aspects⁶⁹. It is therefore recommended that research data be published and well documented, even if their value or tangible benefit is not clearly evident at this stage.

3. Methods for publishing data

Research data can be published in different ways. The choice of publication channel depends on the type of research and the content of the data. The most common is the option (I) of publishing aggregated data as a supplement to the academic article via the publisher. More recently, the options (II) of publishing the data in a repository as independent information objects and (III) of publishing the data description in so-called data journals have been extended. These are journals that specialize in reporting on published/accessible data. Those three ways of publishing data are described in more detail below.

(I) Data may be published by publishers as supplements to publications of research results in scientific articles. These data support and clarify the research results presented in the article. In most cases, they are aggregated data, such as smaller tables or images.

(II) Data may be published as an independent information object in a repository. As described below, there are different types of repositories. In discipline-specific repositories, it is easier for the community to find the data. In addition, they can be better contextualized or linked to other data sets. Discipline-specific repositories also offer corresponding features such as search, analysis and visualization. In cross-disciplinary and especially in institutional repositories, the data are less easily found.

(III) Data journals are dedicated to the publication of information about data published in open or restricted repositories. The information is a detailed documentation of published data, their properties and details of potential subsequent use. The data in the repository and its documentation in the journal are linked to each other by means of a persistent identifier (see below) and can thus be clearly located. Some of these journals offer a peer-review procedure in which the data set and its documentation are reviewed. For example, it is checked whether the data and its documentation match, whether the documentation sufficiently explains the data, what value the data has and whether the file formats are standardised. Examples of such data journals are the Open Access journal "Earth System Science Data"⁷⁰ in the geosciences or the interdisciplinary "Data in Brief".⁷¹

⁶⁷ Based on Martin, Elaine R. (ed.). *New England Collaborative Data management Curriculum. Module 6. Data Sharing & Reuse Politics*. Last accessed 03.11.2020.

https://library.umassmed.edu/docs/necdmc_module6.docx.

⁶⁸ McKiernan, Erin C., Philip E. Bourne, C. Titus Brown, Stuart Buck, Amye Kenall, Jenifer Lin, Damon McDougall et al.: Point of view: How open science helps researchers succeed. *eLife* 5 (2016), p. e16800. <https://doi.org/10.7554/eLife.16800>.

⁶⁹ Steiner, Daniel, Heinz J. Zumbühl and Andreas Bauder: „Two Alpine Glaciers over the Past Two Centuries.” In *Darkening Peaks: Glacier Retreat, Science, and Society*. Ed. by Ben Orlove, pp. 83–99. Berkeley, CA: University of California Press, 2008.

⁷⁰ Copernicus GmbH. “Earth System Science Data.” Last accessed 18.02.2019. <https://www.earth-system-science-data.net>.

⁷¹ Elsevier. „Data in Brief.“ Last accessed 18.02.2019. <https://www.journals.elsevier.com/data-in-brief>.

4. Repositories

Repositories are used to archive, document and publish digital objects. They are storage locations for data, which enable the publication and archiving of data independent of the scholarly article itself in terms of time and space.

Depending on the repository, data, data sets, descriptions of experiments and evaluations, audio-visual objects such as image and video files, models of simulations and also software can be published. In some repositories, entire research data sets can be stored in their complex form as a single unit (e.g. "collection" in PANGAEA).

5. Types of repositories

Repositories can be classified according to various aspects. Most often, they are distinguished by whether they are discipline-specific, cross-discipline/generic or institutional.

Discipline-specific or disciplinary repositories offer the advantage of visibility in the research community and are already established institutions in some disciplines. However, not all academic subject areas have yet established discipline-specific repositories or many are still in the development or project stage.

These repositories usually offer discipline-specific knowledge and know-how in the field of research data management, e.g. with regard to the curation of data or to special services (e.g. visualisation tools), but also specific tools for the analysis, research and visualisation of data.

Examples of discipline-specific repositories:

- Dariah-DE repository (<https://de.dariah.eu/repository>) for the humanities and cultural sciences
- SowiDataNet (<http://sowidatanet.de>) by GESIS for data from the social sciences and economics
- GenBank (<http://www.ncbi.nlm.nih.gov/genbank>) for genetic research data
- Integrated Ocean Observing System (<https://ioos.noaa.gov>) for data from marine environmental research
- NoMaD (<http://nomad-repository.eu>) for data on materials in the field of physics
- PANGAEA (<https://pangaea.de>) for earth science data
- TextGrid repository (<https://textgridrep.org/>) for the humanities (XML/TEI-based).

For interdisciplinary research, the assignment of the resulting data to a subject area may be difficult. Cross-disciplinary repositories offer a solution here. They generally accept very different types of data and provide a good search function. In most cases, however, they do not curate the data or offer other forms of quality control.

Examples of cross-disciplinary, generic repositories:

- ZENODO (<https://zenodo.org>): Digital data from all research areas, EU OpenAIRE⁷² project
- DRYAD (<http://datadryad.org>): Focus on life sciences, not free of charge
- Figshare (<https://figshare.com>): Digital data from all research areas, commercial data service.

Institutional repositories are also emerging more and more. Currently, they offer an alternative if no suitable discipline-specific repository is available. Researchers are happy to take up this offer. For example, if the legal framework for handling data at an external repository location is ambiguous for researchers, they may prefer to publish in their own institution's repository. Institutional repositories are generally available and can be used free of charge for all of the institution's own subject areas.

Some examples of institutional repositories at universities:

⁷² OpenAIRE is an international project with experts from research libraries, national e-Infrastructure and data experts, IT, legal scholars and open scholarship organizations. This project aims to significantly improve the searchability and reusability of research publications and data.

- Technical University of Berlin (<https://depositonce.tu-berlin.de/>)
- Humboldt-Universität zu Berlin (<https://edoc.hu-berlin.de>)
- Free University of Berlin (<https://refubium.fu-berlin.de>)
- University of Bielefeld (<https://pub.uni-bielefeld.de>)
- Heidelberg University ([https://heidata.uni-heidelberg.de/](https://heidata.uni-heidelberg.de))

6. Selection of a repository

In order to find an appropriate repository, the cross-disciplinary directory re3data (<https://www.re3data.org>) can be used. This is a DFG-funded project that lists German and international repositories for research data, with more than 2.580 entries at present.⁷³. Here you may select the discipline, type of data or country. It is also possible to filter by very detailed criteria, for example for repositories that charge a fee for data upload or where data use is restricted.

Based on OpenAIRE, the following recommendations for the procedure when selecting a repository for data publication are given in descending order of priority:⁷⁴

1. choose an external discipline-specific repository that is recognised in the discipline
2. find a suitable repository via re3data.org
3. select an institutional repository, or
4. use a cost-free multidisciplinary repository

Specific criteria for selecting a suitable repository:⁷⁵

1. certification, e.g. Core Trust Seal
2. (automated) assignment of persistent identifiers, e.g. DOI, handle
3. access to data: open, restricted or inaccessible
4. clear terms of use for data authors and users, e.g. fees, embargo periods

The options for choosing a license may also influence the choice of a suitable repository.

7. Licences

As part of the publication process, it is decided under which license the data is released. This decision regulates their use by third parties. Widely used are the free licenses of Creative Commons (CC).⁷⁶

A distinction is drawn between the following:

- CC0 (Public Domain)
- CC BY (Attribution)
- CC BY-ND (Attribution – No derivative works)
- CC BY-NC (Attribution – Non-commercial)
- CC BY-SA (Attribution – Share-alike: Distribution under Equal Terms)
- CC BY-NC-SA (Attribution – Non-commercial - Share-alike: Distribution under Equal Terms)
- CC BY-NC-ND (Attribution – Non-commercial - No derivative works)

The granting of a Creative Commons license has no influence on the copyright. However, it gives you the opportunity to grant rights of use for research data in a simple and standardized way. When granting

⁷³ re3data.org. Last accessed 02.11.2020. <https://www.re3data.org>.

⁷⁴ OpenAIRE. "How to select a repository?" Last accessed 02.11.2020. <https://www.openaire.eu/opendatapilot-repository-guide>.

⁷⁵ re3data.org. „FAQ“ Last accessed 02.11.2020. <https://www.re3data.org/faq>.

⁷⁶ Creative Commons Corporation. "Creative Commons." Last accessed 29.10.2020. <https://creativecommons.org>.

CC0, the author(s) renounce all copyright and related rights, whereas all other CC licenses may grant or restrict the rights of use to different extents.

Software as a research date requires a separate license to meet the special requirements of this format (agreement to install, modify, execute, purpose or location of use, number of users, etc.). The Creative Commons licenses do not cover this. It is recommended to use one of the common software licenses, such as the MIT license, GNU General Public License (GPL), GNU Lesser General Public License (LGPL) or the Apache license.

It is recommended to restrict the use of the research data as little as possible. This facilitates the subsequent use by third parties. If no license is granted, subsequent use is not permitted without the consent of the copyright holder.

8. Persistent Identifiers

Persistent identifiers (PIPs) are used to make digital publications findable in the long term and to solve the problem of "dead" links, as well as to improve the documentation of research data, especially their machine readability. Every object is identified by a unique name. This name is then included wherever reference is made to this object, i.e. the resource is linked to the identifier, not to a specific location. The persistence of PIPs is not guaranteed by technical means, but by contractual regulations.

There are different types of persistent identifiers, because potentially everything that is distinguishable and nameable can be provided with a persistent identifier. Two persistent identifiers are presented below: Digital Object Identifier (DOI) for data identification and Open Researcher and Contributor ID (ORCID) for unique identification of researchers. Other persistent identifiers used for scholarly work are, for example, Uniform Resource Name (URN), which, however, is not used worldwide but only regionally, the "International Geo Sampling Number" for geological samples, which is assigned via the System for Earth Sample Registration SESAR, or the identifiers of the Research Organization Registry (ROR)⁷⁷ for research institutions. The allocation of persistent identifiers may involve costs that should be priced into data management.

a) Digital Object Identifier (DOI)

The Digital Object Identifier (DOI) is very common because it is citable and its allocation has been free of charge in Germany since 2013. The International DOI Foundation (IDF) ensures uniform standards and workflows for the use of DOIs⁷⁸ and there has been an ISO standard for this purpose since May 2012. DOIs are unique sequences of alphanumeric characters. Permitted characters are:

a-z, A-Z, 0-9, . (dot), - (hyphen), _ (underscore), : (colon) and / (slash). Each DOI consists of two parts, a prefix that identifies the issuing organization and a suffix that identifies the object.

Examples for the structure of the suffix:

- Original DOI: 10.1234/abc123
- DOI of a new version: 10.1234/abc123.1
- DOI of a part: 10.1234/abc123/2

The DOI allows a distinct linking between the data and the resulting publications. Thus, the data remains permanently citable. DOIs are often assigned by repositories or institutions such as libraries. DataCite⁷⁹ distributes so-called number ranges to these institutions (registrar), which then assign them individually (registrant).

⁷⁷ Research Organization Registry: „About.” Last accessed 29.07.2020, <https://ror.org/about/>.

⁷⁸ Paskin, Norman. Digital object identifier (DOI®) system In *Encyclopedia of library and information sciences*. Ed. by Marcia J. Bates and Mary Niles Maack. 3. ed., pp. 1586–1592. Boca Raton, FL: CRC Press, 2010. <https://doi.org/10.1081/E-ELIS3-120044418>.

⁷⁹ DataCite. Last accessed 02.11.2020. <https://datacite.org/>.

In contrast to DOIs, handles as a global reference system for large amounts of data (and the basic technology for DOIs) are less persistent and not secured by a standard. They are therefore particularly suitable for referencing data before publishing.

b) Open Researcher and Contributor ID (ORCID)

To ensure all scientific papers are clearly assigned to the author, the so-called ORCID is used. Since researchers usually work at different institutions in the course of their academic career, their contact details may change. Sometimes the names of researchers also change in the course of their professional life. In order to be able to easily allocate all publications over time and with the change of institutions and names, researchers have the option of registering with ORCID (<https://orcid.org/>). Furthermore, by registering with ORCID, they can avoid having to enter the same personal data over and over again, for example when submitting data or articles for publication. Above all, this prevents confusion when identical names occur within the same discipline.

What you should know about ORCID:⁸⁰

1. stands for Open Researcher and Contributor ID
2. 16-digit (alpha) numeric code
3. protects your unique scholarly identity (also across name changes, typing errors or identical names)
4. is used by journals, funding agencies and institutions
5. is maintained by researchers themselves
6. lasts longer than an e-mail address
7. ORCID creation takes about 30 seconds
8. is run by a non-profit initiative
9. continuous growth (November 2020: 9.931.141 ORCIDs)
10. links to Web of Science, Scopus, Zenodo, DataCite, etc.

Didactic methods and exercises

Methods:

- Brainstorming
 - Which publication channels for research data can you think of?
- Statement slam
 - Persistent identifiers such as ORCID take time to set up and are not very useful afterwards.
 - I will publish my data so that my article will be cited more often.
 - Research is funded publicly to a large extent, so the resulting data is also a public good.
 - The reuse of data does not save costs, since research data management also generates large costs.
 - Of course, I will always collect my own data: I will not adapt my questions to existing data.
 - The re-use of data requires more knowledge than the collection of new data.
 - The re-use of my data could result in exciting new collaborations.
 - If I publish my data, my research becomes completely transparent and even the slightest errors are revealed.

⁸⁰ Cf. "Ten things you need to know about ORCID right now", <https://orcid.org/blog/2014/04/29/tenthingsyouneedtoknow>, last accessed 29.10.2020.

- The publication of research data does not contribute to enhancing my reputation.
- When I publish my research data, someone might pre-empt me and publish findings based on my data.
- Research data are a resource whose preservation and safeguarding represents a value for the future.
- The management and publication of research data generates costs that I cannot bear.
- Published data does not provide any further benefit.
- My research data belongs to me!
- Estimation question
 - How many researchers are registered with ORCID right now?
- Flip and turn
 - Combine CC pictograms with their meaning and set up any possible alternative linkages.
- Ideas out loud
 - What did I learn about data publishing that is relevant to my work?

Exercise:

- Find a discipline-specific repository in re3data.
- Register your own ORCID (or update your own ORCID profile if it already exists)
- Consider and write down answers to these questions: In your last article (published/read), were there any additional materials provided (tables, graphs, etc.)? What kind of data would you need to be able to reproduce a study?
- Constraints in science: Obstacles to the publication of data and possible solutions
 - List five reasons why some data cannot be published and shared with others
 - Consider also whether there are possibilities to reduce or remove restrictions

Training materials

- Template: Statement slam
- Poster: How to publish research data?
- Teaching script: Data publication

Additional sources

- Data Archiving and Networked Services: „The PID Forum.“ Last accessed 20.07.2020, <https://www.pidforum.org>.
- DataCite. Last accessed 29.07.2020, <https://datacite.org>.
- Ebel, Thomas und Alexia Meyermann. „Hinweise zur Anonymisierung von quantitativen Daten. Version 1.0.“ *Forschungsdaten Bildung informiert* Nr. 3 (2015). Frankfurt a.M.: Forschungsdatenzentrum Bildung am DIPF, 2015. Last accessed 29.07.2020, https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-3.pdf.
- Forschungsdaten.info: “Persistente Identifikatoren.“ Last accessed 20.07.2020. <https://www.forschungsdaten.info/themen/veroeffentlichen-und-archivieren/persistente-identifikatoren/>.
- Henneken, Edwin A. and Alberto Accomazzi. *Linking to data-effect on citation rates in astronomy*. arXiv: 2011, <https://arxiv.org/abs/1111.3618>.
- International DOI Foundation. „DOI.“ Last accessed 29.07.2020, <https://www.doi.org>.
- McKiernan, Erin C., Philip E. Bourne, C. Titus Brown, Stuart Buck, Amye Kenall, Jenifer Lin, Damon McDougall et al.: Point of view: How open science helps researchers succeed. *eLife* 5 (2016). <https://doi.org/10.7554/eLife.16800>.

- Meyermann, Alexia and Maike Porzelt: „Hinweise zur Anonymisierung von qualitativen Daten. Version 1.0.“ *Forschungsdaten Bildung informiert* Nr. 1 (2014). Frankfurt a.M.: Forschungsdatenzentrum Bildung am DIPF, 2014. Last accessed 29.07.2020. https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-1.pdf.
- ORCID Inc. „ORCID.“ Last accessed 29.07.2020, <https://orcid.org>.
- Paskin, Norman: "Digital object identifier (DOI®) system." In *Encyclopedia of library and information sciences*, ed. by Marcia J. Bates and Mary Niles Maack. 3rd ed. Boca Raton, FL: CRC Press, 2010. <https://doi.org/10.1081/E-ELIS3-120044418>.
- Pienta, Amy M., George C. Alter and Jared A. Lyle: *The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data*. Ann Arbor, Mi: University of Michigan – Deep Blue, 2010. <http://hdl.handle.net/2027.42/78307>.
- Piwowar, Heather A and Todd J. Vision: Data reuse and the open data citation advantage. *PeerJ* 1 (2013), p. e175. <https://doi.org/10.7717/peerj.175>.
- Piwowar, Heather A., Roger S. Day and Douglas B. Fridsma: Sharing detailed research data is associated with increased citation rate. *PLoS ONE* 2 (2007). <https://doi.org/10.1371/journal.pone.0000308>.
- Poisot, Timothée E., Ross Mounce and Dominique Gravel: Moving toward a sustainable ecological science: don't let data go to waste! *Ideas in Ecology and Evolution* 6 (2013). <https://doi.org/10.4033/iee.2013.6b.14.f>.
- Project FREYA: "The FREYA project." Last accessed 20.07.2020, <https://www.project-freya.eu/en/about/mission>.
- Steiner, Daniel, Heinz J. Zumbühl and Andreas Bauder: "Two Alpine Glaciers over the Past Two Centuries." In *Darkening Peaks: Glacier Retreat, Science, and Society*. Ed. by Ben Orlove. Berkeley, CA: University of California Press, 2008.
- Verbund FDB: „Informierte Einwilligung.“ Last accessed 29.07.2020, <https://www.forschungsdaten-bildung.de/einwilligung>.
- Vierkant, Paul, Daniel Beucke, Gernot Deinzer, Sarah Hartmann, Sebastian Herwig, Kathrin Höhner, Uwe Müller et al.: *Autorenidentifikation anhand der Open Researcher and Contributor ID (ORCID) - Positionspapier*. Berlin: Humboldt-Universität zu Berlin – edocServer, 2018. <https://doi.org/10.18452/19528>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al.: "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3 (2016). <https://doi.org/10.1038/sdata.2016.18>.



Template: Statement slam

Persistent identifiers such as ORCID take time to create and are not very useful afterwards.

I will publish my data so that my article will be cited more often.

Research is funded publicly to a large extent, so the resulting data is also a public good.

The re-use of data does not save costs, since research data management also generates large costs.

Of course, I will always collect my own data: I will not adapt my questions to existing data.

The re-use of data requires more knowledge than the collection of new data.

The re-use of my data could result in exciting new collaborations.

If I publish my data, my research becomes completely transparent and even the slightest errors are revealed.



The publication of research data does not contribute to enhancing my reputation.

When I publish my research data, someone might pre-empt me and publish findings based on my data.

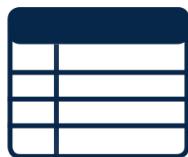
Research data are a resource whose preservation and safeguarding represents a value for the future.

The management and publication of research data generates costs that I cannot bear.

Published data does not provide any further benefit.

My research data belongs to me!

HOW TO PUBLISH RESEARCH DATA?



DOCUMENT THE DATA

Always document your data from the beginning of the research process in order to make the research data reproducible. Provide relevant metadata as well. Stick to discipline-specific metadata standards.

Further information:
<https://tinyurl.com/RDdocu>



CHOOSE A REPOSITORY

Search for a suitable subject-specific repository relevant for your research community. If you do not find what you are looking for choose a multidisciplinary or institutional repository.

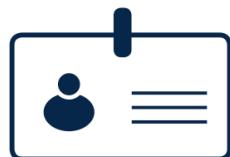
Further information:
www.re3data.org



GRANT A LICENSE

Choose a license suitable for your research data (e.g. Creative Commons). Try to keep the terms of use as open as possible and as restricted as necessary.

Further information:
<https://creativecommons.org>
<https://choosealicense.com>



PERSISTENT IDENTIFIER

Make sure that your data gets a DOI to make it findable in the long term. Create an ORCID for yourself so that your scientific work can be clearly assigned to you.

Further information:
www.doi.org
<https://orcid.org>



LEGAL ASPECTS

The publication of research data may be opposed by various legal and/or ethical aspects. Check this before publication.

Further information:
<https://tinyurl.com/RDlegal>



PUBLISHING

Upload your research data in a suitable file format to the selected repository and let the world know about it!

If you have any questions, the staff of the repository will be happy to help you.

Further information:
<https://tinyurl.com/RDfileformats>

15. Teaching script: Data publication							
Component	Target	Time	Content	Working modes	Material	Inhalation/ exhalation	Voices sound
						Alternative	Notes
Orientation	The participants get orientation	1	The content is presented	Presentation	PPTX	In	-
Ways of publication	The participants get to know different ways of publication	5	(I) as a supplement to the scientific article (II) in a repository as independent information objects (III) Data description in a so-called data journal	Presentation	PPTX	In	-
	Participants learn how to find repositories using concrete examples	10	Exercise for finding repositories (re3data)	Individual work on a notebook	Notebooks/ Tablets/ Smartphone (info for participants in invitation)	Ex	-
	Participants describe their experience of searching for a repository	3	Discussion of the results	Exchange in the whole group		Ex	Yes
Research a repository	The participants develop criteria for the selection of a repository	5	The participants answer the question: 'What criteria do you have in mind when selecting a suitable repository?'	Ideas out loud	Flip chart, pen	Ex	S: present prepared flip chart
Select data for publication	The participants learn criteria for the selection of data for publication	2	Criteria are presented	Presentation	PPTX	In	-

Licences	The participants get to know CC licences	Trainer introduces CC licences and their combinations. The participants answer the question: What happens if I want to use an image with CC-SA in a CC-BY publication?	Presentation and discussion	PPTX	In and ex	-	
	The participants learn about different persistent identifiers	Various persistent identifiers are presented and their benefits are illustrated	Presentation	PPT	In	-	Motivate the participants to create an ORCID
Persistent Identifier	ORCID	1	The participants answer the question: How many researchers were registered with ORCID in November 2020? (9.931.141)	Expectation query	Current number	Ex Yes	How many people in the room have an ORCID?
Coffee break							A small prize (e.g. a small pack of gummy bears) increases the motivation to participate
Pros and cons of the publication	The participants will discuss the advantages and disadvantages of data publication	Each participant draws a sheet of paper with a statement and has 3 minutes preparation time. Afterwards the participant takes a stand on the statement in a 60 second presentation. A discussion can follow. 1 min introduction 4 min preparation 20 min lectures 20 min discussion	Method: Statement	Theses	Ex	Yes	Shorter with less participants
Reflection	Participants make a connection between learning content and their everyday life	What have I learnt about publishing research data that is directly relevant to my regular work?	Individual work	Harvest sheet	Ex	-	

Duration of the unit: 91 minutes



Unit 16: Re-use of research data

Learning objectives

The Participants will get to know different sources of information for the discovery of research data.

Participants will learn how to find data using specific examples and how to evaluate the found data in terms of its reusability for their research.

Participants will learn which elements belong to the common standards for data citation.

Key aspects

1. Why reuse data?
2. Searching for data
3. Access and terms of use
4. Key questions for the evaluation of reusability
5. Citation

Contents

1. Why reuse data?

The benefits of data reuse emerge at several levels,⁸¹ which are briefly discussed in detail here: the researcher – both reusing and publishing – the discipline and the societal level:

- for researchers who reuse data
 - less effort and costs, since no own data collection
 - secondary analyses on new research questions and/or with new methods

⁸¹ Cousijn, Helena, Amye Kenall, Emma Ganley, Melissa Harrison, David Kernohan, Thomas Lemberger, Fiona Murphy et al.: "A data citation roadmap for scientific publishers." *Scientific data*, 5, 180259 (2018). <https://doi.org/10.1038/sdata.2018.259>.

- comparisons over time
- comparisons of different samples
- links to other sources
- new collaborations
- for researchers who publish data
 - citations
 - transparency
 - enhancing the scholarly reputation
 - new collaborations
- for the discipline
 - reproducibility of the research
 - more efficient research
 - enabling new research
 - preservation and safeguarding of data for the future, especially relevant for non-repeatable data collections (historically unique phenomena) and for data from vulnerable groups that are difficult to access
 - applicability in teaching
- for the public benefit
 - transparency of research
 - trust in science
 - protection of population and environment through less frequent primary data gathering
 - economic exploitation also by the private sector

Researchers can spare the effort of their own data collection by using already existing data sets. By reusing them, they expand their own research base. New collaborations may also result from this. However, the reuse of research data involves the effort to read and understand the data.

2. Searching for data

Research data is currently still difficult to find. Many directories and (meta) search engines are under construction and they vary greatly in volume, quality and reputation. Finding suitable data for reuse usually requires a search in different sources. The following are some examples:

- Direct browsing of discipline-specific and multidisciplinary repositories (cf. unit 15 point 5, p. 110)
- Search by means of meta search engines⁸², e.g.
 - B2FIND data search (<http://b2find.eudat.eu>)
 - BASE (<https://www.base-search.net>)
 - DataCite Metadata Search (<https://search.datacite.org>)
 - Europäisches Datenportal (<https://www.europeandataportal.eu/de>)
 - Google Dataset Search (<https://datasetsearch.research.google.com/>)
- Search in data journals, e.g. "Earth System Science Data"⁸³ in geosciences or the interdisciplinary "Data in Brief".⁸⁴

Other portals can also be found on the information website forschungsdaten.info.⁸⁵

⁸² Meta search engine – a search engine that submits a query to several search engines at the same time and displays the aggregated results.

⁸³ "Earth System Science Data." Last accessed 19.02.2019. <https://www.earth-system-science-data.net>.

⁸⁴ Elsevier. „Data in Brief.“ Last accessed 19.02.2019. <https://www.journals.elsevier.com/data-in-brief>.

⁸⁵ Forschungsdaten.info. „Forschungsdaten finden.“ Last accessed 09.07.2020, <https://www.forschungsdaten.info/themen/finden-und-nachnutzen/forschungsdaten-finden/>.

3. Access and terms of use

Access to research data is either open (unrestricted), conditional or restricted (only for specific purposes/research, for specific groups of people, users must provide information or meet other formal requirements) or even completely inaccessible. Sometimes fees are charged for providing data (e.g. for sending a DVD with the data). The terms of use are determined by the repository's own terms of use on the one hand, and by those that are specified by the data authors, such as the licences used (cf. unit 15 point 6, p. 111), on the other. If this information is not evident from the metadata, it should be clarified or negotiated in the context of the request for reuse.

4. Key questions for the evaluation of reusability

Once the legal reusability has been checked, the content should be examined. The metadata usually provide only basic information as to whether this data set could be suitable for the intended purpose. If the data set appears to be suitable for one's own work at first glance, the suitability must be examined more closely. You can picture this examination similar to a more in-depth study of a scholarly article, in which the details of the collection, evaluation and interpretation are checked and assessed very carefully. Therefore, it is recommended to invest some time in order to review the following questions:

- Is the specific research question well-documented?
- How was the data collected?
- Are the collection and processing methods used appropriate to the research question and do they correspond to the current state-of-the-art in my field of expertise?
- Is the procedure of data collection accurately recorded and comprehensibly documented?
- Which collection instruments were used for collecting the data? And what settings or parameters?
- Are reports and protocols of the data gathering as well as their specifics included in the data set?
- Is the description of the data set available and sufficient to understand the data and its context of origin?
- Which criteria for data selection were applied?
- Has the data been processed since the data collection? If so, how, e.g. handling of missing values? Weighting?
- Are precise descriptions of the variables available, e.g. what variables are there, how are they coded, etc.?
- Is all the information understandable and consistent?
- Is the source trustworthy?

Only if these questions can be answered sufficiently is it possible to determine whether the data set is suitable for subsequent use.

5. Citation

The citation of data sets serves several purposes. First of all, it acknowledges the author's performance in producing the data and creating the data set.^{86⁸⁷} Furthermore, it ensures transparency in academic research and corresponds to good scientific practice. At the same time, the citing of data provides the basis for further reuse: other researchers learn which data have been used and where they can be found. Citations allow the author to find out what influence his or her work has and for what purposes the data is reused. Reuse without citation would be plagiarism.

⁸⁶ DataCite. „Cite Your Data.“ Last accessed 19.02.2019. <https://datacite.org/cite-your-data.html>.

⁸⁷ Cousijn, Helena, Amye Kenall, Emma Ganley, Melissa Harrison, David Kernohan, Thomas Lemberger, Fiona Murphy et al.: “A data citation roadmap for scientific publishers.” *Scientific data*, 5, 180259 (2018). <https://doi.org/10.1038/sdata.2018.259>.

In 2014, FORCE11 defined data citation principles⁸⁸ that cover the purpose, function and attributes of citations. These principles recognize the dual need to create citation practices that are both machine-readable and comprehensible to humans. The principles are grouped according to facilitating comprehension and are not ranked by importance:

1. importance
2. credit and attribution
3. evidence
4. unique identification
5. access
6. persistence
7. specificity and verifiability
8. interoperability and flexibility

Some subject areas already have their own recommendations (e.g. psychology from APA⁸⁹). In general, citing data should be similar to citing a research article. Among the common standard information are:

- Author
- Publication date
- Title
- Publisher (name of the data centre/institution that published the source)
- Resource type (for example, data set)
- Persistent identifier
- if applicable, Version number, in case of several published versions

Examples:

- Markowski, Radoslaw; Gebethner, Stanislaw; Grabowska, Miroslawa; Grzelak, Paweł; Jasiewicz, Krysztof et. al. (2006): Polish National Election Study 2000 (PGSW). Version: 1.0.0. GESIS Data Archive. Data set. doi.org/10.4232/1.4334
- U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2013). Treatment episode data set -- discharges (TEDS-D) -- concatenated, 2006 to 2009 [Data set]. doi:10.3886/ICPSR30122.v2

For the reuse of works with a Creative Commons license, you should also make sure the license is correctly attributed. This particularly means the source of the work should include the following information:

- Source
- Name of the license including the version and link to the license description
- if applicable, processing information (since version 4.0)
- If applicable, title of the work (since version 4.0)

Didactic methods and exercises

Methods:

- Chatter
 - How do I want my records to be cited?

⁸⁸ Data Citation Synthesis Group, M. Martone (ed.): *Joint Declaration of Data Citation Principles*. San Diego, CA: FORCE11, 2014. <https://doi.org/10.25490/a97f-egyk>.

⁸⁹ American Psychological Association: "About APA." Last accessed 27.08.2020. <https://www.apa.org/about>.

Exercises:

- Find a data set in the discipline-specific repository.

Training materials

- Worksheet: Citation
- Teaching script: Re-use of research data

Additional sources

- Ball, Alex and Monica Duke: *How to Cite Datasets and Link to Publications*. Edinburgh: Digital Curation Centre, 2015. Last accessed 29.07.2020, <http://www.dcc.ac.uk/resources/how-guides/cite-datasets>.
- Cousijn, Helena, Amye Kenall, Emma Ganley, Melissa Harrison, David Kernohan, Thomas Lemberger, Fiona Murphy et al.: "A data citation roadmap for scientific publishers." *Scientific data*, 5 (2018). <https://doi.org/10.1038/sdata.2018.259>.
- Data Citation Synthesis Group, M. Martone (ed.) "Joint Declaration of Data Citation Principles." San Diego, CA: FORCE11, 2014. <https://doi.org/10.25490/a97f-egyk>.
- Forschungsdaten.info: „Das Zitieren von Daten. Prinzipien und Varianten.“ Last accessed 29.07.2020, <https://www.forschungsdaten.info/themen/bewahren-und-nachnutzen/zitieren-von-daten/>.
- Pampel, Heinz: *Auf dem Weg zu Open Science – Strategien und Erfahrungen am GFZ*. Präsentation. Potsdam: GFZ - GFZpublic, 2016. Last accessed 02.11.2020. <http://gfzpublic.gfz-potsdam.de/pubman/item/escidoc:1453909>.



Worksheet: Citation

Cite the following datasets.

Example 1

Creator: Risan, Patrick

Title: Accommodating Trauma in Police Interviews. An Exploration of Rapport in Investigative Interviews of Traumatized Victims, 2013

Resource type: Data set

Publisher: NSD – Norwegian Centre for Research Data

DOI: 10.18712/NSD-NSD2407-V1

Date: 2017

Citation:

.....
.....
.....
.....

Example 2

Creator: Parton, William J.; Gutmann, Myron P.; Hartman, Melannie D.; Merchant, Emily R.; Lutz, Susan M.

Title: Great Plains Population and Environment Data: Biogeochemical Modeling Data, 1860-2003 [United States]

Resource type: Data set

Publisher: ICPSR – Interuniversity Consortium for Political and Social Research

Date: 04.10.2012

DOI: 10.3886/ICPSR31681.v1

Citation:

.....
.....
.....
.....



16. Teaching script: Re-use of research data							
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound
						Alternative	Notes
Research	The participants learn about different resources for searching research data	5	Introduction and presentation of research possibilities	Presentation	PPTX	In	-
	The participants learn to search for data using a concrete example	10	The participants search and find data in the repository	Practice on the Internet	Notebooks/ Tablets/ Smartphone	Ex	-
	The participants compare their experiences	5	The participants exchange information	Ideas out loud		Ex	Yes
	The participants activate their previous knowledge	2	The participants answer the question: How do I want my data sets to be cited?	Method: Chatter		Ex	Yes
Citation	The participants learn which information is part of the common standards for data citation	3	Presentation of two examples	Presentation	PPTX	In	-
	The participants can apply the newly acquired knowledge	5	The participants receive a worksheet and write down citations	Exercise	Worksheet: Re-use	Ex	-

Duration of the unit: 30 minutes



Unit 17: Legal aspects

Learning objectives

Participants learn about the different levels of legal regulations that could be relevant to research data management.

Participants will understand how they are allowed to publish personal data with the help of informed consent, anonymisation and access control.

Key aspects

1. Levels of regulation
2. Data and privacy protection
3. Act on Copyright and Related Rights
4. Contract law

Contents

1. Levels of regulation

Depending on the nature of the data, legal and ethical frameworks, which operate at very different levels, must be considered closely in research data management (see Figure 6). For example, when it comes to personal data, the European General Data Protection Regulation GDPR, the German federal data protection act called *Bundesdatenschutzgesetz* (BDSG), the *Landesdatenschutzgesetze*, which are the federal data protection acts of the 16 German states, and, if applicable, their respective state higher education acts apply, as well as, depending on the specific case, further data protection regulations, e.g. in the Social Code, Genetics Law, Pharmaceuticals Law and School Law. However, there may also be guidelines, policies or even specific ethical guidelines at institutional level. Basically, researchers are subject to the "Guidelines for Safeguarding Good Research Practice"⁹⁰ of the German Research Foundation.

⁹⁰ DFG, "Guidelines for Safeguarding Good Research Practice. Code of Conduct." Zenodo 2019, <http://doi.org/10.5281/zenodo.3923602>.

Patent law What has to be considered when research data (can) reach patentability?	Copyright law Is research data even subject to copyright law?	Competition law Is data used unfairly in business transactions?	Data protection Which research data is sensible and needs special protection?
Science law Can licensing and publication requirements for research data be mandated?	Fundamental rights Which constitutional constraints have to be considered?	International law What legal regulations exist outside Germany?	EU law What are the benefits of e.g. the "European Data Economy" for research data?
Contract law Are there agreements on the "intellectual property" of research data?	Labour/Service law Who "owns" the research data collected at universities?	Terms of funding What conditions are set by funding bodies (DFG; industry)?	Policies What legal obligations can policies develop?

Figure 6: Examples of the legal fields of action that may affect the publication of research data. According to Thomas Hartmann: „terra incognita – digitale Forschungsdaten auf der Suche nach einer rechtlichen Heimat“. Contribution to the workshop "Legal aspects of digital research data" on 31.01.2018, Europa-Universität Viadrina, Frankfurt (Oder) http://www.forschungsdaten.org/index.php?Datei:Hartmann_TerraIncognita-Forschungsdaten-RechtlicheHeimat.pdf.

In some cases, it is useful to restrict access to the data if ethical or legal aspects are involved. Furthermore, the research itself may be subject to confidentiality agreements, e.g. contract research.

2. Data and privacy protection

In some fields, data protection aspects can determine and limit the handling of data, for example in the social sciences and in medicine. This is always the case when processing personal data or data that could relate to individuals. Article 4 of the GDPR (DSGVO in German) defines personal data as "any information relating to an identified or identifiable natural person". Persons are considered identifiable if they can be recognised, directly or indirectly, in particular by means of association with an identifier such as a name, an identification number, location data, online identification or one or more characteristics which reveal the physical, psychological, genetic, mental, economic, cultural or social identity of these natural persons. Sensitive personal data are specific personal data which require enhanced protection. According to article 9 GDPR, this includes "data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership" and "genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation". Therefore, the personal details must be permanently removed from the data or the person in question must have given his or her "informed consent" to the processing and use of the data.

The EU General Data Protection Regulation (EU GDPR) which has been in effect since May 2018 largely harmonises European data protection law and aims to protect the fundamental rights and freedoms of natural persons, in particular their right to the protection of personal data. The EU GDPR generally prohibits the handling of personal data, unless another legal regulation or the consent of the person affected permits this. Such consent is an informed and unequivocally given consent in the form of a declaration or other unambiguous act.

With an informed consent, the person is informed about his/her rights, the processing of his/her data, its use and the purpose of the study. Only when informed does the person agree to participate in the study under these conditions. The information can also be given verbally (for example, in the case of children or illiterate persons). If the publication of research data is planned, this information will be

included in the disclosure. It is vital to include all planned purposes of use in the consent (e.g. long-term archiving, data publication, further analyses on other issues, etc.), as it is often not feasible to obtain the extension of consent afterwards.

In order to facilitate the handling of the data, there are various ways of changing the data so that they no longer contain any references to specific persons. For example, numerical data may be anonymised or aggregated. For this purpose, the personal information is either completely removed or replaced by other information that cannot be traced back to individual persons (e.g. zip code instead of street address). Video and audio recordings could be edited and also anonymised or depersonalised. Some examples are the modification (pixelation, black bar) of people or information as well as the revision of sound recordings in order to distort voices. Using pseudonyms, i.e. replacing a piece of information with a similar piece of information that is no longer person-related, is also a way of processing qualitative data without informed consent.

Further protective measures include access restrictions by means of password protection and the assignment of access rights. For example, access to the data can be granted only to certain IPs or only to defined groups of people (see unit 11: Access control, p. 88).

The rapid evolution of digital media today leads more than ever to careless publication of personal data – for example in the form of photographs or videos. The right to one's own image, however, includes the right of every individual to decide what happens to images or other forms of representation of oneself in public. These may only be distributed or published with a consent. A consent can be both: financial compensation or answering questions in front of a camera. Exceptions to this rule:

- contemporary historical images
- pictures in which the person is only depicted incidentally and not as the main motif
- pictures of meetings or similar

3. Act on Copyright and Related Rights

Copyright law aims to protect the creative output of creative people and the creative industries. It grants authors the sole right to publish, process, reproduce, perform and distribute their work. In order for a work to be considered "protected by copyright", it must be the result of a creative effort and originality. According to art. 2 paragraph 1 of the German Act on Copyright and Related Rights (UrhG)⁹¹ this includes "works of literature, science and art". Accordingly, ideas or information are not protected by copyright, since they only lay the foundation of a work, but are not the work itself.⁹² The copyright expires at the latest 70 years after the death of the author*in and the work is then considered to be in the public domain.

The so-called ancillary copyrights – also called "related rights" – are protective rights for e.g. photographers, singers, interpreters or mediators of creative contents. The ancillary copyright protects artistic or scholarly achievements and includes, among other things, the protection of database creators, the protection of photographs or the protection of the producer of audio recordings (further protection rights can be found in the German Act on Copyright, part 2 "Related rights"). The ancillary copyrights are entitled to the person who has carried out the respective protected work.

In addition to copyright aspects, research data management may also require commercial property rights to be considered. For example, employment contracts might have been formulated in such a way that the institution at which they work, rather than the researchers themselves, holds the legal rights of use for the data. Furthermore, contractual regulations on the rights of use are also to be found in the funding guidelines of sponsors or in the policies of groups, departments or institutions involved. Nonetheless, the copyright remains with the creators of the data. It is therefore advisable to check who

⁹¹ Cf. https://www.gesetze-im-internet.de/englisch_urhg/index.html, last accessed 29.10.2020.

⁹² Kreutzer, Till and Henning Lahmann. *Rechtsfragen bei Open Science: Ein Leitfaden*. Hamburg: Hamburg University Press, 2019, <https://doi.org/10.15460/HUP.195>, p. 25.

will be the rightsholder of the data before starting the research. Within the framework of cooperation agreements, these rights can be written down (see Contract law).

If the research or its results are likely to be commercially viable, it is advisable to contact the office for knowledge transfer of your own institution. They can clarify how the licensing and reuse of the data is regulated. They also provide detailed information on patent law.

4. Contract law

Further rights and obligations for the handling of research data may arise from existing contractual agreements. These might be contracts with funding agencies as well as with scientific or economic cooperation partners.

Didactic methods and exercises

Exercise:

- At which stages of the data management process are legal issues relevant?
- Where do you get advice on legal aspects of research data management?

Training materials

- Exercise: De-identification of qualitative data
- Answers: De-identification of qualitative data
- Checklist: Requirements for a consent according to GDPR
- Teaching script: Legal aspects

Additional sources

- Deutsche Forschungsgemeinschaft: "Guidelines for Safeguarding Good Research Practice. Code of Conduct." Zenodo 2019, <http://doi.org/10.5281/zenodo.3923602>.
- Deutsche Gesellschaft für Kartographie, Gesellschaft für Kartographie und Geomatik. *Urheberrecht leicht gemacht – Karten und Geodaten für Unterricht, Wissenschaft und Beruf nutzen und publizieren*. 01.10.2017. Oldenburg, DGfK, 2017. Last accessed 29.07.2020, http://www.dgfk.net/download/171026_Urheberrecht_4a.pdf.
- Guibault, Lucie and Andreas Wiebe (ed.). *Safe to be open: Study on the protection of research data and recommendations for access and usage*. Göttingen: Universitätsverlag Göttingen, 2013. <https://doi.org/10.17875/gup2013-160>.
- Hartmann, Thomas: „Zur urheberrechtlichen Schutzhfähigkeit von Forschungsdaten.“ In *Law as a Service (LaaS) - Recht im Internet- und Cloud-Zeitalter* herausgegeben von Jürgen Taeger. Band 1, S. 505–515. Edewecht: OIWIR, Oldenburger Verlag für Wirtschaft, Informatik und Recht, 2013. <http://hdl.handle.net/11858/00-001M-0000-0014-1208-E>.
- Helbig, Kerstin: Quiz FDM-Rechtsexpertise. Zenodo, 2019. <https://doi.org/10.5281/zenodo.3479148>.
- Klimpel, Paul and John H. Weitzmann: "Forschen in der digitalen Welt – Juristische Handreichung für die Geisteswissenschaften." *DARIAH-DE Working Papers* Nr. 12 (2015). <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-5-0>.
- Kreutzer, Till and Henning Lahmann. *Rechtsfragen bei Open Science. Ein Leitfaden*. Hamburg: Hamburg University Press, 2019. <https://doi.org/10.15460/HUP.195>.
- Lauber-Rönsberg, Anne, Philipp Krahn and Paul Baumann: *Kurzfassung: Gutachten zu den rechtlichen Rahmenbedingungen des Forschungsdatenmanagements im Rahmen des DataJus-Projekts*. Dresden: Technische Universität, 2018. Last accessed 29.07.2020. https://tu-dresden.de/gsw/jura/igetem/jfbimd13/ressourcen/dateien/dateien/DataJus/DataJus_Zusammenfassung_Gutachten_12-07-18.pdf?lang=de.

- Leibniz Universität Hannover und Technische Informationsbibliothek: *FAQs zu rechtlichen Aspekten im Umgang mit Forschungsdaten*. Zenodo, 2018, <http://doi.org/10.5281/zenodo.1173546>.
- Meyermann, Alexia and Maike Porzelt: „Datenschutzrechtliche Anforderungen in der empirischen Bildungsforschung – eine Handreichung. Version 1.0.“ *Forschungsdaten Bildung informiert* Nr. 6 (2017). Last accessed 29.07.2020, https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-6.pdf.
- Scheller, Jürgen: „Rechtliche Rahmenbedingungen der Verwendung von Videos in der Schul- und Unterrichtsforschung. Diskrepanzen zwischen Datenschutzrecht, Förder- und Genehmigungsauflagen. Version 1.0.“ *Forschungsdaten Bildung informiert* Nr. 5 (2017). Last accessed 29.07.2020, https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-5.pdf.
- Technische Universität Dresden: „Forschungsprojekt DataJus.“ Last accessed 29.07.2020, <https://tu-dresden.de/gsw/jura/igetem/jfbimd13/forschung/forschungsprojekt-datajus>.
- Verbund Forschungsdaten Bildung: „Informierte Einwilligung.“ Last accessed 29.07.2020, <https://www.forschungsdaten-bildung.de/einwilligung>.

Exercise: De-identification of qualitative data



1. In this example interview transcript, where would you have concerns for the risk to disclose the identity of the interviewee? What direct and indirect identifiable information do you note in the text that might concern you? Highlight any words, phrases or sections that you think need to dealt with.
2. How might you de-identify the text to reduce the risk to disclose the identity of the interviewee?

Source: Seymour, Jane (2010-2012). Managing suffering at the end of life: a study of continuous deep sedation until death. [Data Collection]. Colchester, Essex: Economic and Social Research Council. 10.5255/UKDA-SN-850749

Case and interview

Mr Tom Jeavons, aged 63, was suffering from metastatic cancer resulting from a primary site in the bladder. His wife, Sue (58), had been his main carer for many months as he struggled with severe pain, anxiety and other symptoms. Eventually, she received support from the hospice at home team, based at their nearby hospice – St Barbara. 11 days before his death, he was admitted to their inpatient unit, where he died. The case was identified by the staff there as a “critical case”, involving palliative sedation and the difficulties staff experienced in controlling his complex symptoms. Other interviews carried out were with the hospice consultant, Dr Jane O’Connor and three nurses: Elaine McDonald, Claire Smith and Mark Ferguson. Mr and Mrs Jeavons’ GP, Dr Paul Hyde, was also interviewed which added a different medical perspective, making this an unusual case.

Central themes in all of the interviews were his intractable and distressing symptoms and the repeated requests from Mr Jeavons for euthanasia. His wife mentions earlier discussions with Mr Jeavons about the possibility of going to a Dignitas clinic, but he was already too ill to travel. She also expresses how concerned she was about what Mr Jeavons’s adult children might witness when he was dying in the hospice.

INT: So, really, it's as I said to you: I want you to tell me what you can remember about Mr Jeavons' care in the last week of his life ... or about Mr Jeavons in the last week of his life.

RESP: Yeah, erm, 11 days, Tom was in St Barbara's Hospice for the last 11 days of his life so...

INT: So if you'd like to talk about that period...

RESP: Yeah.

INT: ...that'd be great.

RESP: Prior to him going in, and we was coping with his care at home, but then he was becoming less and less mobile: he couldn't go to the toilet; he had a frame, and everything that you added in that was, it was a step to help him but a downward step to the end of how he could cope. We had a Bariatric bed brought into the other room but he insisted in sleeping in his chair. We had St Barbara's here and, erm, the GP, and, er, we also had him assessed at home as to whether or not we could care for him completely at home. And Tom was about 20-something stone, so he wasn't easy to manoeuvre and, and the one thing that concerned me was the fact that, erm, they needed four people to move him, you know, if he wanted to go to the toilet or if he wanted to go on a bedpan or anything, and we had the bed in there – which he wouldn't sleep in. And, erm, basically the, logically trying to be able to do everything for him and keep him comfortable, we'd have to wait for an on-call four nurses – could be in the middle of the night – and, and sort of the idea of being able to cope, erm, for his safety and wellbeing was, was really compromised. He didn't want to go into St Barbara's, he didn't want to die in hospital, erm, but I just felt I had to take that



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Cite as: Katherine Powis (2018) Exercise: De-identification of qualitative data. UK Data Service, UK

UK Data Service



Answers: De-identification of qualitative data

The mark-up shows how the text was de-identified to reduce the risk to disclose the identity of the interviewee and other individuals mentioned in the text.

Mr Tom Jeavons, aged 63, was suffering from metastatic cancer resulting from a primary site in the bladder. His wife, Sue (58), had been his main carer for many months as he struggled with severe pain, anxiety and other symptoms. Eventually, she received support from the hospice at home team, based at their nearby hospice St Barbara. 11 days before his death, he was admitted to their inpatient unit, where he died. The case was identified by the staff there as a "critical case", involving palliative sedation and the difficulties staff experienced in controlling his complex symptoms. Other interviews carried out were with the hospice consultant, Dr Jane O'Connor and three nurses: Elaine McDonald, Claire Smith and Mark Ferguson. Mr and Mrs Jeavons' GP, Dr Paul Hyde, was also interviewed which added a different medical perspective, making this an unusual case.

Central themes in all of the interviews were his intractable and distressing symptoms and the repeated requests from Mr Jeavons for euthanasia. His wife mentions earlier discussions with Mr Jeavons about the possibility of going to a Dignitas clinic, but he was already too ill to travel. She also expresses how concerned she was about what Mr Jeavons's adult children might witness when he was dying in the hospice.

INT: So, really, it's as I said to you: I want you to tell me what you can remember about Mr Jeavons' care in the last week of his life ... or about Mr Jeavons in the last week of his life.

RESP: Yeah, erm, 11 days, Tom was in St Barbara's Hospice for the last 11 days of his life so...

INT: So if you'd like to talk about that period...

RESP: Yeah.

INT: ...that'd be great.

RESP: Prior to him going in, and we was coping with his care at home, but then he was becoming less and less mobile: he couldn't go to the toilet; he had a frame, and everything that you added in that was, it was a step to help him but a downward step to the end of how he could cope. We had a Bariatric bed brought into the other room but he insisted in sleeping in his chair. We had St Barbara's here and, erm, the GP, and, er, we also had him assessed at home as to whether or not we could care for him completely at home. And Tom was about 20-something stone, so he wasn't easy to manoeuvre and, and the one thing that concerned me was the fact that, erm, they needed four people to move him, you know, if he wanted to go to the toilet or if he wanted to go on a bedpan or anything, and we had the bed in there – which he wouldn't sleep in. And, erm, basically the, logically trying to be able to do everything for him and keep him comfortable, we'd have to wait for an on-call four nurses – could be in the middle of the night – and, and sort of the idea of being able to cope, erm, for his safety and wellbeing was, was really compromised. He didn't want to go into St Barbara's, he didn't want to die in hospital, erm, but I just felt I had to take that decision to say, erm, when the guy came out to assess him, erm, he said, 'We can do it but, you know, you've got to say what you're going to do at three o'clock on Saturday, early hours of

Commented [PKM1]: Delete and replace with "This gentleman"

Commented [PKM2]: Delete

Commented [PKM3]: Delete

Commented [PKM4]: Delete

Commented [PKM5]: Delete

Commented [PKM6]: Delete

Commented [PKM7]: Delete

Commented [PKM8]: Delete

Commented [PKM9]: Delete and replace with "The couple's"

Commented [PKM10]: Delete

Commented [PKM11]: Delete and replace with "the patient"

Commented [PKM12]: Delete and replace with "her husband"

Commented [PKM13]: Delete and replace with "his"

Commented [PKM14]: Delete. Replace with [your husband's]

Commented [PKM15]: Delete. Replace with [your husband]

Commented [PKM16]: Delete and replace with [he]

Commented [PKM17]: Delete and replace with [the hospice]

Commented [PKM18]: Delete and add [hospice at home]

Commented [PKM19]: Delete and replace with [he]

Commented [PKM20]: Delete and replace with [the hospice]



This work is licensed under a [Creative Commons Attribution 4.0 International License](#).



Checklist: Requirements for a consent according to GDPR

To check...	Yes	No
Overall		
Is consent obtained before the collection and use of personal data?		
Does the consent relate only to data processing that is not already legitimised by law? (The legal consequences of "unnecessary" consent are controversial.)		
Form		
Is the unambiguousness or the proactive act of consent guaranteed? Consent must be given unambiguously, i.e. by an active act of the person giving consent (e.g. by ticking a box).		
Is informed consent "part of a larger document"? If so, it must be clearly distinguishable from the other facts of the document: Are requirements for the "visual" highlighting of the data protection consent respected?		
Has a duplicate copy of the document been provided for? (Keep the original with the person responsible, the copy with the person concerned)		
Has the opportunity for questions been granted? In this respect, the following wording is recommended: "...I had the opportunity to ask questions. These were answered fully and comprehensively. ..." The name of the person who answered the questions, if applicable, should be handwritten on the consent form.		
Voluntariness		
Did the person concerned have a real choice between approval or rejection?		
Is it guaranteed that the fulfilment of a contract or the provision of a service was not made dependent on consent if consent is not absolutely necessary for fulfilment (prohibition of tying)?		
Informedness		
Has the person concerned received all the necessary information (including advantages and disadvantages)? In particular: - data use (purpose, objective, benefits, chances and risks) - persons who may have access to data - the types of data processed - data transfer (to whom, if necessary, storage at which location, country)		
Is all the information specified in Art. 13 GDPR or Art. 14 GDPR provided? In particular: - contact person and contact details (person responsible, data protection officer, ...) - legal basis of the agreement - recipient - storage duration - rights of the person concerned (access, correction, deletion, withdrawal of consent)		
Are the person responsible and his or her representatives clearly identified? Are all necessary contact details available to the person concerned?		
Is there a (comprehensible) explanation of the consequences that a refusal to give consent may have for the person concerned?		



To check...	Yes	No
When processing special categories of data (Art. 9 GDPR), does the declaration of consent explicitly refer to these data?		
Definiteness		
Does the informed consent relate to a specified case? General consents are invalid; separate consents must be obtained / given for different purposes.		
Is the declaration of consent clearly separated from any other (data protection-related) information? It should be avoided that the person concerned cannot recognize whether and, if so, to what he or she actually consents or should consent due to the lack of clarity of the document.		
Revocability		
Does the consent form indicate that consent may be revoked at any time?		
Does the consent form indicate that revocation always applies only to the planned processing that takes place after revocation?		
Is the revocation of consent (at least) as simple as giving consent itself?		
Is there a (clear) indication of the consequences of revocation?		
Informed consent of minors		
Is Art. 8 for processing by means of "information society services" respected?		
If parental consent is given: at the latest when the person concerned comes of legal age, further processing is only permitted with the consent of the person concerned.		
Is there a mechanism to stop the processing of data at the point in time "x"?		
Verifiability		
Is there proof that the consent was given by the person concerned?		
Is there proof that consent has been given in a form sufficient to meet the requirements of the GDPR?		
Is the consent granted recorded? If so: Have sufficient technical and organisational measures been taken to protect the records? (evidence)		
Is it possible to access the granted consent at any time?		

Source:

Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS) (2016). EU DS-GVO: Anforderungen an eine Einwilligung. Online available: <https://www.gesundheitsdatenschutz.org/download/einwilligung.pdf>. Last accessed: 24.11.2019. Licensed under Creative Commons Share-Alike (4.0 Germany) license. [transl. MB]

17. Teaching script: Legal aspects							
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound
						Alternative	Notes
Introduction	The participants activate previous knowledge	10	The participants are divided into groups of 3. Based on the stages of RDM, they answer: In which of the stages can legal issues be relevant? Which legal areas? Explain the task: 2 min Cooperation: 5 min Recall results: 3 min	Exercise	Paper, pencils	Ex Yes	
Legal aspects	The participants learn which legal regulations may have to be considered when publishing research data	5	Copyrights, data protection, personal rights, rights of third parties – each with examples	Presentation	PPTX	In -	
	The participants know about what kind of laws there are, and what personal data and informed consent are	5	The participants learn the basics of the applicable laws (EU, Germany, federal state, university), policies, guidelines, contracts and definitions of personal data, anonymisation, pseudonymisation and informed consent	Presentation	PPTX	In Yes	
	The participants are aware of where to get advice	5	The participants answer the question: Where can you get advice on legal aspects of RDM?	Ideas out loud	Flip chart	Ex Yes	L: Exercise: Identification of qualitative data (15 min) Checklist: Requirements for informed consent according to GDPR

Duration of the unit: 25 minutes



Unit 18: Institutional infrastructure

Learning objectives

Participants get to know the services for research data management at colleges/universities or other research institutions.

Participants will learn about the services for handling research data at their own institution or at a typical institution.

Key aspects

1. Institutional research data management
2. Infrastructure for research data management

Contents

1. Institutional research data management

At each educational institution, the topic of research data management is dealt with quite differently. There are institutions that already have a coordination office for years. At other institutions, appropriate services are still being developed or planned.

A central coordination office provides a better overview of the internal services for managing research data. Usually, advice as well as general and discipline-specific trainings are offered here. Support material such as flyers, posters and FAQs are available to staff, researchers and students.

2. Infrastructure for research data management

Universities must provide their researchers with the best possible support in research data management using the appropriate tools and services, both organisationally and technically. Depending on the capacities of the institutions, the following services could be available:

- Advice on the management of research data in connection with project proposals or applications for external funding
- Possibly specialist areas of focus, e.g. consulting on discipline-specific issues in the field of digital humanities or social sciences
- Transfer to internal and external services

- Support in the creation of a data management plan
- Training or Workshops on research data management
- Information events
- Information material
- Legal advice on research data (also in cooperation with experts)
- Cloud services
- Sync and share solutions
- Backup Service
- Database Service
- Research data repository (sometimes integrated in a publication server)
- Versioning software
- Long-term archiving
- Allocation of Persistent Identifier
- DMP tools
- Survey tools
- Electronic Lab Notebooks

Didactic methods and exercises

Methods:

- Ideas out loud
 - What other or additional services are available at your institution?

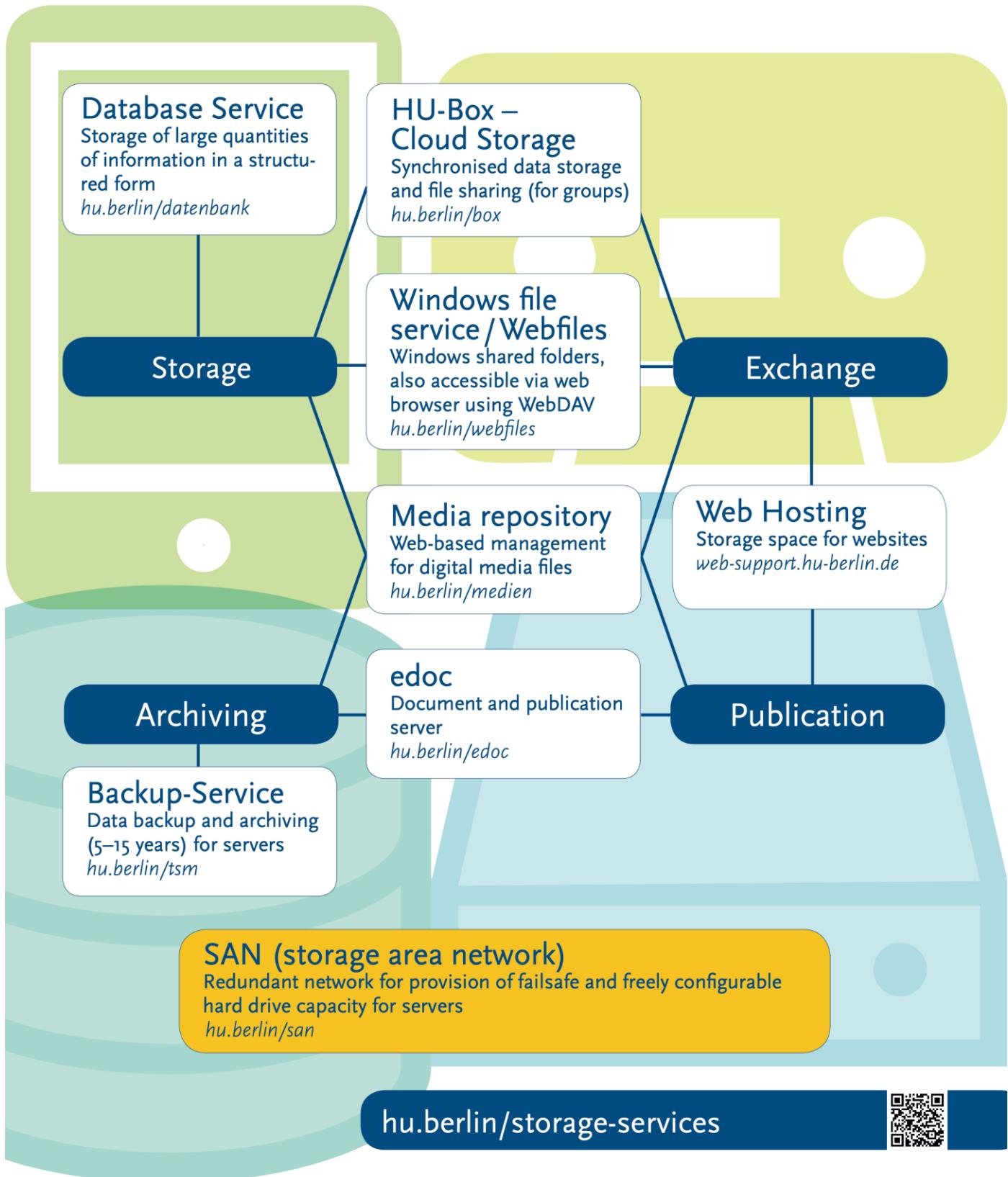
Training materials

- Storage services of the Computer and Media Service of the Humboldt-Universität zu Berlin⁹³
- Teaching script: Institutional Infrastructure

⁹³ Humboldt-Universität zu Berlin: "Storage services of the CMS for research data management." Last accessed 02.11.2020. <https://www.cms.hu-berlin.de/en/dl-en/storage-services>.



Storage services of the CMS for research data management



18. Teaching script: Institutional infrastructure									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voice sound	Alternative	Notes
Research data management at an academic institution	The participants get to know the RDM services at colleges/ universities or other research institutions	3	Presentation of possible services	Presentation	PPTX	In	-		
	The participants know the services of an exemplary institution for handling research data	5	The services of an exemplary institution are presented and the relevant contact persons are named	Presentation	PPTX	In	-		Example Humboldt-Universität zu Berlin: HU-Box, backup service, database service, media repository, long-term storage, edoc publication server, persistent identifier
	The participants connect what they have learned with their own work routine	2	The participants answer the question: What other/further services are available at your institution?	Answer in turn	-	Ex	Yes	L: Talk with a person from the institution's IT centre (7 min)	

Duration of unit: 10 minutes



Unit 19: Training exercise

Learning objectives

The participants can transfer the newly acquired knowledge to different problems.

The participants make use of their newly acquired knowledge in practice.

Key aspects

1. Training exercise

Contents

1. Training exercise

The subject matter learned on research data management will now be applied in a training exercise. On the basis of an individual example, an original data management plan is to be created in which all of the points discussed are outlined.

The data management plan used here is significantly shorter than, for example, the plan used in Horizon 2020. We decided on this version because the latter is very extensive and it would not be possible to process it within this unit in terms of time required. The aim is to review the various aspects of research data management when completing the plan, and thus to repeat the content of the workshop. Any questions that arise in the process can be clarified in the workshop.

Training materials

- Worksheet: Data management plan
- Exemplary solution: Data management plan
- Research data quiz
- Solutions: Research data quiz
- Teaching script: Training exercise



Worksheet: Data management plan

Project name:

Research funder:

Funding program:

Primary researcher/researcher/project leader:

ID Primary researcher/researcher/project leader:

Project description:

.....
.....
.....
.....
.....
.....
.....
.....

Creation date:

Update:

Data collection:

.....
.....
.....
.....
.....
.....
.....
.....

Data storage:

.....
.....
.....
.....
.....





Data documentation:

.....
.....
.....
.....
.....
.....

Data availability:

.....
.....
.....
.....
.....
.....

Data archiving and preservation:

.....
.....
.....
.....
.....

Responsibilities:

.....
.....
.....
.....
.....
.....





Exemplary solution: Data management plan

Project name: Analysis of the inclusive educational competence of educators in Brandenburg (ANIECE-BB)

Research funder: German Federal Ministry of Education and Research

Funding programme: Qualification of pedagogical specialists for inclusive education

Principal investigator/researcher/project leader: Kerstin Helbig

ID Principal investigator/researcher/project leader: <http://orcid.org/0000-0002-2775-6751>

Project description:

The project deals with the competence of educators in the field of inclusive education in Brandenburg day-care centres. By means of focus groups, the educators are interviewed about their current approach to inclusive education. In addition, data from the Federal Statistical Office will be reused. The data collection will be used to analyse the importance of inclusion in Brandenburg and to determine the need for support and necessary future training and services in the field of inclusive education.

Creation date: version 1 from March 16, 2020

Update: version 2.3 from April 26, 2020

Data collection:

Focus groups are organized and surveyed throughout Brandenburg. The answers are stored as video recordings and subsequently transcribed. The evaluation of the answers is done with MAXQDA. Excerpts from the videos will also be used for teaching and further education. Existing data will be used as well. A secondary analysis of the statistics of children and persons working in day care facilities (EVAS 22541) of the Federal Statistical Office will be carried out. The statistics are part of the data of the Child and Youth Services Statistics (KJH). The data are evaluated with the help of the statistics program R. The data are representative, since it is a complete survey.

Data storage:

During the project storage and backup will be ensured by the project manager in cooperation with the responsible IT representative of the Computer and media service (CMS) of the University. The infrastructure of Humboldt-Universität zu Berlin will be used for this purpose. The research data will be stored in the HU-Box and secured with a password. Only authorized employees have access. A backup of the data is made once a day. Version control will be automated. File naming is done according to the following standard:

[focus group]_[location]_[YYYYMMDD].mp4

[statistics]_[file type]_[YYYYMMDD]_[version].csv

Files are stored in formats that are as open and standardized as possible. The file formats PDFA, CSV, MPEG-4 (audio track WAVE) and, if necessary, TIFF are used for this purpose. Where conversion to an open format is not possible, original formats are stored.

Data documentation:

Metadata will be created via GESIS - Leibniz Institute for the Social Sciences according to [the DDI standard](#). In addition, the metadata are included in the German national educational portal forschungsdaten-bildung.de. Additional documentation of the research data is planned as well. The following documents will be created:



- Transcription manuals
- Focus group guidelines
- QDA files
- R syntax
- Consent forms
- Anonymization measures

Keywords will be assigned according to the subject-specific [thesaurus for the social sciences \(TheSoz\)](#). The study is classified via GESIS using the [classification of social sciences](#).

Data availability:

The digital research data obtained will be published openly under a Creative Commons Attribution CC-BY 4.0 International license, provided that there are no data protection concerns. Further data is provided with restrictive access. The GESIS - Leibniz Institute for the Social Sciences is used for making the data available. The data has great value, especially in teaching. In addition, they can serve as a basis for comparison for further nationwide or state-wide studies. The potential for re-use can therefore be considered very large. Consequently, the aim is to make access as open as possible.

Data archiving and preservation:

The research data underlying a publication, but also other relevant milestone versions of the project files are archived for at least ten years. Data for which there is no legal archiving basis will be deleted shortly before the end of the project. The data protection officer of the Humboldt University of Berlin is involved in this process. The expected total size of the remaining data is about 100 GB. The long-term archiving for at least 10 years will be provided by GESIS - Leibniz Institute for the Social Sciences. In addition, the project results and all relevant research data will be stored for 15 years at servers of Humboldt-Universität zu Berlin using the backup service of the university.

Responsibilities:

The project manager is responsible for the secure storage and long-term archiving of the generated digital research data together with the institute's IT representative. An additional three project months are planned for the corresponding preparation of the research data for publication or access. Making the data available and archiving by GESIS - Leibniz Institute for the Social Sciences is free of charge. Likewise, the use of the HU backup service does not incur any additional costs.

Source: Based on Helbig, Kerstin. Muster-DMP für die Erstellung eines Datenmanagementsplans im Rahmen eines BMBF-Antrags. Online available at: <https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-bmbf>.

Research data quiz

1	2	3	4	5	6	7	8		9	10	11	12
						I						
						G						
						T						
		N				L			P		E	
						O					P	
L	A	G		F		A						
		E			J	R		T	I			
C					E				C	I		
E	S		R					D				
E		N	C			I		A			C	
-	T		K	R					I		C	
E		P	U		D	E		A		Y		
				N					N			
E		A		C	N						S	
						I					I	
				L							B	
				E								
					E							
					R							

1. Part of the research data life cycle.
2. Legal regulation for provision and use of data.
3. Your research is based on it.
4. Organizes your research.
5. When you save all aspects of your work not just selected data to avoid loss, you create this.
6. Four basic aspects of Research Data Management.
7. Makes your data unique.
8. Long-term storage option.
9. Data about your data.
10. Final step of your research.
11. A place where things are stored and can be found.
12. 2nd FAIR principle.

Solution: Research data quiz

1	2	3	4	5	6	7	8		9	10	11	12
						D						
						I						
						G						
						I						
		M				T						
		A				A				P	R	
		N				L				U	E	
	D	A			O			M	B	P		
L	A	G		F	B	A		E	L	O		
I	T	E		A	J	R		T	I	S		
C	A	M	B	I	E	C		A	C	I		
R	E	S	E	A	R	C	H		D	A	T	A
E	N	E	N	C	P	T	I	A	T	O	C	
-	C	T	T	K	R	I	V		T	I	R	C
U	E		P	U	I	D	E		A	O	Y	E
S			L	P	N	E				N		S
E			A		C	N						S
			N		I	T						I
				P	I							B
				L	F							L
				E	I							E
				S	E							
					R							

Kristin Meier, Leibniz Centre for Agricultural Landscape Research (ZALF) e.V., licensed under CC-BY 4.0

19. Teaching script: Training exercise						
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation
						Voice sound
DMP-Exercise	The participants create their own DMP	12	T explains the task: 2 min The participants write a DMP: 10 min	Exercise	Worksheet: Data management plan	Ex -
	The participants identify problem areas and discuss ideas for solutions	8	The participants discuss their problems and solutions	Ideas out loud: Participants present their problems and the group develops solutions. A sample solution is handed out	-	Ex Yes

Duration of the unit: 20 minutes



Unit 20: 7 steps of concept development

Learning objectives

Participants learn about the 7 steps of concept development according to Harald Groß.⁹⁴

The participants know how to create a teaching concept.

Key aspects

1. Open the subject
2. Clarify requirements (3T formula)
3. Organize, focus, reduce
4. Develop teaching scripts
5. Create methods and exercises
6. Create teaching material
7. Review concept

Contents

1. Open the subject

The first step of concept development deals with the subject to be taught. At this point there should be no limits to creativity. It is recommended to write down, draw or make notes of everything that comes to mind, regardless of the general setting of the event.

In this step the subject matter is not yet defined, this will happen later. It is rather a basis for selection: What does the topic include?

⁹⁴ Harald Groß: „Arbeitsheft Nr. 4: Königsdisziplin Stoffreduktion.“ In Harald Groß (ed.): *Didaktik*. Orbium Seminare, internal seminar material.

2. Clarify requirements (3T formula)

The 3T formula was already introduced and discussed in unit 12: Formal framework (p. 92).

3. Organize, focus, reduce

Often the subject matter to be covered is too extensive for the time available. Therefore, it is important to select the essentials for the workshop. This means being thorough rather than exhaustive. In this way, the participants gain specialist thinking and can grasp the connections within the topic more easily. The teacher sets priorities, thus ensuring transparency.

A great help in reducing or selecting material is to identify the most important terms. As a rule of thumb, no more than seven such terms should be chosen.

Once these terms have been defined, the main related topics can be attributed. One should choose as many subjects/terms as necessary, but at the same time as few as possible.

For this step, it is a good idea to create a workshop map early on. More information on this can be found in unit 2: Orientation (p.18).

4. Develop teaching scripts

As soon as the units of the workshop are decided on, it is time to complete them with content and methods. Teaching scripts are ideal for this. A chart, containing an overview of the modules covered, learning objectives, the time required, content, methods and material to be used, will help to both prepare for and carry out the workshop. The modules of a topic can be structured quite finely in order to get a better sense of timing. Short, medium and long versions (S-M-L) of exercises, methods or explanations should also be planned and written down to allow for alternatives during the implementation (see unit 21: Didactic methods, p. 155).

5. Create methods and exercises

There is a wide range of teaching methods to be used in different settings. Each of these methods has its own goal and may support the different basic forms of learning, e.g. to analyse, observe, transfer, etc. Each of the methods also supports a different social and working format, e.g. individual or group work.

The selection of methods is based on the following criteria: goal, subject, participants, energy, setting and teaching style.

Every teacher has to choose methods that he/she feels comfortable with and that suit him/her. The number of participants often determines the amount of time needed for an exercise, which must be considered when writing the teaching script (S-M-L variants are recommended here, see unit 21 point 2).

If the teacher notices participants "slackening", several activating methods might be chosen to raise the energy level.

Of course, the method should suit the subject and be feasible under the given circumstances.

6. Create teaching material

A comprehensive lesson also includes teaching material. These may be slides as well as hand-outs, worksheets, scripts or photo protocols. All of these materials should help the participants to process the new content and serve as a potential reference work. Again, different objectives can be pursued with different types of material.

Study material can be distributed before, during or after the event. Each of these options has its advantages:

- Before – offers an orientation to the participants; can be used as preparation for the subject; makes the workshop more transparent

- During – during the course of the event, training material may be used to direct the participants' attention and enhance their ability to stay focused.
- Afterwards – thus the participants are given the opportunity to repeat and remember the contents after the event.

7. Review concept

Before employing the developed teaching concept in practice, it should be thoroughly reviewed. Does the plan suit the target group? Do you include a topic-related and a social introduction? Are there sufficient work phases? Is the time planned carefully? Are the guiding and learning objectives actually being pursued?

Didactic methods and exercises

Methods:

- Ideas out loud
 - Which steps of concept development do you remember from the previous units?

Exercises:

- Create a detailed plan for your course. Answer the following questions and write down the answers:
 - What is the target audience for my course?
 - What is the time frame for the course?
 - What context is it embedded in?
 - What do I wish to convey?⁹⁵

Training materials

- Template for a teaching script
- Teaching script: 7 steps of concept development

Additional sources

- Blumer, Eliane and René Schneider: „Modul 9-3: Train the Trainer: Methodik & Didaktik.“ In Mastrandrea, Elena, Nicolas Prongué, René Schneider and Niklaus Stettler. *Kursbuch Forschungsdaten*. Chur: HTW Chur, 2017. Last accessed 29.07.2020, https://campus.hesge.ch/researchdatamanagement/?page_id=4928.
- Lehner, Martin: *Viel Stoff – wenig Zeit. Wege aus der Vollständigkeitsfalle*. 4th ed. Bern: Haupt, 2013.

⁹⁵ Eliane Blumer and René Schneider: „Modul 9-3: Train the Trainer: Methodik & Didaktik.“ In Mastrandrea, Elena, Nicolas Prongué, René Schneider and Niklaus Stettler. *Kursbuch Forschungsdaten*. Chur: HTW Chur, 2017. Last accessed 29.07.2020. https://campus.hesge.ch/researchdatamanagement/?page_id=4928.

Teaching script:						
Component	Target	Time	Content	Working modes	Material	Inhalation/ exhalation
						Voices sound
						Alternative
						Notes

Duration of the unit:



Developed within the FD-Mentor project
Funding code: 16FDI010 and 16FDI011
Project duration: 1. May 2017 - 30. April 2019

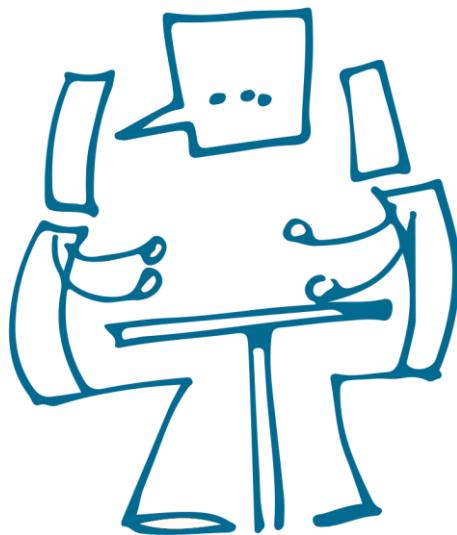


This work is licensed under a
Creative Commons
Attribution 4.0
International License
(https://creativecommons.org/
licenses/by/4.0/)



20. Teaching script: 7 steps of concept development						
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation
					Voices sound	Alternative
Introduction	The participants develop first ideas on the topic of concept development	3	The participants answer the question: Last time we already mentioned the steps of concept development during the topic "formal framework". What do you remember?	Ideas out loud	Flip chart	Ex Yes
	Steps of concept development according to Harald Groß	12	The participants learn the 7 steps of concept development according to Harald Groß	The 7 steps are presented	PPTX Presentation	In -
						If a lot has already been previously mentioned, then special priorities can be set here, e.g. content reduction

Duration of the unit: 15 minutes



Unit 21: Didactic methods

Learning objectives

Participants will be able to distinguish between the different methods.

The participants know the different goals to be achieved with certain methods.

Key aspects

1. Choice of method
2. Short-Medium-Long versions (S-M-L)
3. Activating methods
4. Overview of the methods used
5. Other methods from our pilot workshops

Contents

1. Choice of method

When choosing the method, the trainer should be guided by the following aspects⁹⁶:

- Goal
- Subject
- Participants
- Setting
- Energy
- Style

The first question to be answered is: What do I wish to achieve by using this method? As a next step, consider the subject and decide which methods would be appropriate. Another criterion is the number

⁹⁶ Based on Groß, Harald, Betty Boden and Nikolaas Boden: *Munterrichtsmethoden : 22 aktivierende Lehrmethoden für die Seminarpraxis*. 3rd, completely revised ed.. Berlin: Schilling, 2012, pp. 127–137.

of participants, as not every method is suitable for very small or large groups. Furthermore, the duration of the implementation will depend on it, which brings us to the next point: the setting of the event. Of course, the method must fit into the time frame of the class and be feasible in the given location.

A good indicator for the right choice of a method is also the level of energy in the room. It is good to observe the participants and decide whether to use an activating method or one that requires less movement but stimulates the mind.

Last but not least, the trainer's own style should be considered. You should always choose a method you feel comfortable with as a teacher. It should be a method that matches your own teaching style.

2. Short-Medium-Long versions (S-M-L)

It is difficult to decide on methods before the training. There are plenty of unknowns at this stage: How many participants are actually attending? How much energy do they bring? How much previous knowledge do they have? What are their interests? How much do they enjoy contributing?

In order to maintain flexibility in time management, it is recommended to prepare three versions for some exercises: a short, a medium and a long version (S-M-L). This allows you to respond well to the needs of the participants and, depending on the time at hand, to choose the method variant best suited to them in order to stay on schedule. This ensures a smooth and relaxed process for the participants, who do not need to be interrupted at points of particular interest, as well as for the workshop instructor, who can meet the time schedule without any problems.

3. Activating methods

Sometimes, when the energy level of the participants is low, either because of the time of day or because a topic required a lot of attention, it is helpful to bring more dynamic into the group.⁹⁷ This works especially well with movement and humour. Sometimes it helps to choose a method of grouping that gets the participants moving more than, for example, just cooperation with neighbours. However, there are also methods with more of a gaming character which do not involve RDM content, such as gotcha. In order to facilitate subsequent use, some recommendations of methods suitable as activations are given below:

- Landscape setting
- Question ball
- Gotcha
- 7 -Plop

These methods employed by us are described in more detail in the following sub-unit.

4. Overview of the methods used

The methods used in this workshop:

1. Landscape setting⁹⁸
2. We and I⁹⁹
3. Expectation query
4. Harvest sheet¹⁰⁰

⁹⁷ Will, Hermann: *Mini-Handbuch Training und Seminar*. Weinheim: Beltz, 2016, pp. 144-152.

⁹⁸ Zamyat M. Klein: *Kreative Seminarmethoden: 100 kreative Methoden für erfolgreiche Seminare*, Offenbach: Gabal, 2003. 16–17.

⁹⁹ Orbium Seminare. "Muniterrichtsmethode 45: Wir und ich." Last accessed 03.11.2020, <https://www.orbium.de/methodensammlung/muniterrichtsmethode-45-wir-und-ich/>.

¹⁰⁰ Based on Groß, Harald, Betty Boden and Nikolaas Boden: *Muniterrichtsmethoden: 22 aktivierende Lehrmethoden für die Seminarpraxis*. 3. vollständig überarb. Aufl. Berlin: Schilling, 2012, pp. 112–115.

5. Chatter¹⁰¹
6. Flip and turn¹⁰²
7. Question ball
8. Truth or lie
9. Gotcha¹⁰³
10. 7 -Plop
11. User manual
12. LEGO® SERIOUS PLAY®¹⁰⁴
13. Estimation question¹⁰⁵
14. Blind decision¹⁰⁶
15. Mind map
16. Inventory¹⁰⁷
17. Ideas out loud
18. Keyword strips¹⁰⁸
19. Statement slam¹⁰⁹
20. Interim balance¹¹⁰
21. Schema-X¹¹¹
22. Fried Egg
23. Five-Finger-Feedback
24. Energy level

5. Other methods from our pilot workshops

On the second day of the pilot workshops the participants had the task to describe a method they know and like using the method schema X. In addition to the methods used in this workshop, we will introduce five of our participants' favourite methods below:

25. Photo reflection
26. Competency pizza

¹⁰¹ Groß, Munterrichtsmethoden, pp. 116–119.

¹⁰² Groß, Munterrichtsmethoden, pp. 98–104.

¹⁰³ Groß, Harald. *Munterbrechungen: 22 aktivierende Auflockerungen für Seminare und Sitzungen*. 2nd ed. Berlin: Schilling, 2012, pp. 56–58.

¹⁰⁴ LEGO Group. "LEGO® Serious Play." Last accessed 02.11.2020. <https://www.lego.com/en-us/seriousplay>.

¹⁰⁵ Groß, Munterbrechungen, pp. 84–87.

¹⁰⁶ Groß, Munterrichtsmethoden, pp. 52–53.

¹⁰⁷ Groß, Munterrichtsmethoden, pp. 79–83.

¹⁰⁸ Orbium Seminare. "Munterrichtsmethode 46: Stichwortsalat." Last accessed 03.11.2020. <https://www.orbium.de/Methodssammlung/munterrichtsmethode-46-Stichwortsalat/>.

¹⁰⁹ Groß, Munterrichtsmethoden, pp. 34–38.

¹¹⁰ Groß, Harald. *Munterrichtsmethoden: 22 weitere aktivierende Lehrmethoden für die Seminarpraxis*. Vol. 2. Berlin: Schilling, 2014, pp. 57–60.

¹¹¹ Groß, Munterrichtsmethoden, pp. 72–76.

- 27. Data set XYZ
- 28. RDM slam
- 29. Topic memory

Didactic methods and exercises

Methods:

- Interim balance
 - Which methods did you get to know in the course of the workshop and what was their objective?
- Schema-X
 - Create your own method in groups and write down the basic details. Present your results to the whole group.

Exercises:

- The participants make suggestions on possible selection criteria for teaching methods. The answers are collected on a pin board and supplemented with additional answers.

Training materials

- Template: Schema-X
- Explanation of the methods
- Teaching script: Didactic methods

Additional sources

- Groß, Harald, Betty Boden and Nikolaas Boden: *Munterichtsmethoden: 22 aktivierende Lehrmethoden für die Seminarpraxis*. 3rd ed. Berlin: Schilling, 2011.
- Groß, Harald: *Munterichtsmethoden: 22 weitere aktivierende Methoden für die Seminarpraxis*. Vol. 2. Berlin: Schilling, 2014.
- Groß, Harald: *Munterbrechungen: 22 aktivierende Auflockerungen für Seminare und Sitzungen*. Berlin: Schilling, 2010.
- Klein, Zamyat M.: *Kreative Seminarmethoden. 100 kreative Methoden für erfolgreiche Seminare*. 8th ed. Offenbach: GABAL, 2003.
- Will, Hermann: *Mini-Handbuch Training und Seminar*. Weinheim: Beltz, 2016.



Template: Schema-X

Title:	
Goal:	Description:
Duration:	Required materials:



Explanation of methods

1. Landscape setting	
<p>Goal:</p> <ul style="list-style-type: none"> • get to know each other • activation • getting participants to speak 	<p>Description:</p> <p>The participants position themselves in the room so that there is enough space for everyone. The teacher asks a question (e.g. about the field of work, age, previous knowledge) and the participants have to find their similarities and differences in order to group themselves accordingly in the room.</p> <p>For some questions, even lines can be formed, e.g. the distance between the place of birth and the event location or the number of years of professional experience.</p> <p>No oral answers or explanations are required in front of the whole group. The positioning in the room is answer enough.</p> <p>To lighten the atmosphere, it is advisable to mix professional questions with personal ones.</p>
<p>Duration: Depending on number of questions: 10-15 minutes</p>	<p>Required materials:</p> <ul style="list-style-type: none"> • possibly a crepe tape to mark a line

2. We and I	
<p>Goal:</p> <ul style="list-style-type: none"> • get to know each other • getting participants to speak 	<p>Description:</p> <p>The participants are split into groups of three people each, preferably in such a way that they know each other as little as possible.</p> <p>Each group receives a flip chart with the "us-and-me" template. In the middle, they enter the identified commonalities. From work to private life, everything that the participants are interested in can be recorded here.</p> <p>In addition, all participants get their own field on the sheet – the "Me" field. The characteristics that make the participant unique are noted here.</p> <p>At the end, each team presents their collection.</p>
<p>Duration: Explanation and group formation: 3 minutes Working time: 15-20 minutes Presentation: 3-5 minutes per group</p>	<p>Required materials:</p> <ul style="list-style-type: none"> • flip charts with "us-and-me" template • pens

3. Expectation query	
Goal: <ul style="list-style-type: none"> query expectations draw attention 	Description: Participants receive cardboard cards on which they write down in keywords what they expect from the workshop or topic. At the end, the expectations are presented to the group and pinned to the workshop map.
Duration: Write down the expectations: 5 minutes Presentation of expectations: 10 minutes	Required materials: <ul style="list-style-type: none"> cardboard cards (A5 or A6) pens pins workshop map

4. Harvest sheet	
Goal: <ul style="list-style-type: none"> reminders on subject matter 	Description: The participants receive a sheet (the harvest sheet) on which they make notes in keywords, drawings, half-sentences during the workshop.
Duration: Throughout the workshop	Required materials: <ul style="list-style-type: none"> printed harvest sheet pens

5. Chatter	
Goal: <ul style="list-style-type: none"> exchange consolidation repetition of contents preparation for discussion 	Description: The participants exchange information on a question with a neighbour.
Duration: 2-3 minutes per question	Required materials: <ul style="list-style-type: none"> none

6. Flip and turn	
Goal: <ul style="list-style-type: none"> work out coherences independently 	Description: Participants receive cards with key terms of a process, model or theory. In groups, they should now arrange the cards and present their results to the other participants.
Duration: Explanation: 3 minutes Execution of exercise: 5-10 minutes Presentation of results and discussion: 10-15 minutes	Required materials: <ul style="list-style-type: none"> cards with key terms adhesive tape pins

7. Question ball	
Goal: <ul style="list-style-type: none"> • activation • query keywords • recollection 	Description: The teacher asks a question to which the participants should answer in brief using keywords. The teacher throws the ball to the person who should answer the question first. The participant throws the ball to any person after answering the question, etc.
Duration: Depending on question: 3-5 minutes	Required materials: <ul style="list-style-type: none"> • soft ball

8. Truth or lie	
Goal: <ul style="list-style-type: none"> • remembering • consolidation • stimulating discussion 	Description: The teacher mentions various scenarios. The participants indicate by hand sign or individually with pen and paper whether they consider the respective scenario to be true or a lie.
Duration: Depending on scenarios: 3-5 minutes	Required Materials: <ul style="list-style-type: none"> • prepared scenarios, possibly as an exercise sheet • pens, if applicable.

9. Gotcha	
Goal: <ul style="list-style-type: none"> • activation • increase concentration 	Description: The participants stand in a circle with the palm of their left hand pointing upwards (at chest height, to their left). The right index finger hovers over the open hand of the right neighbour. At a signal, the participants try to grab the finger of the left neighbour on the one hand and to escape the right neighbour on the other hand.
Duration: 5 minutes	Required materials: <ul style="list-style-type: none"> • none

10. 7 -Plop	
Goal: <ul style="list-style-type: none"> • activation • increase attention 	Description: The participants sit on their seats. One after the other, they will count to 70 together. There are a few rules, though: <ul style="list-style-type: none"> • the number 7 must not be spoken. Instead, PLOP is said, • this also happens with all numbers divisible by 7 • and for all numbers in which the digit 7 appears • and for numbers with the checksum 7. When someone makes a mistake, the others are happy for him and with him.
Duration: 5 minutes	Required materials: <ul style="list-style-type: none"> • none

11. User manual	
Goal: Transfer what has been learned: <ul style="list-style-type: none">• Participants use their previous knowledge and what they have learned and transfer it to a new context• Participants recognize criteria of good documentation	Description: The workshop leader provides various items. Each group will receive an object and should describe it and its use. The group creates a user manual for the item.
Duration: Introduction: 2 minutes Descriptions: 15 minutes Test and exchange: 5 minutes per group	Required materials: <ul style="list-style-type: none">• 1 object per group, e.g. a frisbee, a can opener, a jump rope etc.• flipchart sheets• pens

12. LEGO® SERIOUS PLAY®	
Goal: <ul style="list-style-type: none">• Transfer what has been learned in a playful way	Description: LEGO® SERIOUS PLAY® combines playful model making by means of LEGO® bricks with reaching learning goals. It can be used for different learning scenarios. Among other things, the method is suitable for planning the creation of metadata, the formats for the collection/generation of metadata, standards and automation. The method involves several parallels between the capture and communication of the research process and the documentation and creation of a Lego model. The method encourages researchers to discuss how metadata is collected and disseminated, which in turn provides an opportunity to point out further resources in this area.
Duration: Introduction: 5 minutes 1. Design: 20 minutes 2. Design: 15 minutes	Required materials: <ul style="list-style-type: none">• LEGO® bricks• pens• paper• instructions for different groups

13. Estimation question	
Goal: <ul style="list-style-type: none">• activate knowledge base• estimating• working out connections	Description: The teacher asks a question to which the participants probably do not know the exact answer. The participants should guess and estimate.
Duration: Depending on question: 3-5 minutes	Required materials: <ul style="list-style-type: none">• none

14. Blind decision	
Goal: <ul style="list-style-type: none"> • query knowledge • recollect contents • refresh contents 	Description: The participants answer a yes-or-no question with their eyes closed by hand signals (thumbs up, thumbs down, thumbs in the middle). Only when all participants have given a sign the eyes may be opened. The decisions are then discussed.
Duration: Depending on the number of questions: 5-10 minutes	Required materials: <ul style="list-style-type: none"> • none

15. Mind map	
Goal: <ul style="list-style-type: none"> • remembering • repeating • refreshing • collecting 	Description: The participants each write down on a large sheet of paper everything that comes to mind on a given topic. With lines, arrows or circles connections can be represented.
Duration: Depending on the scope of the question: 10-15 minutes	Required materials: <ul style="list-style-type: none"> • large sheets of paper (a3) • pens

16. Inventory	
Goal: <ul style="list-style-type: none"> • remembering • repeating • refreshing 	Description: The participants write down everything they know about a given topic on a large piece of paper. They should collect a minimum number of (for example) 17 pieces of information. The number should be small enough to be feasible, but at the same time large enough to challenge participants to explore their knowledge in more detail. Subsequently, the results or parts thereof can be presented
Duration: Without presentation of results: 10-15 minutes With presentation of results: 20-30 minutes	Required materials: <ul style="list-style-type: none"> • large sheets of paper (a3) • pens

17. Ideas out loud	
Goal: <ul style="list-style-type: none"> • activation • recollection • query knowledge 	Description: The teacher asks a question and the participants call out the solution without being asked to do so directly. There is no particular order of answers.
Duration: Depending on question: 3-5 minutes	Required materials: <ul style="list-style-type: none"> • none

18. Keyword strips	
Goal: <ul style="list-style-type: none"> • recollection • review of contents • reproduction 	Description: <p>The most important keywords of the previously processed content are written on small pieces of paper and placed in an envelope.</p> <p>The participants are divided into groups of 3 and each group receives an envelope with a set of keywords. One participant from the group draws a term and is supposed to explain it to the others. The others may supplement and correct the information. When everything has been said about the keyword, the envelope is passed on to the next person to draw the next term and so on.</p>
Duration: Explanation: 2 minutes Execution of exercise: 5-10 minutes	Required materials: <ul style="list-style-type: none"> • keywords on small sheets of paper • envelopes

19. Statement slam	
Goal: <ul style="list-style-type: none"> • dealing with a topic • arguing 	Description: <p>A provocative statement on a specific topic was written down on each sheet beforehand. The sheets are rolled up individually.</p> <p>Each participant draws a role and prepares his/her opinion on the statement.</p> <p>The opinions of each of the group are then presented.</p>
Duration: Explanation: 2 minutes Preparation of the opinion: 3 minutes Presentation of the statements: 15 minutes	Required materials: <ul style="list-style-type: none"> • sheets of paper with provocative statements • elastic bands

20. Interim balance	
Goal: <ul style="list-style-type: none"> • recap • review of contents 	Description: <p>The participants look for an exercise partner. In the first round, person A interviews person B and asks two questions:</p> <ul style="list-style-type: none"> • What did you learn today? • What are you going to do with it? <p>In the second round the roles are reversed.</p>
Duration: Explanation: 2 minutes Execution of exercise: 5-10 minutes	Required materials: <ul style="list-style-type: none"> • none

21. Schema-X	
Goal: <ul style="list-style-type: none"> • structuring of contents • organizing contents • review of subject 	Description: The participants are divided into four groups. Each group receives a flipchart sheet with a pre-drawn scheme (questions, categories, headlines). This must now be completed in teams. The results are subsequently presented.
Duration: Approx. 30 minutes	Required materials: <ul style="list-style-type: none"> • flipchart sheets with pre-drawn scheme • pens

22. Fried Egg	
Goal: <ul style="list-style-type: none"> • give and receive feedback 	Description: A flipchart sheet shows a fried egg including the egg yolk. The participants receive a card on which they write down what they thought was the 'yellow of the egg' (the essential) in the workshop. Then the participants come in turn to the flipchart and pin their answers to the flipchart. They comment on their selection.
Duration: Approx. 10-15 minutes	Required materials: <ul style="list-style-type: none"> • flipchart sheet with drawn fried egg • presentation cards • pens • pins • pinboard

23. Five-Finger-Feedback	
Goal: <ul style="list-style-type: none"> • give and receive feedback 	Description: The participants give feedback using their hand. In doing so, they may address the following options: <ul style="list-style-type: none"> • Thumb: I liked that! • Forefinger: I'll just pick that up for me! • Middle finger: I did not like that! • Ring finger: I can transfer this to my everyday work! • Little Finger: That got a little too short!
Duration: Approx. 10-15 minutes	Required materials: <ul style="list-style-type: none"> • none

24. Energy level	
Goal: <ul style="list-style-type: none"> • focus attention • learn about students' well-being 	Description: The teacher asks the participants about their current energy level on a scale from 0 (none) to 10 (it doesn't get any better). The participants answer with the number and if they wish, they can add something else.
Duration: Depending on group size: 1-5 minutes	Required materials: <ul style="list-style-type: none"> • none

25. Photo reflection	
<p>Goal:</p> <ul style="list-style-type: none"> • arouse interest • activate previous knowledge • retrieve experience 	<p>Description:</p> <p>The seminar leader spreads out the photos so that all are visible to everyone at the same time and names the topic or the work assignment. Participants select one photo each. The selection process arouses interest and the formation of associations between the photo and the subject. This activates prior knowledge.</p> <p>When the associations are presented, the experiences of the participants and their level of previous knowledge become apparent.</p>
<p>Duration:</p> <p>Selection: 1 minute Reflection: 2 minutes Presentation: 1-2 minutes per person</p>	<p>Required materials:</p> <ul style="list-style-type: none"> • selection of photos • alternative: other pictures, objects, statements or quotations instead of photos

26. Competency pizza	
<p>Goal:</p> <ul style="list-style-type: none"> • query previous knowledge: • overview of the participants' competences • recognize priorities 	<p>Description:</p> <p>The workshop host presents the competency pizza to the participants. The participants position themselves on the pizza according to their skills with adhesive dots or other markings.</p> <p>Other forms of structuring (matrix, diagram, table, ...) or content, such as opinion polls, can also be processed with this method. It is also suitable for quick topic-related feedback.</p>
<p>Duration:</p> <p>5-10 minutes</p>	<p>Required materials:</p> <ul style="list-style-type: none"> • prepared structure of the competencies as a pizza • adhesive dots or pens

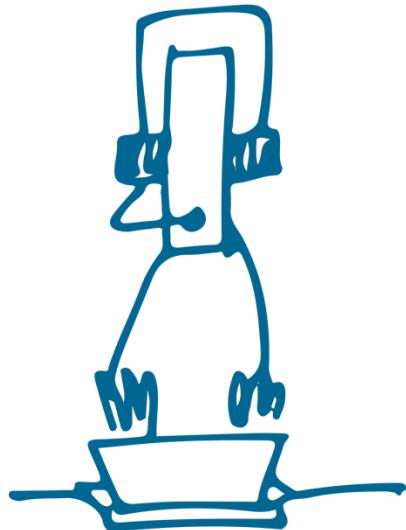
27. Data set XYZ	
<p>Goal:</p> <ul style="list-style-type: none"> motivation and awareness raising for RDM conveying the usefulness of RDM 	<p>Description:</p> <p>This method is suitable for</p> <ul style="list-style-type: none"> participants for whom RDM is new participants who are more critical homogeneous groups of researchers <p>The participants work in groups with a prepared data set. Their task is to assess the quality of this data and the data documentation. The results are documented in such a way that they can be easily presented.</p> <p>All groups present and explain their evaluation of the data set.</p>
<p>Duration:</p> <p>Selection: 1 minute Reflection: 2 minutes Presentation: 1-2 minutes per person</p>	<p>Required materials:</p> <ul style="list-style-type: none"> worksheet with documentation and data set containing appropriate quality deficiencies (alternative: the data sets or deficiencies can be varied between the groups to avoid repetition in the presentation) paper sheets for presentation of results pens

28. RDM slam	
<p>Goal:</p> <ul style="list-style-type: none"> presentation exercise argumentation exercise activation 	<p>Description:</p> <p>The participants work in three groups. Each group collects arguments and chooses one person to represent them at the slam.</p> <p>The chosen speakers are performing. The opposing groups evaluate the performance.</p>
<p>Duration:</p> <p>Preparation: 10 minutes Battle per person: 1 minute Voting: 1 minute</p>	<p>Required materials:</p> <ul style="list-style-type: none"> prize for the winner(s), for example, chocolate bar(s) presentation cards pens

29. Topic memory	
<p>Goal:</p> <ul style="list-style-type: none"> • consideration of participants' specific interests (and appreciation) • recording of further topics • further development of the subject matter 	<p>Description:</p> <p>The topics brought in by the participants, for example through questions that cannot be dealt with at the moment, are collected continuously.</p> <p>Topics that are dealt with at a later time will be removed.</p> <p>The memory is dynamic and is constantly updated.</p> <p>For topics not covered in the current workshop, i.e. topics still in the memory at the end of the event, a joint approach can be considered with the participants.</p>
<p>Duration:</p> <p>Continuously</p> <p>Fixed time frame (e.g. 20 minutes) at the end for „leftovers“ and how to deal with them.</p>	<p>Required materials:</p> <ul style="list-style-type: none"> • depending on your preference: • cards on clothesline, on bulletin board, on table, in box... • balloons • poster

21. Teaching script: Didactic methods							
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound
							Alternative Notes
Methods used	The participants recapitulate and apply the newly acquired knowledge	8	In groups of 2 people, the participants consider which methods have been used in the course of the workshop and which goal they each have pursued	Method: Inventory	Paper, pencils	Ex	Yes
	The participants repeat the methods and consolidate the new knowledge	15	T goes through the methods used together with the participants	Presentation	PPTX	In and ex	Yes
	The participants learn the S-M-L principle	2	T presents examples of different variants of various methods	Presentation	-	In	-
Method development	The group is divided into 4 teams. The teams should now develop their own method and write down the basics according to Schema-X. At the end the method is presented to the rest of the group.	30	The group is divided into 4 teams. The teams should now develop their own method and write down the basics according to Schema-X. At the end the method is presented to the rest of the group.	Explanation of the task: 2 min Group finding: 1 min Processing of the task: 10 min Performance: 12 min	Method: Schema-X	Flip charts with Schema-X, pens, pin board, pins or crepe tape	Ex
	The participants develop their own teaching method						Yes

Duration of the unit: 55 minutes



Unit 22: Feedback and farewell

Learning objectives

- Participants know the functions of evaluation.
- Participants know the measures used to collect feedback.
- Participants learn methods of giving feedback.
- Participants are able to repeat what they have learned.
- Participants transfer the acquired knowledge into their everyday work.

Key aspects

1. Functions of the evaluation
2. Aspects of the evaluation
3. Constructive use of feedback
4. Repetition
5. Farewell

Contents

1. Functions of the evaluation

Evaluation can be used for various purposes:¹¹²

¹¹² Reinmann, Gabi, Florian Alexander, Eva Häuptle and Johannes Metscher: *Wissenschaftliche Begleitung von Blended Learning in der Lehrerfortbildung : Konzept, Methodik, Ergebnisse, Erfahrungen und Empfehlungen am Beispiel "Intel® Lehren – Aufbaukurs Online*. Münster: Monsenstein und Vannerdat, 2009. <https://nbn-resolving.org/urn:nbn:de:bvb:384-opus4-10892>.

- Legitimacy – in certain situations it may be required by the employer, the user, the institution or the public to justify the purpose and the benefit of an event. In particular, for events that involve financial costs, justification to the sponsor may be necessary.
- Optimisation – Feedback from participants of an educational event can be helpful in order to improve the content or implementation of the training. The effectiveness of the teaching methods thus becomes apparent and the overall training service can be optimised.
- Check – similar to legitimacy, checking of whether or not a training event is worth the money and effort may be necessary. In so doing, it is also possible to check whether there is sufficient interest in the topic.
- Dialogue – the results of an evaluation could serve as a foundation for an open dialogue. If the results are disclosed and communicated, transparency is created.

2. Aspects of the evaluation

In the field of adult education, evaluation is used to record the input, teaching performance and benefits of qualification measures. According to Wesseler,¹¹³ evaluation is based on the following aspects:

- Learning performance of the participants (exams as an option)
- Teaching performance of the instructor (subject-related, didactic, and communicative performance)
- Characteristics of the curriculum (comprehensibility, appropriate level, sustainability...)
- General framework (resources, contexts, materials, locations, time frame, group size...)

3. Constructive use of feedback

Dealing with feedback, especially negative feedback, is a challenge. Taking up criticism properly is difficult when everything you have done is questioned and you have to admit to weaknesses and mistakes. However, constructive and honest feedback may contribute to improving your working methods.

4. Repetition

At the end of the workshop, the topics discussed should be briefly repeated. Simple methods like "Ideas out loud", "inventory" or "interim balance" can be used to consolidate the subject matter covered.

5. Farewell

To conclude the course formally, the teacher thanks the participants. You can also distribute study material and memory aids at this point. This marks the end of the workshop and the certificates of attendance are handed out.

Didactic methods and exercises

Methods:

- Mind map
 - Evaluation of workshops
- Fried Egg
 - Feedback on this workshop

¹¹³ Wesseler, Matthias: "Evaluation and Evaluationsforschung." In Rudolf Tippelt and Aiga von Hippel (ed.), *Handbuch Erwachsenenbildung/Weiterbildung*. 3rd, revised ed. Wiesbaden: VS Verlag für Sozialwissenschaften, 2011, pp. 1031–1048. https://doi.org/10.1007/978-3-531-91834-1_65.

- Inventory
 - What did you learn today and in the overall workshop? Name at least 12 keywords / terms / half sentences.
 - What are the benefits for your work?

Training materials

- Example of a feedback form for the Train-the-Trainer workshop on research data management
- Example of a certificate of attendance for participation in the Train-the-Trainer workshop on research data management
- Teaching script: Feedback and farewell

Additional sources

- Kirkpatrick, Donald L. and James D. Kirkpatrick: *Evaluating Training Programs – The four Levels*. 3rd ed. San Francisco, CA: Berrett-Koehler, 2006.
- Quilling, Eike and Hans J. Nicolini: *Erfolgreiche Seminargestaltung. Strategien und Methoden in der Erwachsenenbildung*. 2nd, revised ed. Wiesbaden: VS Verlag für Sozialwissenschaften, 2009.
- Reinmann, Gabi, Florian Alexander, Eva Häuptle and Johannes Metscher: Wissenschaftliche Begleitung von Blended Learning in der Lehrerfortbildung: Konzept, Methodik, Ergebnisse, Erfahrungen und Empfehlungen am Beispiel "Intel® Lehren – Aufbaukurs Online". Münster: Monsenstein und Vannerdat, 2009, <https://nbn-resolving.org/urn:nbn:de:bvb:384-opus4-10892>.
- Wessel, Matthias: "Evaluation und Evaluationsforschung." In *Handbuch Erwachsenenbildung/Weiterbildung* ed. by Rudolf Tippelt and Aiga von Hippel. 3rd ed. Wiesbaden: VS Verlag für Sozialwissenschaften, 2009. https://doi.org/10.1007/978-3-531-91834-1_65.

Train-the-trainer Workshop on Research Data Management

I found the following topics and contents particularly important:

- Introduction to research data management
- Introduction to didactics
- Organisational aspects
- "Experiencing" activating methods

The following topics or contents should have been dealt with in more detail or additionally:

The following topics or contents should have been dealt with in a shorter or less detailed way:

The relationship between lecture and personal work was...

- too lecture-intensive
- just right
- too much own work

The overall level of requirements of the workshop was...

- too high
- just right
- too low

The overall practical relevance of the workshop was...

- very well
- sufficient
- too small

After the workshop, do you feel able to prepare and conduct your own events?

- Yes
- Not yet
- I cannot judge yet

How did you like the workshop overall?

This is how workshops should be
Well

Was quite interesting
I expected more



Developed within the FD Mentor project
Funding code: 16FDM010 and 16FDM011
Project duration: 1. May 2017 - 30. April 2019

Contact: fdmentor@hu-berlin.de
Twitter: @fd_mentor
<https://hu.berlin/fdmentor>



This work is licensed under a
Creative Commons Attribution 4.0
International License
(<https://creativecommons.org/licenses/by/4.0/>).



How much knowledge and experience did you have in research data management before the workshop:

No knowledge and experience
Little knowledge and experience

Extended knowledge and experience
Extensive knowledge and experience

How did you hear about the event?

Website
Posting
Word of mouth

Twitter
E-mail
Other, namely:

Would you be interested in a more in-depth event on the topic of research data?

Yes

No

Please name your field(s) and level of study (e.g. MA, PhD, Prof.):

Please name your hero(s) and level of study (e.g. 1st, 2nd, 3rd year).

Other comments or suggestions for improvement:

Thank you very much for your feedback!





Participation certificate

Ms./Mr.

has successfully participated in the workshop

Train-the-trainer research data management

on [DD.MM.YYYY] and [DD.MM.YYYY]. The workshop comprised a total of 16 training hours.

Topics:

- Didactic approach
- Digital research data
- Research data policies
- Data management plan
- Order and structure
- Documentation and metadata
- Storage and backup
- Long-term archiving
- Access security
- Event framework
- Publication of research data
- Re-use of research data
- Legal aspects
- Institutional infrastructure
- Concept development
- Didactic methods
- Feedback and quality assurance

The program was developed within the project FDMentor. The project was funded by the German Federal Ministry of Education and Research (funding code 16FDM010 and 16FDM011).

[Place], [DD.MM.YYYY]

[Signature]

22. Teaching script: Feedback and farewell									
Component	Target	Time	Content	Working modes	Material	Inhalation/exhalation	Voices sound	Alternative	Notes
Evaluation	The participants know which functions evaluation has	2	The functions are presented	Presentation	PPTX	In	-		
	The participants consider what criteria there are for feedback	5	Participants consider which criteria/dimensions can be used to evaluate a training (2 min) Participants pin their cards in a grouped manner and T reads out the result (3 min)	Individual work & plenum	Moderation cards, pens	Ex	-		
	The participants get to know the dimensions in which feedback is acquired	2	The dimensions are introduced	Presentation	PPTX	In	-		
Recap	The participants remember the learning contents	12	The participants recall what they have learned in the workshop	Method: Inventory	Paper A3, pencils	Ex	Yes	S: Method: Ideas out loud	At least 12 keywords
	The participants transfer what they have learned for their everyday work	5	The participants answer the question: What can you use for yourself?	Ideas out loud	-	Ex	Yes	Harvest sheet	
Handouts & homework	The participants consolidate the learning contents, reproduce and transfer it to their everyday work	1	The participants receive worksheets with homework	-	Work sheets	In	-		

		Participants provide initial, more general feedback on the workshop. The participants write general feedback on cards, pin the cards to the pin board at the perceived distance to the yellow of the egg and explain in one sentence why they chose that position Writing down: 4 min Presentation: 5 min	Method: Fried egg	Pin board; prepared large sheet with picture of egg yolk; cards for participants and thick pencils	Ex	Yes	Method: Five-Finger-Feedback	Point out: Statement should be limited to ONE sentence
Feedback	The participants evaluate the workshop	9						
Farewell	Ending	2	T thanks the participants for their participation and feedback. T says goodbye to the participants. The participants can give further input if they will think of something later. The certificates of participation are handed out		-	-	In	-
Feedback	T receives detailed feedback with questionnaire	12	The participants fill in the questionnaire	-	Prepared questionnaire (paper and/or online)	-	-	Provide sufficient time

Duration of the unit: 50 minutes

General resources on the topic

Bezjak, Sonja, April Clyburne-Sherin, Philipp Conzett, Pedro Fernandes, Edit Görögh, Kerstin Helbig, Bianca Kramer et al.: Open Science Training Handbook. Version 1.0 Zenodo, 2018. <http://doi.org/10.5281/zenodo.1212496>.

Briney, Kristin: *Data Management for Researchers: Organize, maintain and share your data for research success*. Exeter: Pelagic Publishing, 2015.

Büttner, Stephan, Hans-Christoph Hobohm and Lars Müller (ed.): *Handbuch Forschungsdatenmanagement*. Bad Honnef: Bock+Herchen, 2011. <https://doi.org/10.34678/opus4-208>.

CESSDA Training Working Group: "Data Management Expert Guide." Bergen: CESSDA ERIC, 2017-2018. Zugriff am 29.07.2020. <https://www.cessda.eu/DMGuide>.

Christodoulou, Michail, Stefanos Kachrilas, Ahmed Dina, Andreas Bourdoumis, Junaid Masood, Noor Buchholz and Athanasios Papatsoris. "How to conduct a successful workshop: The trainee's perspective." *Arab Journal of Urology* 12, 1 (2014): pp. 12–14. <https://doi.org/10.1016/j.aju.2013.08.004>.

Clare, Connie, Maria Cruz, Elli Papadopoulou, James Savage, Marta Teperek, Yan Wang, Iza Witkowska et al.: *Engaging Researchers with Data Management: The Cookbook*. Open Book Publishers, 2019, <https://doi.org/10.11647/OPB.0185>.

Cohn, Ruth: Von der Psychoanalyse zur themenzentrierten Interaktion: von der Behandlung einzelner zu einer Pädagogik für alle. Stuttgart: Klett, 1975.

Corti, Louise, Veerle Van den Eynden, Libby Bishop and Matthew Woppard: *Managing and Sharing Research Data: A Guide to Good Practice*. Los Angeles, CA: Sage, 2014.

Döring, Klaus W.: *Handbuch Lehren und Trainieren in der Weiterbildung*. Weinheim: Beltz, 2008.

FDMentor and DINI/nestor-AG Forschungsdaten: "Materialkatalog zum Forschungsdatenmanagement." Version 1.0. Zenodo, 2018. <http://doi.org/10.5281/zenodo.1209284>.

Forschungsdaten.info: "Forschung und Daten managen." Last accessed 29.07.2020, <https://www.forschungsdaten.info>.

Forschungsdaten.org: „Hauptseite.“ Last accessed 29.07.2020, <http://www.forschungsdaten.org/index.php/Hauptseite>.

Groß, Harald: *Munterbrechungen. 22 aktivierende Auflockerungen für Seminare und Sitzungen*. Berlin: Schilling, 2010.

Groß, Harald: *Muniterrichtsmethoden. 22 weitere aktivierende Methoden für die Seminarpraxis*. Vol. 2. Berlin: Schilling, 2014.

Groß, Harald, Betty Boden and Nikolaas Boden: *Muniterrichtsmethoden. 22 aktivierende Lehrmethoden für die Seminarpraxis*. 3rd ed. Berlin: Schilling, 2011.

Haussmann, Martin and Holger Scholz: *bikablo. Das Trainerwörterbuch der Bildsprache*. 2nd ed. Eichenzell: Neuland, 2007.

Haussmann, Martin and Holger Scholz: *bikablo 2.0: neue Bilder für Meeting, Training & Learning*. Eichenzell: Neuland, 2009.

Jensen, Uwe, Sebastian Netscher and Katrin Weller (ed.): *Forschungsdatenmanagement sozialwissenschaftlicher Umfragedaten. Grundlagen und praktische Lösungen für den Umgang mit quantitativen Forschungsdaten*. Opladen: Verlag Barbara Budrich, 2019, <https://doi.org/10.3224/84742233>.

Klein, Zamyat M.: *Kreative Seminarmethoden. 100 kreative Methoden für erfolgreiche Seminare*. 8th ed. Offenbach: GABAL, 2003.

Pavelin, Katrina, Sangya Pundir and Jennifer A. Cham. "Ten Simple Rules for Running Interactive Workshops." *PLOS Computational Biology* 10, 2 (2014): p. e1003485. <https://doi.org/10.1371/journal.pcbi.1003485>.

Pryor, Graham: *Managing Research Data*. London: Facet Publishing, 2012.

Ray, Joyce M. (ed.): *Research data management: practical strategies for information professionals*. West Lafayette, IN: Purdue University Press, 2014.

Rice, Robin and John Southall: *The Data Librarian's Handbook*. London: Facet Publishing, 2016.

Szepansky, Wolf-Peter: *Souverän Seminare leiten*. Bielefeld: Bertelsmann, 2006.

Train2Dacar: "Research Data Management. E-Learning Platform." Last accessed 29.07.2020, <https://campus.hesge.ch/researchdatamanagement/>.

University of Edinburgh: "MANTRA: Research Data Management Training. Last accessed 29.07.2020, <https://mantra.edina.ac.uk>.