



Heike Neuroth, Stefan Strathmann,  
Achim OBwald, Jens Ludwig (Eds.)

# Digital Curation of Research Data

## Experiences of a Baseline Study in Germany

### Foreword

**vwh** Verlag Werner Hülsbusch  
Fachverlag für Medientechnik und -wirtschaft



Universitätsverlag Göttingen

Neuroth, Strathmann, Oßwald, Ludwig (Eds.)  
Digital Curation of Research Data

**Kontakt / Contact:** [editors@langzeitarchivierung.de](mailto:editors@langzeitarchivierung.de)

c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen, Dr. Heike Neuroth, Forschung und Entwicklung, Papendiek 14, 37073 Göttingen

c/o Göttingen State and University Library, Dr. Heike Neuroth, Research and Development, Papendiek 14, 37073 Göttingen, Germany

**Heike Neuroth, Stefan Strathmann,  
Achim Oßwald, Jens Ludwig (Eds.)**

# **Digital Curation of Research Data**

**Experiences of a Baseline Study  
in Germany**

**vwh**

Verlag Werner Hülsbusch  
Fachverlag für Medientechnik und -wirtschaft

## Digital Curation of Research Data

Herausgegeben von Heike Neuroth, Stefan Strathmann, Achim Oßwald und Jens Ludwig · im Rahmen des Kooperationsverbundes nestor – Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit digitaler Ressourcen für Deutschland · <http://www.langzeitarchivierung.de/>

Edited by Heike Neuroth, Stefan Strathmann, Achim Oßwald and Jens Ludwig · within the context of nestor – Network of Expertise in the Long-Term Storage of Digital Resources for Germany · <http://www.langzeitarchivierung.de/>

### *Bibliografische Information der Deutschen Nationalbibliothek*

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://www.d-nb.de> abrufbar.

### *Bibliographic information of the German National Library*

The German National Library lists this publication in the German National Bibliography; detailed bibliographic data is available online at <http://www.d-nb.de>.

Die Inhalte dieses Buches stehen auch als Onlineversion über die Website von nestor zur Verfügung / This work is available as an Open Access version at the nestor website: <http://nestor.sub.uni-goettingen.de/bestandsaufnahme/index.php?lang=en>

Die digitale Version dieses Werkes ist unter Creative Commons Namensnennung 3.0 lizenziert / The digital version of this work is licensed under a Creative Commons Attribution 3.0 Unported License <http://creativecommons.org/licenses/by/3.0/deed.en>

CC - BY 

Einfache Nutzungsrechte liegen beim Verlag Werner Hülsbusch, Glückstadt.  
The Verlag Werner Hülsbusch, Glückstadt, owns rights of use for the printed version of this work.

**vwh** Verlag Werner Hülsbusch  
Fachverlag für Medientechnik und -wirtschaft

© Verlag Werner Hülsbusch, Glückstadt, 2013 · <http://www.vwh-verlag.de>

in Kooperation mit dem Universitätsverlag Göttingen  
in cooperation with the Universitätsverlag Göttingen

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen, Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und als solche den gesetzlichen Bestimmungen unterliegen.

All trademarks used in this work are the property of their respective owners.

Printed in Poland · ISBN: 978-3-86488-054-4

# Content

<b>Foreword</b>	<b>7</b>
<i>Heike Neuroth, Stefan Strathmann, Achim Oßwald, Jens Ludwig</i>	
<b>1 Digital Curation of Research Data: An Introduction</b>	<b>9</b>
<i>Achim Oßwald, Heike Neuroth, Regine Scheffel</i>	
<b>2 Status of Discussion and Current Activities: National Developments</b>	<b>18</b>
<i>Stefan Winkler-Nees</i>	
2.1 Research Organizations	19
2.2 Recommendations and Policies	22
2.3 Information Infrastructure Institutions	28
2.4 Funding Organizations	33
<b>3 Status of Discussion and Current Activities: The International Perspective</b>	<b>37</b>
<i>Stefan Strathmann</i>	
3.1 International Organizations	37
3.1.1 United Nations Educational, Scientific and Cultural Organization (UNESCO)	38
3.1.2 Organisation for Economic Co-Operation and Development (OECD)	38
3.1.3 European Union (EU)	40
3.1.4 World Health Organization (WHO)	41
3.1.5 Knowledge Exchange	41
3.2 Model Realizations	42
3.2.1 National Science Foundation (NSF)	42
3.2.2 Australian National Data Service (ANDS)	43
<b>4 Methodology: Subject of the Study</b>	<b>46</b>
<i>Heike Neuroth</i>	
4.1 Structure of this Volume	47
4.2 Key questions for mapping research disciplines	48

4.3	Introduction to the Research Area	48
4.3.1	Background	49
4.3.2	Cooperative Structures	49
4.3.3	Data and Metadata	49
4.3.4	Internal Organization	51
4.3.5	Perspectives and Visions	52
<b>5</b>	<b>Summary and Interpretation</b>	<b>54</b>
	<i>Jens Ludwig</i>	
5.1	Cooperative Structures	55
5.2	Data and Metadata	58
5.3	Internal organization	65
5.4	Perspectives and Visions	67
<b>6</b>	<b>Implications and Recommendations on Research Data Curation</b>	<b>69</b>
	<i>Heike Neuroth, Achim Oßwald, Uwe Schwiegelshohn</i>	
	<b>References</b>	<b>79</b>
	<b>Abbrevations</b>	<b>87</b>
	<b>Directory of Authors</b>	<b>91</b>

# Foreword

The relevance of research data today and for the future is well documented and discussed, in Germany as well as internationally. In addition, more and more policy makers are aware of the meaning of research data and the possibilities to access, share, and re-use them. Recently the government of the United Kingdom decided that publicly funded research publications and research data must be freely available and accessible to the public in all situations in which there are no copyright issues or other legal aspects that would prevent it. It is expected that the European Commission will publish similar requirements in the context of the next framework program Horizon 2020 starting at the end of 2013. The newly funded Research Data Alliance (RDA) is one example of the increasing interest in these topics worldwide. Ensuring that research data are accessible, sharable, and re-usable over time makes several further steps possible:

- Research data are documented and could therefore be validated.
- Research data could be the basis for other and new research questions, since they could be an integral part of the (digital) research lifecycle from the very beginning.
- Research data could be re-analysed by using new, innovative digital methods which were unknown at the moment of data acquisition.
- Research data could be used by other disciplines, therefore encouraging interdisciplinary research.

For all of these reasons, it is essential that research data are curated, which means that they are kept accessible and interpretable over time. A standardized questionnaire was developed in order to understand whether the approaches and methods of dealing with research data within the academic disciplines are different or whether there are similarities in terms of solutions as well as challenges and problems. This questionnaire was distributed to representatives from those disciplines in Germany that were identified as familiar with or already expert in research data curation.

The results of this survey have been published in German in 2012 in the handbook *“Langzeitarchivierung von Forschungsdaten – Eine Be-*



*standsaufnahme*".<sup>1</sup> This publication is the English-language translation of the main chapters of this handbook. The original German version also contains detailed analyses of the situation regarding the curation of research data of eleven disciplines ranging from humanities and social sciences to the natural sciences and medicine. Colleagues from these eleven disciplines were asked to describe the state-of-the-art regarding their methods of handling and experiences with research data curation in the questionnaire. These chapters have not been included in this English publication.

The last chapters of this English-language publication analyse the responses from all disciplines, compare the similarities as well as differences, and conclude with some overall implications and recommendations for stakeholders, policy-makers, key-players, and scholarly societies.

The editors of this volume, as well as the additional editors of the German version, have been working together closely for many years, such as in the context of nestor – the German competence network for digital preservation. When we started the data curation discussion in Germany at universities and in research disciplines, the terms “long-term preservation” and “digital preservation” were already established in Germany. Today we prefer the term “digital curation”. As a result we used these earlier terms in cases where we are following the original German handbook or for citations. In all other cases the more modern term “data curation” is used.

We would like to express our special thanks to Hanna-Lena Stolz and Dr. Kathleen M. Smith for their valuable support in translating the main chapters of the German handbook. Without their help, we would never have been able to share our experiences, thoughts, and conclusions on this important and urgent topic. We are eager for exchange with the broader data community, across geographic and linguistic borders, across academic disciplines, across funding agencies, and many other levels.

With best regards,

*Heike Neuroth, Stefan Strathmann, Achim Oßwald and Jens Ludwig*

---

<sup>1</sup> Neuroth et al. (2012).