

Das DINI-Positionspapier „Forschungsdaten“ im nationalen und internationalen Kontext

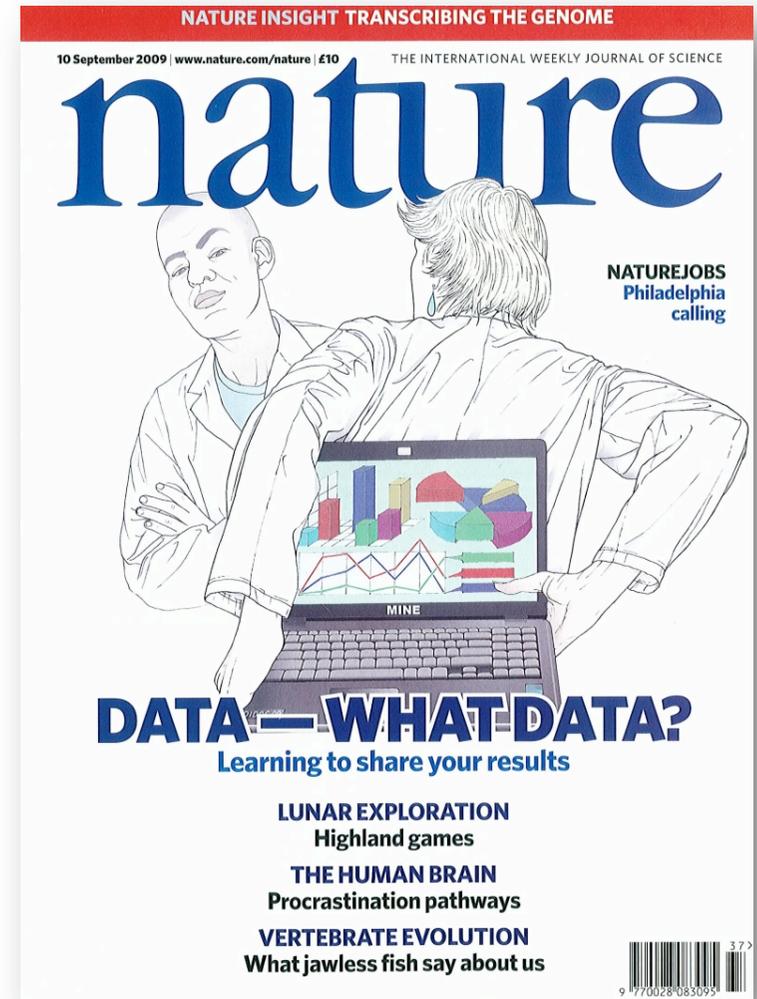
Heinz Pampel, Helmholtz-Gemeinschaft

10. DINI-Jahrestagung

Kassel, 30.09.2009

AGENDA

- Relevanz für die Helmholtz-Gemeinschaft
- Herausforderungen
- Positionen
 - Wissenschaftspolitik
 - Disziplinen
 - Infrastruktureinrichtungen
- DINI-Positionspapier Forschungsdaten
 - Handlungsfelder
- Fazit



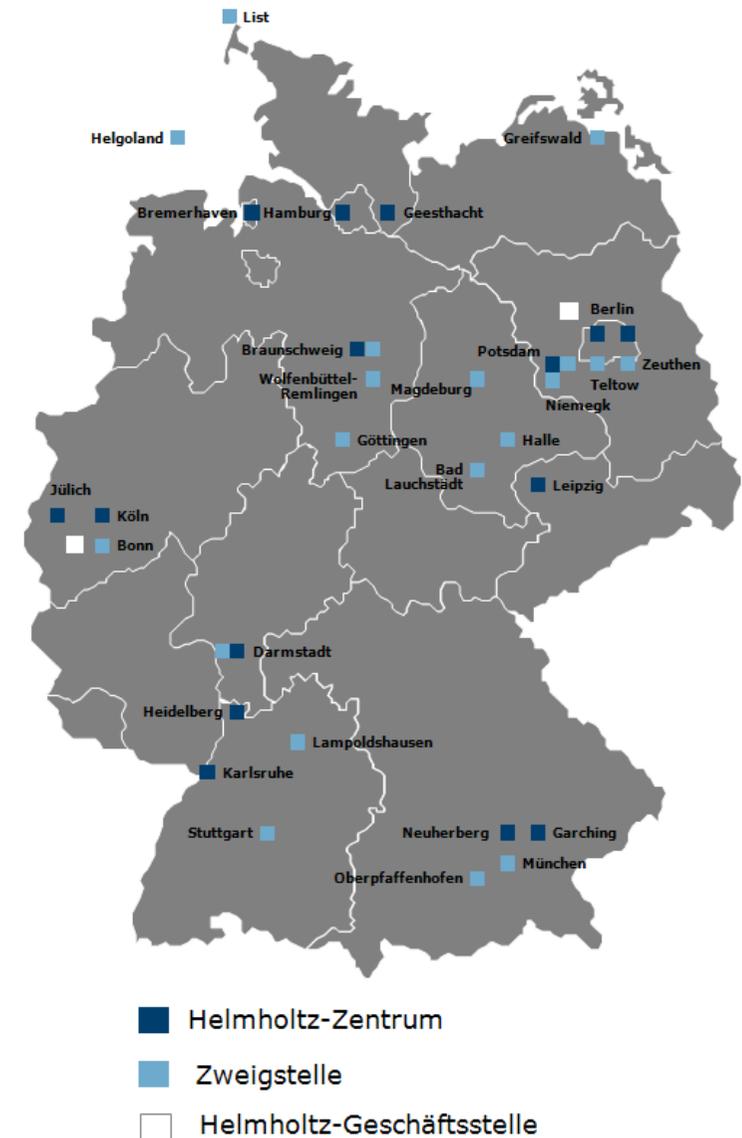
[1]

HELMHOLTZ-GEMEINSCHAFT

- 16 naturwissenschaftlich-technische und medizinisch-biologische Forschungszentren

- Mission:

„Wir leisten Beiträge zur Lösung großer und drängender Fragen von Gesellschaft, Wissenschaft und Wirtschaft durch strategisch-programmatisch ausgerichtete Spitzenforschung in den Bereichen Energie, Erde und Umwelt, Gesundheit, Schlüsseltechnologien, Struktur der Materie, Luftfahrt, Raumfahrt und Verkehr.“



RELEVANZ FÜR DIE HELMHOLTZ-GEMEINSCHAFT

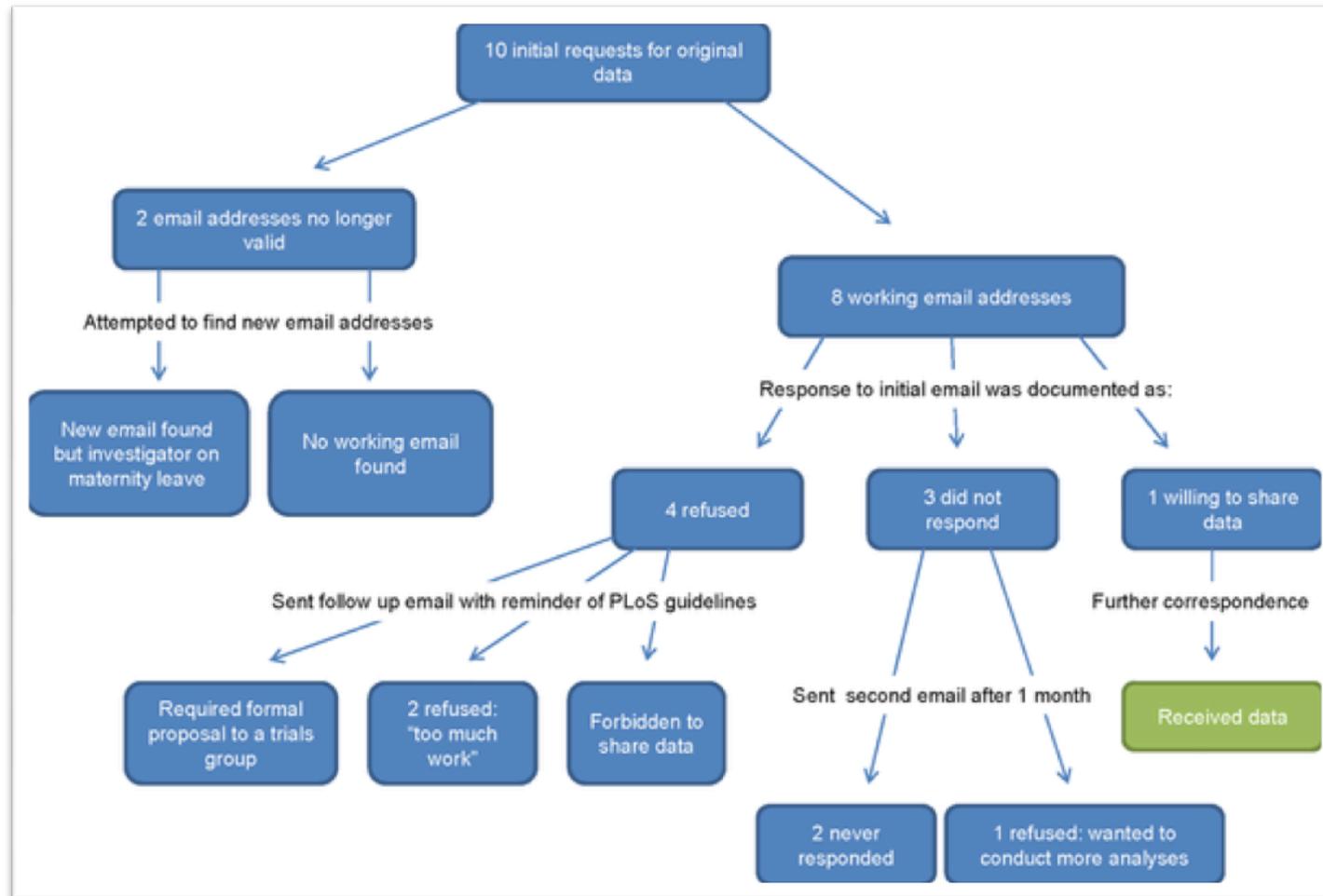
- International bedeutende Großgeräte
- Datenintensive Forschung
- Hoher technischer, organisatorischer und finanzieller Aufwand
- Weltdatenzentren (WDC)
- Datenzentren, Rechenzentren, Bibliotheken
- Erstunterzeichner der „Berliner Erklärung“
- Helmholtz Open Access Projekt
- Beteiligung an nationalen und internationalen Initiativen



HERAUSFORDERUNGEN

- Savage and Vickers:

Empirical Study of Data Sharing by Authors Publishing in PLoS Journals, 2009



[2]

POSITIONEN: WISSENSCHAFTSPOLITIK

- Deutsche Forschungsgemeinschaft (DFG), 1998

Empfehlung 7

Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Institution, wo sie entstanden sind, für zehn Jahre aufbewahrt werden.

Erläuterungen

Auf die Aufzeichnungen später zurückgreifen zu können, ist schon aus Gründen der Arbeitsökonomie in einer Gruppe ein zwingendes Gebot. Noch wichtiger wird dies, wenn veröffentlichte Resultate von anderen angezweifelt werden.



[3]

POSITIONEN: WISSENSCHAFTSPOLITIK

- „Berlin Declaration“, 2003

Establishing open access as a worthwhile procedure ideally requires the active commitment of each and every individual producer of scientific knowledge and holder of cultural heritage. Open access contributions include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material.

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities

Preface

The Internet has fundamentally changed the practical and economic realities of distributing scientific knowledge and cultural heritage. For the first time ever, the Internet now offers the chance to constitute a global and interactive representation of human knowledge, including cultural heritage and the guarantee of worldwide access.

We, the undersigned, feel obliged to address the challenges of the Internet as an emerging functional medium for distributing knowledge. Obviously, these developments will be able to significantly modify the nature of scientific publishing as well as the existing system of quality assurance.

In accordance with the spirit of the Declaration of the Budapest Open Access Initiative, the ECHO Charter and the Bethesda Statement on Open Access Publishing, we have drafted the Berlin Declaration to promote the Internet as a functional instrument for a global scientific knowledge base and human reflection and to specify measures which research policy makers, research institutions, funding agencies, libraries, archives and museums need to consider.

Goals

Our mission of disseminating knowledge is only half complete if the information is not made widely and readily available to society. New possibilities of knowledge dissemination not only through the classical form but also and increasingly through the open access paradigm via the Internet have to be supported. We define open access as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community.

In order to realize the vision of a global and accessible representation of knowledge, the future Web has to be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible.

Definition of an Open Access Contribution

Establishing open access as a worthwhile procedure ideally requires the active commitment of each and every individual producer of scientific knowledge and holder of cultural heritage. Open access contributions include original scientific research results, raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material.

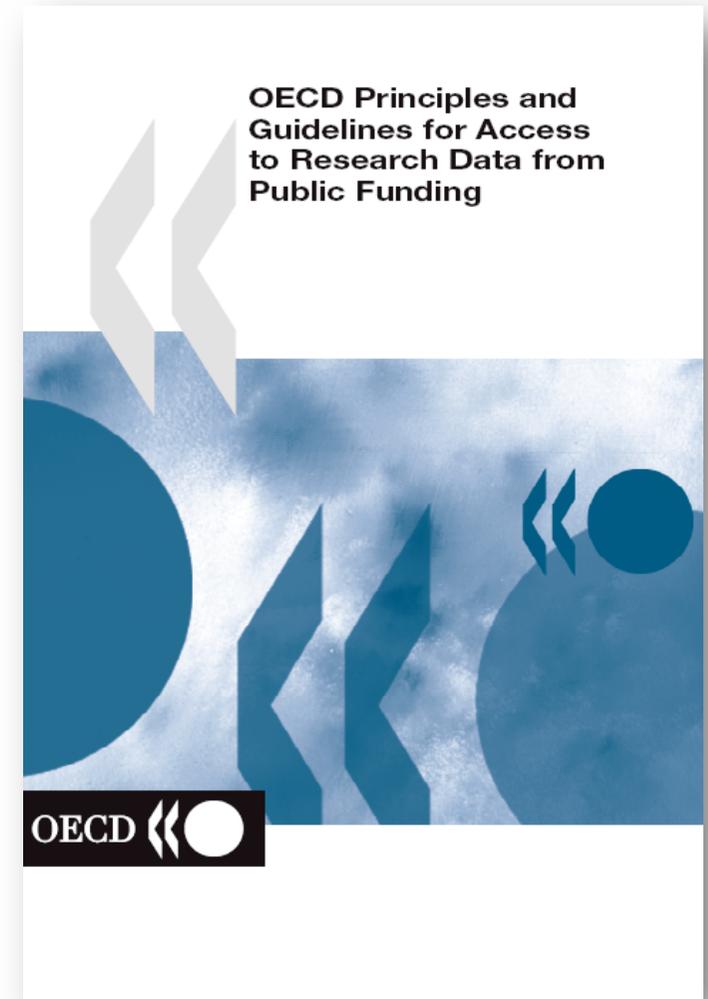
[4]

POSITIONEN: WISSENSCHAFTSPOLITIK

- OECD, 2007

These Principles and Guidelines [...] provide broad policy recommendations to the governmental science policy and funding bodies of member countries on access to research data from public funding. They are intended to promote data access and sharing among researchers, research institutions, and national research agencies, while at the same time, recognising and taking into account, the various national laws, research policies and organisational structures of member countries.

The ultimate goal of these Principles and Guidelines is to improve the efficiency and effectiveness of the global science system. They are not intended to hinder its development with onerous obligations and regulations or impose new costs on national science systems.



[5]

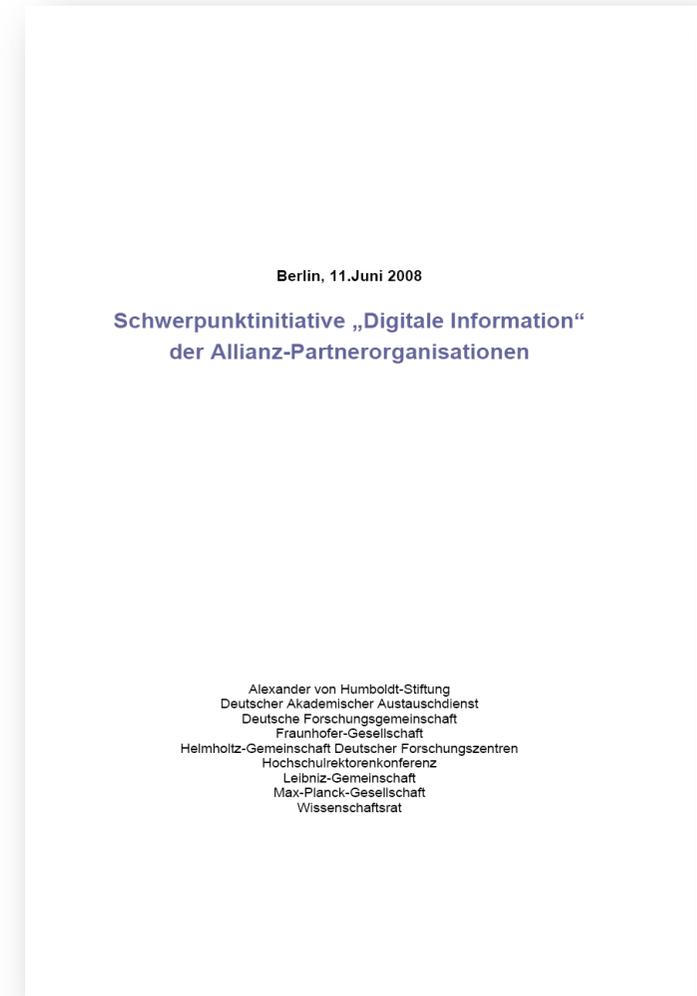
POSITIONEN: WISSENSCHAFTSPOLITIK

- Schwerpunktinitiative „Digitale Information“, 2008

Es ist unbestreitbar, dass viele dieser Daten nach einer relativ kurzen Phase der Auswertung durch Einzelne oder kleine Gruppen dem Vergessen oder gar dem Verfall ausgesetzt sind. Hier sehen alle Wissenschaftseinrichtungen einen dringenden Handlungsbedarf hinsichtlich der systematischen Sicherung, Archivierung und Bereitstellung dieser Daten für die Nachnutzung durch Dritte.

Die Aktivitäten der Allianz-Initiative sind auf drei Bereiche gerichtet:

- *Primärdaten-Policy*
- *Zusammenarbeit zwischen Fachwissenschaftlern und Informationswissenschaftlern anzustoßen*
- *Vernetzte Repositorien- und Archivstrukturen für Forschungsprimärdaten*



[6]

POSITIONEN: WISSENSCHAFTSPOLITIK

■ EUROHORCs / ESF, 2009

The collection of research data is a huge investment. Permanent access to such data, if quality controlled and in interoperable formats, will allow better use to be made of this investment because it allows other researchers to (re)use them. Furthermore it allows re-analysis and could play a role in ensuring research integrity.

Permanent preservation and open access, such as promoted by the Alliance for Permanent Access, will be the rule for repositories. EUROHORCs and ESF Member Organisations will address how to best promote and ensure such permanent access to data generated with their funding.

The cover of the report features the logos of the European Science Foundation (ESF) and EUROHORCs. The ESF logo includes the text 'EUROPEAN SCIENCE FOUNDATION' and 'SETTING SCIENCE AGENDAS FOR EUROPE'. The EUROHORCs logo includes the text 'EUROHORCs' and 'EUROPEAN HEADS OF RESEARCH COUNCILS'. The title of the report is 'EUROHORCs and ESF Vision on a Globally Competitive ERA and their Road Map for Actions'. Below the title is a 'Contents' section with three items: '1•Foreword', '3•The EUROHORCs and ESF Vision', and '25•Annex 1 EUROHORCs-ESF Task Force Members'. The 'Foreword' section is highlighted. The foreword text discusses the role of EUROHORCs and ESF in shaping a European Research Area of Excellence and the importance of a Road Map for Actions to help construct the ERA. It mentions the approval of a Vision Statement in April 2008 and the role of the EUROHORCs-ESF Task Force. The foreword is signed by Professor Ian Halliday, President of ESF, and Professor Dieter Imboden, President of EUROHORCs. The website addresses www.eurohorcs.org and www.esf.org are listed at the bottom.

[7]

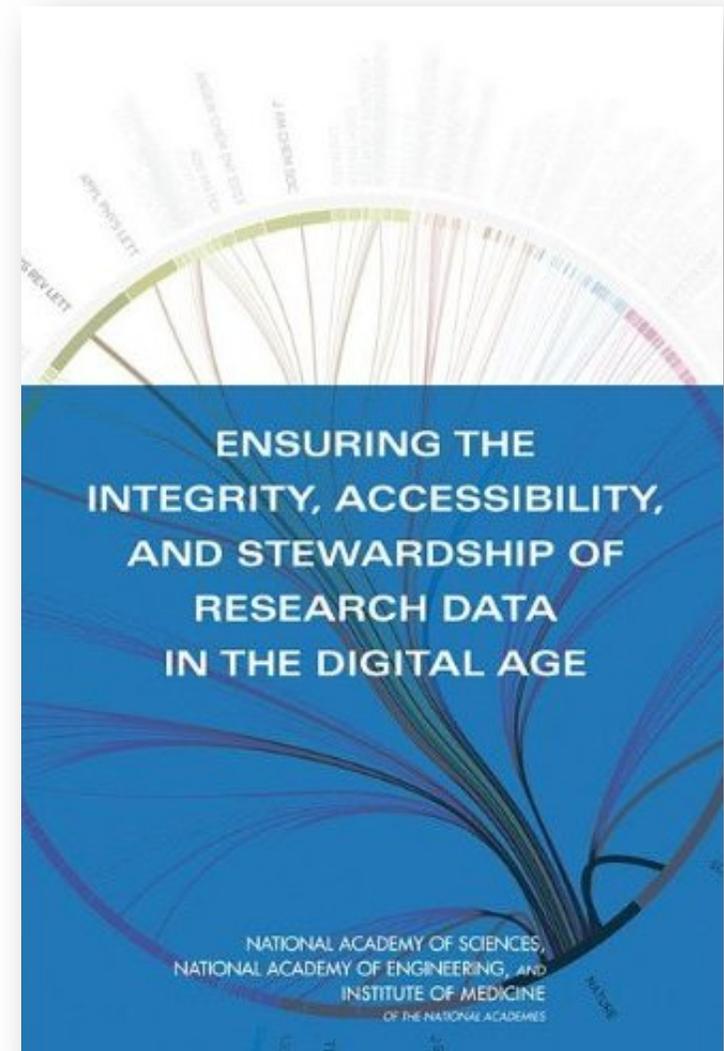
POSITIONEN: WISSENSCHAFTSPOLITIK

- National Academy of Sciences, 2009

Data Integrity Principle: Ensuring the integrity of research data is essential for advancing scientific, engineering, and medical knowledge and for maintaining public trust in the research enterprise. Although other stakeholders in the research enterprise have important roles to play, researchers themselves are ultimately responsible for ensuring the integrity of research data.

Data Access and Sharing Principle: Research data, methods, and other information integral to publicly reported results should be publicly accessible.

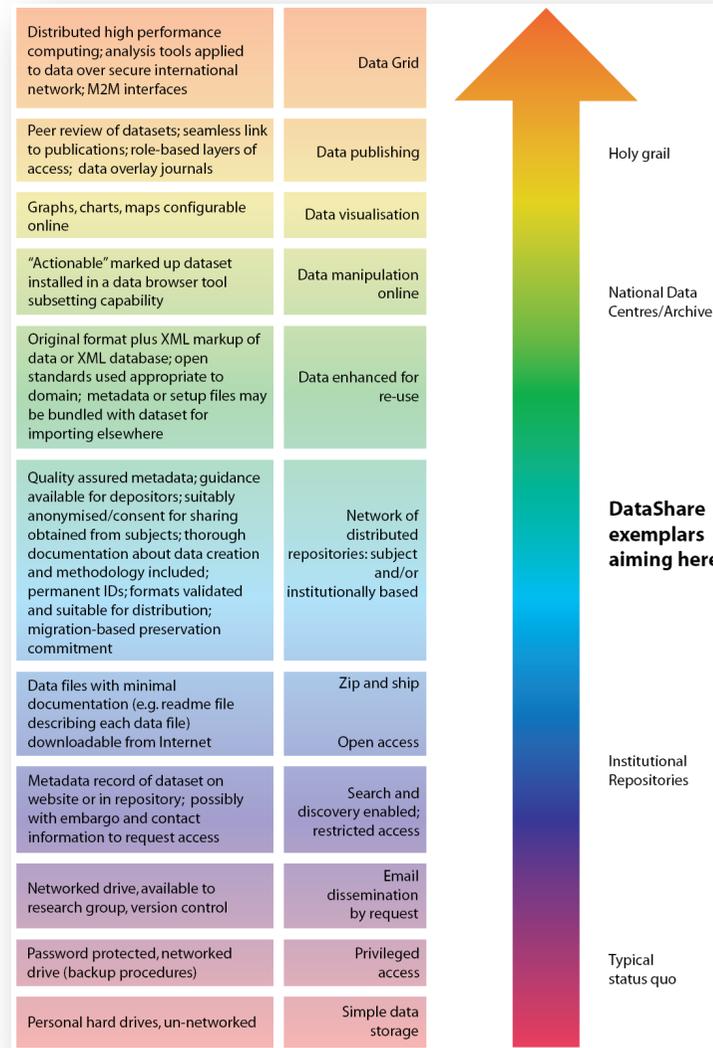
Data Stewardship Principle: Research data should be retained to serve future uses. Data that may have long-term value should be documented, referenced, and indexed so that others can find and use them accurately and appropriately.



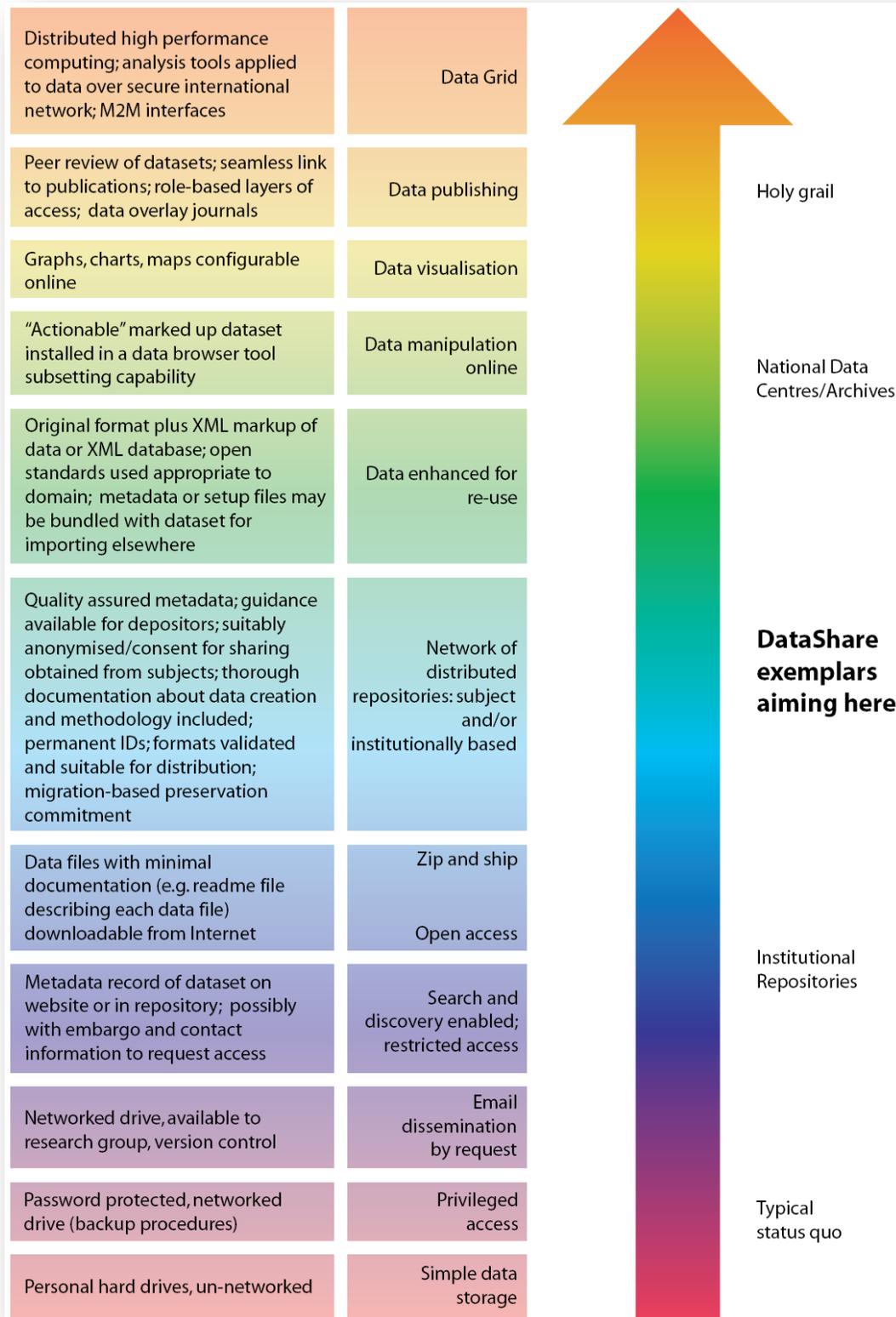
[8]

POSITIONEN: DISZIPLINEN

- Rice: Data Sharing Continuum, 2007



[9]



POSITIONEN: INFRASTRUKTUREINRICHTUNGEN

- RIN: Stewardship of digital research data, 2008

- The principles

1. *Roles and responsibilities*
2. *Standards and quality assurance*
3. *Access, usage and credit*
4. *Benefits and cost-effectiveness*
5. *Preservation and sustainability*

Stewardship of digital research data: a framework of principles and guidelines

Responsibilities of research institutions and funders, data managers, learned societies and publishers

January 2008



www.rin.ac.uk

[10]

POSITIONEN: INFRASTRUKTUREINRICHTUNGEN

- ARL: E-Science Talking Points, 2008
 - *Data management, including collection, organization, description, curation, archiving, and dissemination.*
 - *Creation of new data- and scholarship-based electronic resources for university and/or public use.*
 - *Development of new models, standards, and architectures for various aspects of data management, description, etc.*
 - *Building accessible linkages between all the components and stages of research, from data to researchers to publications.*
 - *Bridging institutional hierarchies and departmental divisions in service of interdisciplinary initiatives.*

E-SCIENCE TALKING POINTS FOR ARL DEANS AND DIRECTORS

by Elisabeth Jones, University of Washington

with contributions from
Wendy Lougee, University of Minnesota
Neil Rambo, University of Washington
Eric Celeste, Consultant to the ARL E-Science Working Group
and guidance from other members of the ARL E-Science Working Group

October 24, 2008
Association of Research Libraries
<http://www.arl.org/rtl/e:science/>



ARL

[11]

POSITIONEN: INFRASTRUKTUREINRICHTUNGEN

■ Brussels Declaration on STM Publishing, 2007

7. Raw research data should be made freely available to all researchers. Publishers encourage the public posting of the raw data outputs of research. Sets or sub-sets of data that are submitted with a paper to a journal should wherever possible be made freely accessible to other scholars.

BRUSSELS DECLARATION ON STM PUBLISHING

by the international scientific, technical and medical (STM) publishing community as represented by the individual publishing houses and publishing trade associations, who have indicated their assent below.

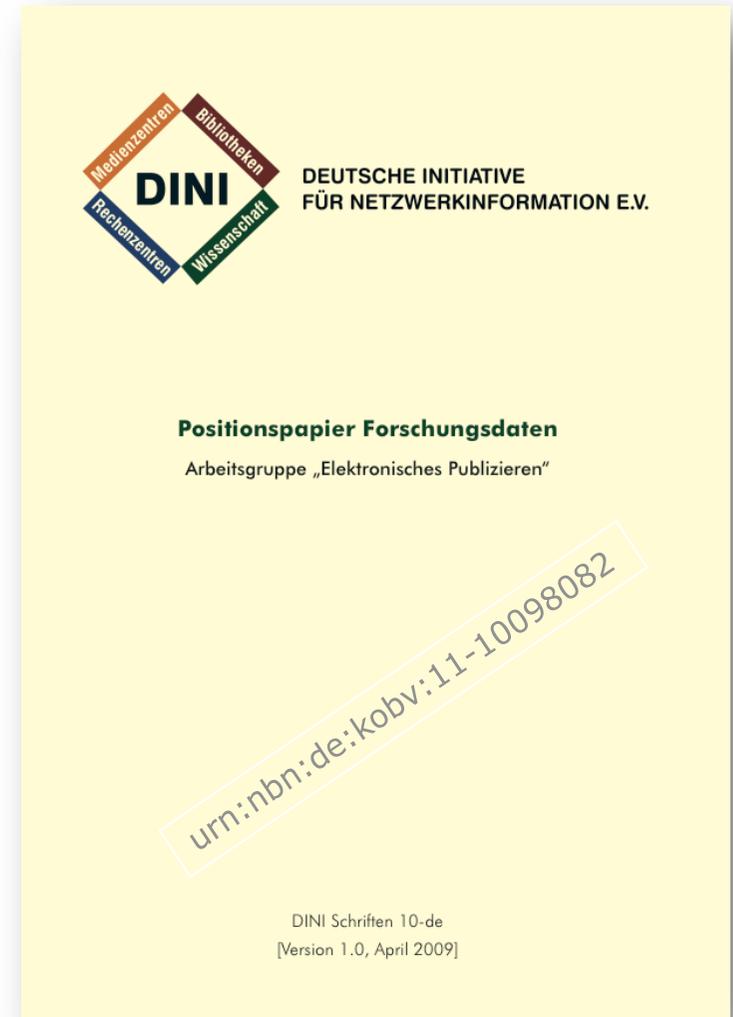
Many declarations have been made about the need for particular business models in the STM information community. STM publishers have largely remained silent on these matters as the majority are agnostic about business models: what works, works. However, despite very significant investment and a massive rise in access to scientific information, our community continues to be beset by propositions and manifestos on the practice of scholarly publishing. Unfortunately the measures proposed have largely not been investigated or tested in any evidence-based manner that would pass rigorous peer review. In the light of this, and based on over ten years experience in the economics of online publishing and our long-standing collaboration with researchers and librarians, we have decided to publish a declaration of principles which we believe to be self-evident.

1. **The mission of publishers is to maximise the dissemination of knowledge through economically self-sustaining business models.** We are committed to change and innovation that will make science more effective. We support academic freedom: authors should be free to choose where they publish in a healthy, undistorted free market
2. **Publishers organise, manage and financially support the peer review processes of STM journals.** The imprimatur that peer-reviewed journals give to accepted articles (registration, certification, dissemination and editorial improvement) is irreplaceable and fundamental to scholarship
3. **Publishers launch, sustain, promote and develop journals for the benefit of the scholarly community**
4. **Current publisher licensing models are delivering massive rises in scholarly access to research outputs.** Publishers have invested heavily to meet the challenges of digitisation and the annual 3% volume growth of the international scholarly literature, yet less than 1% of total R&D is spent on journals
5. **Copyright protects the investment of both authors and publishers.** Respect for copyright encourages the flow of information and rewards creators and entrepreneurs
6. **Publishers support the creation of rights-protected archives that preserve scholarship in perpetuity**
7. **Raw research data should be made freely available to all researchers.** Publishers encourage the public posting of the raw data outputs of research. Sets or sub-sets of data that are submitted with a paper to a journal should wherever possible be made freely accessible to other scholars
8. **Publishing in all media has associated costs.** Electronic publishing has costs not found in print publishing. The costs to deliver both are higher than print or electronic only. Publishing costs are the same whether funded by supply-side or demand-side models. If readers or their agents (libraries) don't fund publishing, then someone else (e.g. funding bodies, government) must
9. **Open deposit of accepted manuscripts risks destabilising subscription revenues and undermining peer review.** Articles have economic value for a considerable time after publication which embargo periods must reflect. At 12 months, on average, electronic articles still have 40-50% of their lifetime downloads to come. Free availability of significant proportions of a journal's content may result in its cancellation and therefore destroy the peer review system upon which researchers and society depend
10. **"One size fits all" solutions will not work.** Download profiles of individual journals vary significantly across subject areas, and from journal to journal

[12]

POSITIONSPAPIER FORSCHUNGSDATEN

- Disziplinarität
 - Forschungsdaten variieren nach Disziplin
 - Umfassendes Verständnis der jeweiligen Daten vonnöten
 - Ausrichtung aller Aktivitäten an den Anforderungen der jeweiligen Disziplin



[13]

POSITIONSPAPIER FORSCHUNGSDATEN

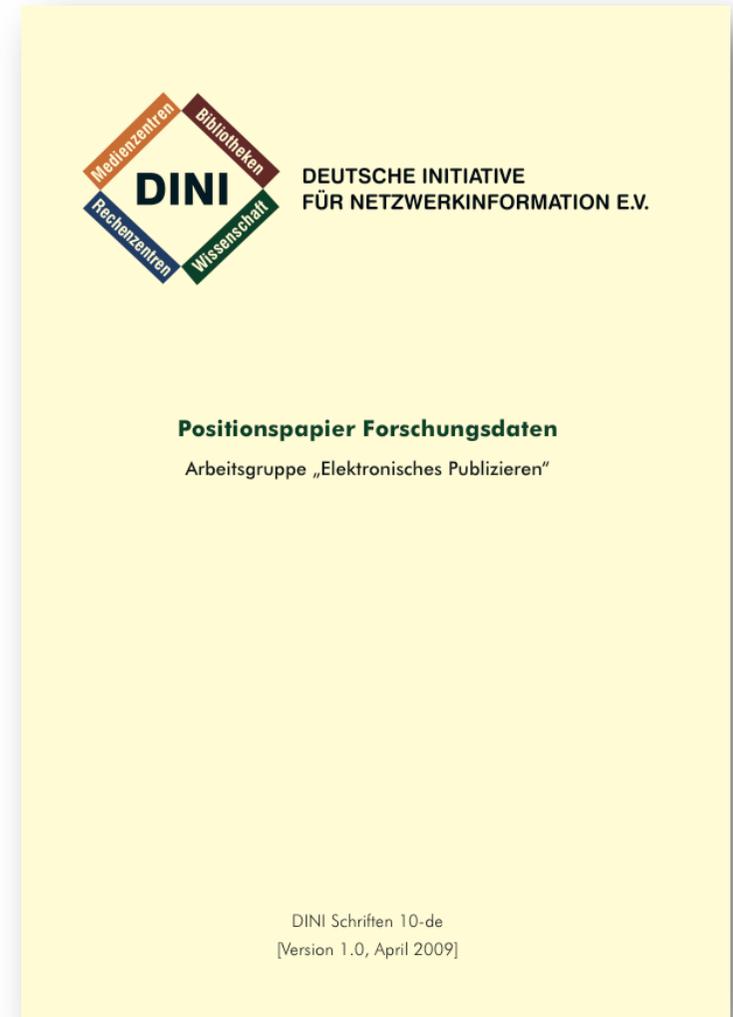
- Organisation
 - Definition eines Workflows
 - Akteure und Rollen
 - Infrastruktureinrichtungen
 - Unterstützung beim Umgang mit Forschungsdaten
 - Aufbau lokaler Dienste
 - Verankerung im lokalen und/oder interdisziplinären Informationsmanagement
 - Kooperation anhand disziplinärer Anforderungen
 - Qualifizierungsangebote (Aus- und Weiterbildung)



[13]

POSITIONSPAPIER FORSCHUNGSDATEN

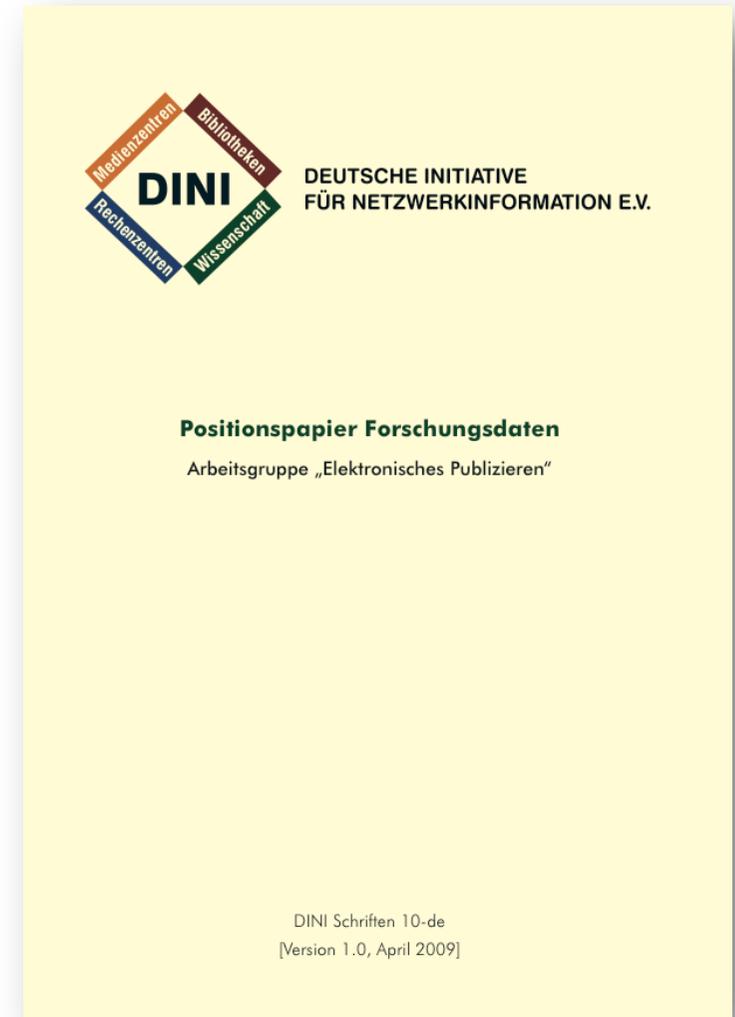
- Zugang
 - Offener Zugang fördert Transparenz und Forschungseffektivität
 - Offener Zugang ist für alle Disziplinen erstrebenswert
 - Jedoch: Disziplinabhängige Aspekte sind zu beachten
 - Offene Standards
 - Einbindung in disziplinspezifische Forschungsinfrastrukturen



[13]

POSITIONSPAPIER FORSCHUNGSDATEN

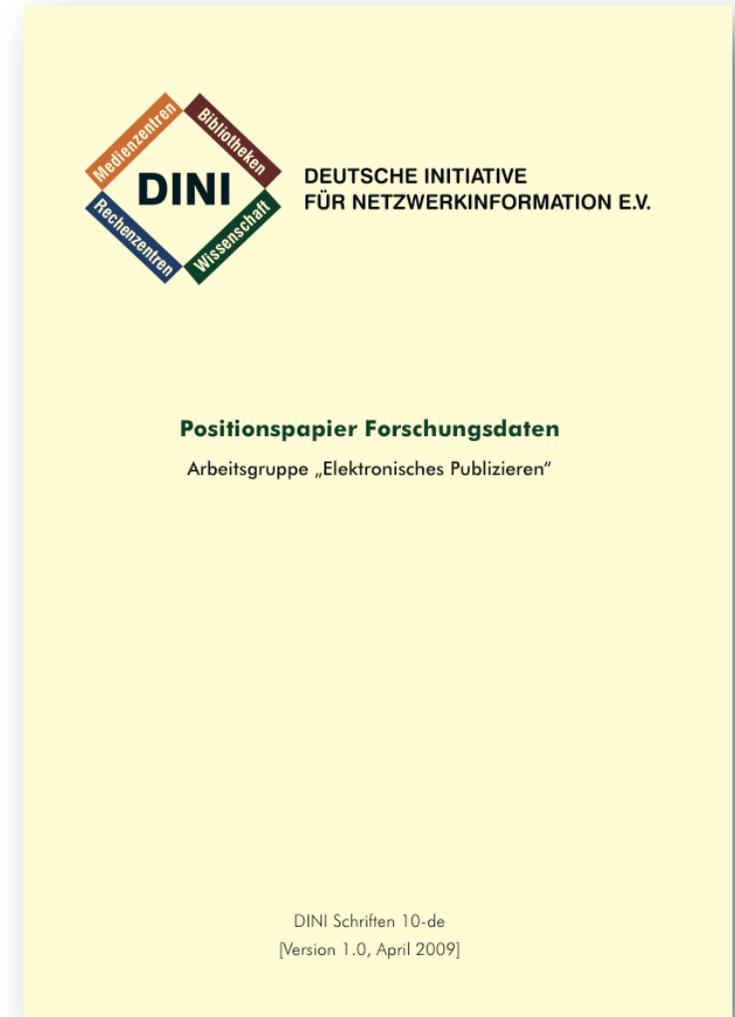
- Qualität
 - Nur qualitätsgesicherte Forschungsdaten können sinnvoll nachgenutzt werden
 - Methoden der Qualitätssicherung variieren
 - Fachwissenschaft
 - Inhaltliche Begutachtung
 - Infrastrukturelle Serviceinstitutionen
 - Ggf. formelle Qualitätssicherung
 - Konzepte zur Zitierung von Forschungsdaten
 - Eindeutige Identifikation
 - Dauerhafte Adressierung



[13]

POSITIONSPAPIER FORSCHUNGSDATEN

- Technik
 - Vertrauenswürdige digitale Infrastrukturen
 - Konzepte der digitalen Langzeitarchivierung
 - Aspekte der Informationssicherheit
 - Interoperable Schnittstellen
 - Disziplinspezifische Standards
 - Nutzerfreundliche Services



[13]

POSITIONSPAPIER FORSCHUNGSDATEN

- Beispiele für disziplinäre Aktivitäten
 - CrystalEye
 - Earth System Science Data (ESSD)
 - GenBank
 - PubChem
 - Publishing Network for Geoscientific & Environmental Data (PANGAEA)
 - World Atlas of Language Structures (WALS)

- Aktuelle Projekte und Aktivitäten in Auswahl
 - DOI-Registrierungsagentur der Technische Informationsbibliothek (TIB)
 - DRIVER
 - Helmholtz Open Access Projekt
 - Kooperative Langzeitarchivierung für Wissenschaftsstandorte (KoLaWiss)
 - PARSE.Insight
 - STD-DOI – Publication and Citation of Scientific Primary Data

FAZIT

- Hoher technischer, organisatorischer und finanzieller Aufwand
- Differierender Umgang mit Forschungsdaten
- Umgang mit Forschungsdaten ist verbesserungswürdig
- Mehrheitlich werden die heutigen technischen Möglichkeiten nicht genutzt
- OECD: „to improve the efficiency and effectiveness of the global science system“
- Nachnutzung und Nachprüfbarkeit durch dauerhafte und offene Zugänglichkeit
- Herausforderung für das System Wissenschaft
- Disziplinspezifische Aspekte

FAZIT

- Wissenschaftspolitik
 - Umsetzbare wissenschaftsnahe Richtlinien
 - Finanzierung
 - ...
- Disziplinen
 - Tiefgreifende Differenzen
 - Kultur des „data sharing“
 - Anreize
 - Qualifikation
 - ...
- Infrastruktureinrichtungen
 - Rollen, Aufgaben und Qualifikationen
 - Bewusstseinsbildung
 - Nutzerfreundliche Werkzeuge und Dienstleistungen
 -

DANKE FÜR DIE AUFMERKSAMKEIT!

■ Autorinnen und Autoren:

- Sünje Dallmeier-Tiessen
- Susanne Dobratz
- Prof. Dr. Stefan Gradmann
- Dr. Wolfram Horstmann
- Eike Kleiner
- Prof. Dr. Peter Schirmbacher
- Dr. Birgit Schmidt
- Frank Scholze
- Dr. Matthias Schulze

■ <http://www.dini.de>



DANKE FÜR DIE AUFMERKSAMKEIT!

Heinz Pampel

pampel@gfz-potsdam.de

<http://oa.helmholtz.de>



QUELLEN

- [1] Nature: Special Data Sharing. In: Nature 461 (2009), Nr. 7261, URL: <http://www.nature.com/news/specials/datasharing>
- [2] Savage CJ, Vickers AJ: Empirical Study of Data Sharing by Authors Publishing. In PLoS ONE 4 (9): e7078, DOI:10.1371/journal.pone.0007078
- [3] DFG: Vorschläge zur Sicherung guter wissenschaftlicher Praxis, 1998, URL: http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf
- [4] Berlin Declaration: Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, 2003, URL: <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>
- [5] OECD: Principles and Guidelines for Access to Research Data from Public Funding, 2007, URL: <http://www.oecd.org/dataoecd/9/61/38500813.pdf>
- [6] Allianz der deutschen Wissenschaftsorganisationen: Schwerpunktinitiative „Digitale Information“, 2008, URL: <http://allianz-initiative.de>
- [7] EUROHORCs / ESF: EUROHORCs and ESF Vision on a Globally Competitive ERA and their Road Map for actions, 2009, URL: http://www.dfg.de/internationales/download/era_road_map_08.pdf
- [8] National Academy of Sciences: Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age, 2009, URL: http://www.nap.edu/html/12615/12615_EXS.pdf
- [9] Rice, R : Data Sharing Continuum, 2007, URL: http://www.disc-uk.org/docs/data_sharing_continuum.pdf
- [10] RIN: Stewardship of digital research data, 2008, URL: <http://www.rin.ac.uk/system/files/attachments/sarah/Stewardship-data-guidelines.pdf>
- [11] ARL: E-Science Talking Points for ARL Deans and Directors, 2008, URL: <http://www.arl.org/bm~doc/e-science-talking-points.pdf>
- [12] STM: Brussels Declaration on STM Publishing, 2007, URL: http://www.stm-assoc.org/2007_02_13_Brussels_Declaration.pdf
- [13] DINI: Positionspapier Forschungsdaten, 2009, URN: <http://nbn-resolving.de/urn:nbn:de:kobv:11-10098082>