

Forschungsdaten-Repositorien

Informationsinfrastrukturen für nachnutzbare Forschungsdaten

Heinz Pampel, Deutsches GeoForschungsZentrum GFZ

12. Oracle Bibliotheken Summit, Bonn, 21.06.2012

Agenda

- Hintergrund
- Forschungsdaten-Repositorien
- re3data.org
- Perspektive

Hintergrund

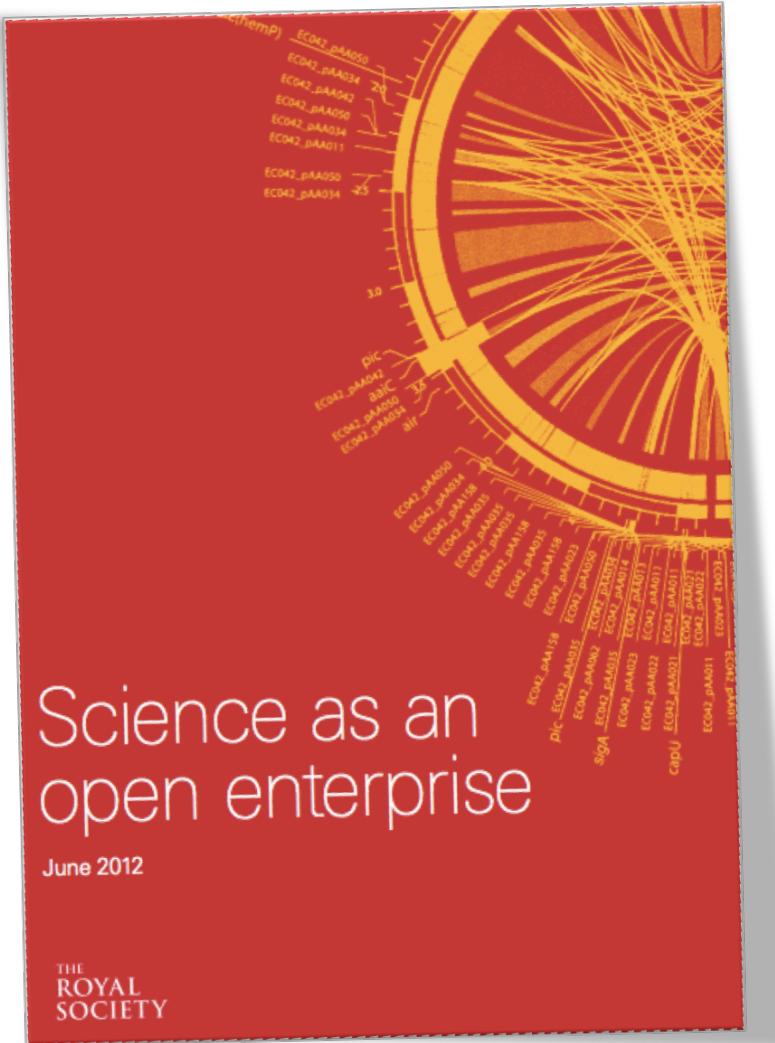
- Forschungsdaten sind wertvoll und einzigartig
- Neue Technologien fördern datenintensive Forschung
- Breite Diskussion über die dauerhafte Zugänglichkeit und Nachnutzung der Daten
- Data Policies als Motor der Diskussion
- Je nach Disziplin diverse Herausforderungen

Enquete-Kommission Internet und digitale Gesellschaft

699 Die Enquete-Kommission empfiehlt dem Bund, Projekte voranzutreiben, die verbindliche
700 Standards für Zugänglichmachung und Erhalt der digitalen Datenbestände aus öffentlicher
701 Forschung entwickeln. Sie empfiehlt weiter die Unterstützung der
702 Wissenschaftsorganisationen bei der Normierung und Standardisierung von (Meta-)Daten,
703 Quellenbezeichnungen (u.a. persistent identifiers) und anderen Formaten. Die Vernetzung
704 von Daten und Repositorien setzt eine eindeutige Referenzierung sowie die Interoperabilität
705 von Daten voraus. Die Interoperabilität muss – ohne Brüche – die Integration von
706 Dienstleistungen in den Arbeitsablauf von Wissenschaftlerinnen und Wissenschaftlern
707 gewährleisten. Zur Implementierung von Standards sowie der Zugänglichmachung gehören
708 auch Fragen der Qualitätssicherung, der Barrierefreiheit und der technischen
709 Voraussetzungen. Entsprechende Akteure sollten darüber hinaus bei der Entwicklung von
710 Standards und Werkzeugen für die Datenanalyse, Datenanreicherung und Datenpublikation
711 unterstützt werden. Die Auswahl der erhaltenswerten Datenbestände soll den
712 Fachdisziplinen obliegen.

Enquete-Kommission Internet und digitale Gesellschaft, 2012.

The Royal Society



„Scientists should communicate the data they collect and the models they create, to allow free and open access, and in ways that are intelligible, assessable and usable for other specialists in the same or linked fields wherever they are in the world. Where data justify it, scientists should make them available in an appropriate data repository. Where possible, communication with a wider public audience should be made a priority, and particularly so in areas where openness is in the public interest.“

Data Policies: Förderer

- NIH, 2003
 - „The NIH endorses the sharing of final research data to serve these and other important scientific goals. The NIH expects and supports the timely release and sharing of final research data from NIH-supported studies for use by other researchers. Starting with the October 1, 2003 receipt date, investigators submitting an NIH application seeking \$500,000 or more in direct costs in any single year are expected to include a plan for data sharing or state why data sharing is not possible.“
- DFG, 2012
 - „Wenn aus Projektmitteln systematisch (Mess-)Daten erhoben werden, die für die Nachnutzung geeignet sind, legen Sie bitte dar, welche Maßnahmen ergriffen wurden bzw. während der Laufzeit des Projektes getroffen werden, um die Daten nachhaltig zu sichern und ggf. für eine erneute Nutzung bereit zu stellen. Bitte berücksichtigen Sie dabei auch - sofern vorhanden - die in Ihrer Fachdisziplin existierenden Standards und die Angebote bestehender Datenrepositorien.“

HORIZON 2020

Document	Reference	Proposal
Proposal for a Regulation of the European Parliament and of the Council establishing Horizon 2020 - The Framework Programme for Research and Innovation (2014-2020)	COM(2011) 809 final 2011/0401 (COD)	„Free, open and unrestricted access to interoperable data and information will be encouraged.“
Proposal for a Regulation of the European Parliament and of the Council laying down the rules for the participation and dissemination in 'Horizon 2020 – the Framework Programme for Research and Innovation (2014-2020)	COM(2011) 810 final 2011/0399 (COD)	„With regard to dissemination through research publications, open access shall apply under the terms and conditions laid down in the grant agreement. With regard to dissemination of other results, including research data, the grant agreement may lay down the terms and conditions under which open access to such results must be provided, in particular in ERC frontier research or any other appropriate areas.“

European Commission, n.d.

Data Policies: Journals

- **Nature, 2011**
 - „Data sets must be made freely available to readers from the date of publication, and must be provided to editors and peer-reviewers at submission, for the purposes of evaluating the manuscript. For the following types of data set, submission to a community-endorsed, public repository is mandatory. Accession numbers must be provided in the paper. Examples of appropriate public repositories are listed below.“
- **AGU, 1996**
 - „Data sets cited in AGU publications must meet the same type of standards for public access and long-term availability as are applied to citations to the scientific literature. Thus data cited in AGU publications must be permanently archived in a data center or centers that meet the following conditions: a) are open to scientists throughout the world. b) are committed to archiving data sets indefinitely. c) provide services at reasonable costs.“

Opportunities for Data Exchange (ODE)

education accreditation & certification

legislation culture & attitude quality

cooperation financing policies

data sharing

publishing & visibility data flow improvements

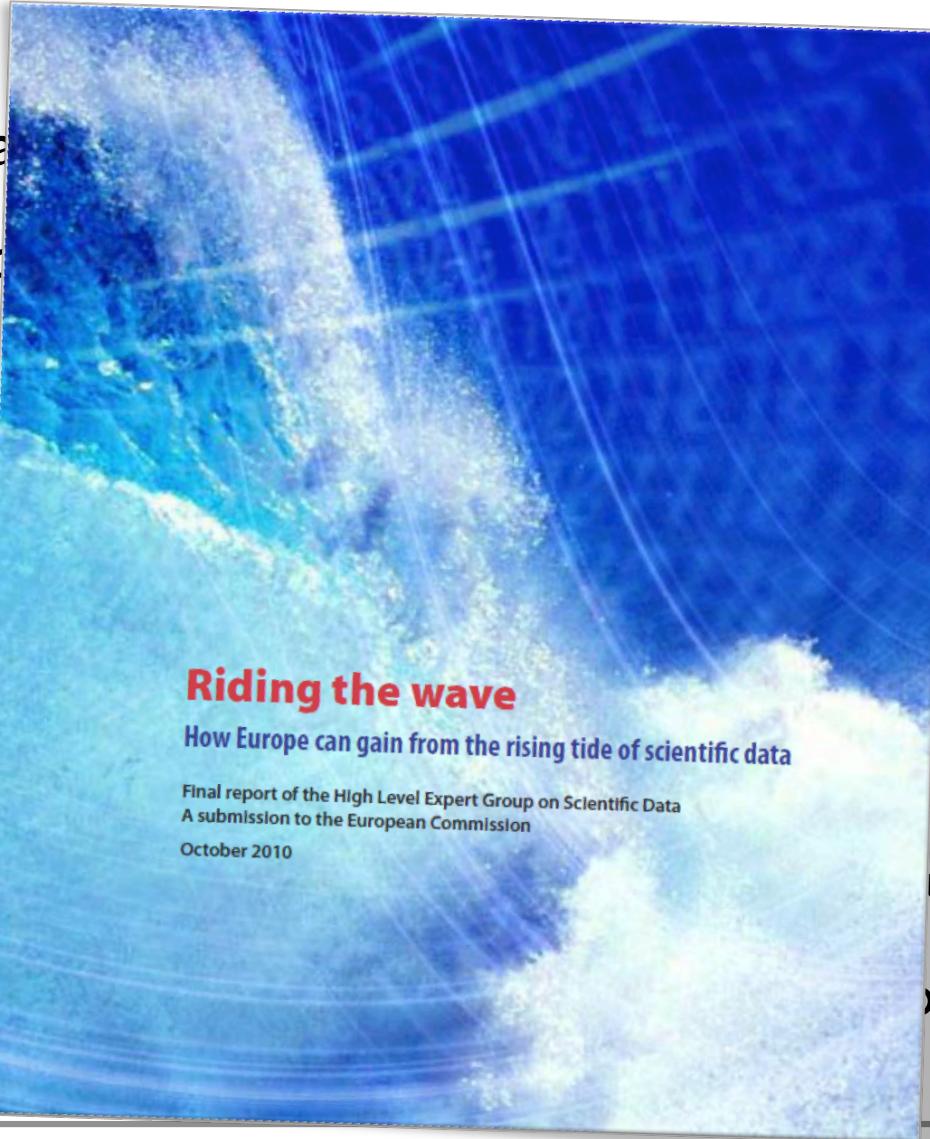
infrastructure disciplines career

efficiency

Schäfer et al., 2011.

Vision 2030

- „Researchers are able to find, access and evaluate the data they need with confidence in their quality, and can evaluate the quality of the data they produce.“
- “Producers of scientific data prefer to deposit their data in international scientific data repositories. This is the only way to make sure that the data is available and accessible to the international research community.”



are able to
can be
a, and they
trusted.“
d access, and
able
d by
orthy.”

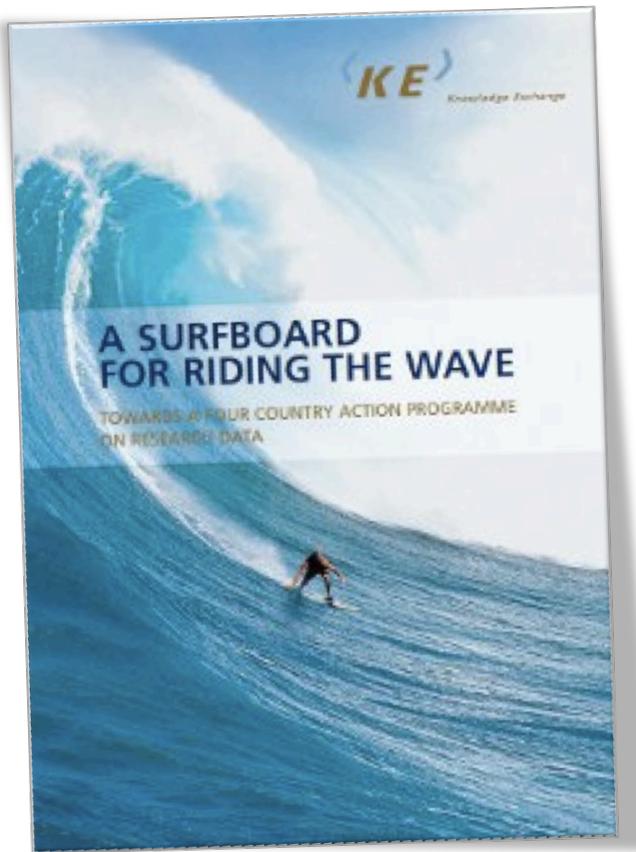
High Level Expert Group on Sci

Vision 2030

- „Researchers and practitioners from any discipline are able to find, access and process the data they need. They can be confident in their ability to use and understand data, and they can evaluate the degree to which that data can be trusted.“
- “Producers of data benefit from opening it to broad access, and prefer to deposit their data with confidence in reliable repositories. A framework of repositories is guided by international standards, to ensure they are trustworthy.”

High Level Expert Group on Scientific Data, 2010.

Handlungsfelder zur Realisierung der Vision 2030



- Anreize für Wissenschaftler
- Ausbildung und Training der beteiligten Akteure
- Aufbau und Vernetzung der Infrastrukturen
 - „The three main challenges in developing an ecosystem of data repositories are (1) gaps in the present data infrastructure and (2) connectivity issues (between the workflow of researchers and the institutional data infrastructure and between institutional and national data infrastructures) and (3) long-term financial basis.”

Finanzierung der Aktivitäten

Van der Graaf, M., & Waaijers, L. 2011.

Publikationsstrategien

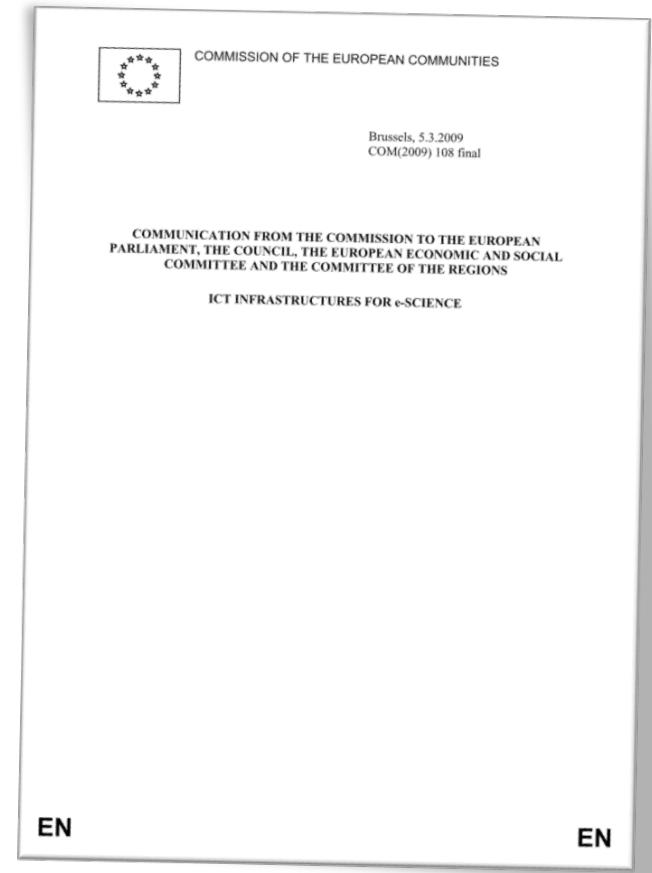
- Die Veröffentlichung von Forschungsdaten als eigenständiges Informationsobjekt in einem Repository (oder einem anderen „Ort“)
- Die Veröffentlichung von Forschungsdaten als Anreicherung einer (interpretativen) Text-Publikation („Enriched Publication“)
- Die Veröffentlichung von Forschungsdaten mit textueller Dokumentation, als sogenanntes „Data Paper“

Angelehnt an Dallmeier-Tiessen, S., 2011.

Forschungsdaten-Repositorien

- EC: ICT infrastructures for e-science, 2009

„The landscape of data repositories across Europe is fairly heterogeneous, but there is a solid basis to develop a coherent strategy to overcome the fragmentation and enable research communities to better manage, use, share and preserve data.“



European Commission, 2009.

Forschungsdaten-Repositorien

The screenshot shows the NCBI GEO homepage. At the top left is the NCBI logo. To its right is the GEO logo with the text "Gene Expression Omnibus". Below the logo are navigation links: "GEO Publications", "FAQ", "MIAME", "Email GEO", and "Login". A breadcrumb trail "NCBI » GEO" is at the top left. A main text block states: "Gene Expression Omnibus: a public functional genomics data repository supporting submissions. Array- and sequence-based data are accepted. Tools are provided to analyze experiments and curated gene expression profiles. [More information »](#)". On the left, there's a "GEO navigation" sidebar with two main sections: "QUERY" and "BROWSE". The "QUERY" section contains four buttons: "DataSets", "Gene profiles", "GEO accession", and "GEO BLAST", each with a search input field and a "GO" button. The "BROWSE" section contains three buttons: "DataSets", "GEO accessions", and "Platforms", "Samples", "Series".

GEO, <http://www.ncbi.nlm.nih.gov/geo/>

The screenshot shows the PANGAEA homepage. At the top right is the PANGAEA logo with the tagline "Data Publisher for Earth & Environmental Science". Below the logo is a globe icon with green and red continents and the word "PANGAEA" in the center. A horizontal menu bar includes "All", "Water", "Sediment", "Ice", "Atmosphere", "Help", "Advanced Search", "Preferences", and "more...". At the bottom, there are links for "About", "Submit Data", "Projects", "Software", "WDC-MARE", and "Contact". A note at the bottom states: "This work is licensed under a Creative Commons License".

PANGAEA, <http://www.pangaea.de>

Forschungsdaten-Repositorien

Scientific Drilling Database
Data from Deep Earth Sampling and Monitoring

+ Home
+ About SDDB
+ News
+ Data Publications
+ Catalogue
+ Authors
+ Dataset
+ Research Programs
+ Sampling Gear
+ Analytical Parameters
+ Publication Process
+ Admin

GFZ icdp

Dataset Description

Citation: Demory, Francois; Nowaczyk, Norbert; Witt, An
Intercorrelation of down-core variations of the high-resolution magnetic susceptibility for COND1-603 and COND1-605. *Scientific Drilling* /GFZ SDDB.1003

[Download Citation \(EndNote\)](#)

Title: Intercorrelation of down-core variations of the high-resolution magnetic susceptibility for COND1-603 and COND1-605

Abstract: Down-core variations of the high-resolution magnetic susceptibility have been obtained for kasten and pilot cores from site COND1-605. The transfer of AMS 14C dating performed on kasten cores to the pilot cores subjected to significant discrepancies in the magnetic susceptibility curve COND1-605 (Vydrovo Shoulder) measured in 2001 and large peaks visible in the first measurement after a 2-year-long storage. This is a first hint for chemically unstable greigite.

[Show in Google Earth](#)

Related Publications:

- N. Fagel, L.Y. Alleman, L. Granina, F. Haas, J. André, Vivianite formation and distribution in the Paleogene of the Western Carpathians, *Planet. Change* 46 (2005), pp. 315-336. doi:10.1016/j.palchange.2005.03.003

SDDB, <http://www.scientificdrilling.org>

The Bern Digital Pantheon Project

Home The Building Research Publications The Team Repository Contact Search

You are here: Home > Repository > BDPP0019

Repository

first previous Record 198 of 776 next last

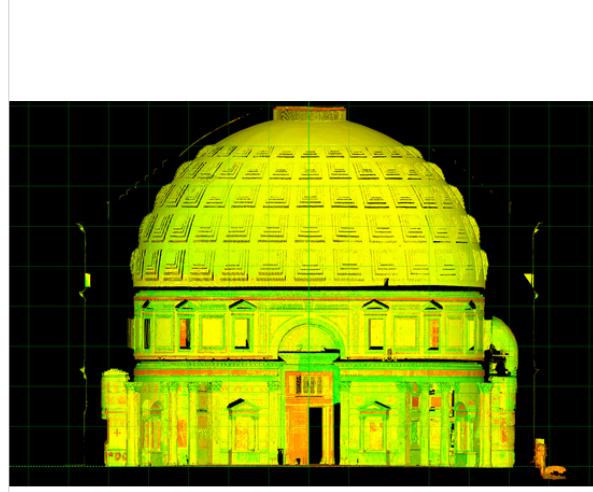
The Bern Digital Pantheon Project: Transverse cross-section, facing north

The Bern Digital Pantheon Project: Transverse cross-section, facing north. In: Gerd Graßhoff, Markus Wäfler, Jon Albers and Christian Berndt (ed.): Digital Repository of the Bern Digital Pantheon Project. Bern 2009, BDPP0019 (Published online at: <http://www.digitalpantheon.ch/BDPP0019>).

BibTeX:

```
@other(BDPP0019,
author = {The Bern Digital Pantheon Project},
title = {Transverse cross-section, facing north},
year = {2009},
keywords = {Medium Visualisation Building Interior},
owner = {berndt},
timestamp = {2009.05.28},
editor = {Graßhoff, Gerd and Wäfler, Markus and Albers, Jon and Berndt, Christian},
address = {Bern},
repository = {Digital Repository of the Bern Digital Pantheon Project}
}
```

Usage of the repository



BDPP, <http://www.digitalpantheon.ch>

Forschungsdaten-Repositorien

PURDUE UNIVERSITY Research Repository | PURR

Home Browse Content Projects Get Started Contact Us Search

You are here: Publications > Datasets > Phosphorus and Potassium Influence on Alfalfa ... > About

Phosphorus and Potassium Influence on Alfalfa Nutrition

By Jeffrey J Volenec

Data from several studies on the influence of phosphorus and potassium nutrition on alfalfa. It includes numeric data such as yield, plant mass, plant counts, and tissue concentration of various nutrients

Listed in Datasets

About Supporting Docs Versions Reviews Questions Wish List

Abstract This is a master file of merged excel files with data from several studies on the influence of phosphorus and potassium nutrition on alfalfa. It includes numeric data such as yield, plant mass, plant counts, and tissue concentration of various nutrients. Conventional wisdom states that plants that are poorly fertilized and do not have adequate phosphorus and especially low potassium will become winterkill. This study showed that plants did not necessarily die in winter with low potassium stress.

The University of Edinburgh

University Homepage IS Homepage Research Support Kontakt Einloggen iS

Information Services

DSpace Startseite > Information Services (IS) > Vice Principal's Office > Dokumentanzeige

A survey of undergraduate technology use and attitudes

Download Dataset Zur Langanzeige

Titel: A survey of undergraduate technology use and attitudes
Autor: Macleod, Hamish; Paterson, Jessie
Datum: 2012-05-09
Citation: Macleod, Hamish; Paterson, Jessie. (2012). A survey of undergraduate technology use and attitudes, 2011 [Dataset]. University of Edinburgh.
Permanent Identifier: <http://hdl.handle.net/10283/144>

An award from ELESIG (The Evaluation of Learners' Experience of eLearning Special Interest Group) provided an opportunity to continue the long and successful collection of information about how our students view ICT at the University of Edinburgh. The original work started in the 1990s but lapsed in recent years due to data collection issues. During the second semester in 2010-11, using a paper questionnaire, first and second years across the University were targeted. The findings raised a number of interesting points, for example some students still find IT challenging; although laptop ownership is high, students do not seem to be planning to carry their laptops to the campus; Facebook appears to be used significantly for academic purposes.

Dateien	Description	Größe	Format	Anzeige	Beschreibung
ICTSurvey.csv	Data file in CSV format	46.51Kb	Text	Download	Data file in CSV format
ICTSurvey.sav	Data file in SPSS format	29.83Kb	Unbekannt	Download	Data file in SPSS format
ICTSurvey-codebook.rtf	Codebook in Rich Text Format	1.563Mb	RTF	Download	Codebook in Rich Text Format
ICTSurvey-codebook.spv	Codebook in SPSS format	137.5Kb	Unbekannt	Download	Codebook in SPSS format

DataShare, <http://databank.is.ed.ac.uk/>

Forschungsdaten-Repositorien

The screenshot shows a Figshare research data page for a dataset titled "Herring Hepatic Transcriptome 34300 contigs.fa". The page includes a preview of the FASTA file content, download links, sharing options, citation information, and metadata like categories, authors, and tags.

Dataset Details:

- Title:** Herring Hepatic Transcriptome 34300 contigs.fa
- Views:** 3677
- Shares:** 7
- Cites:** coming soon
- Published:** on 19 Jan 2012 - 16:37 (GMT)
- Filesize:** 22.40 MB

Categories:

- Molecular biology
- Marine Biology

Authors:

Steven Roberts

Tags:

- contigs
- herring
- fasta
- ngs

Export:

- Export to Ref. Manager

Share this: [Share](#) 0 [Tweet](#) 6 [G+](#) 1

Cite this: Herring Hepatic Transcriptome 34300 contigs.fa. Steven Roberts. Figshare. Retrieved 10:28, May 11, 2012 (GMT) hdl.handle.net/10779/084d34370fbda29bbc67b3c5ecb02575

Preview: **Download:**

```
>ConsensusfromContig1 Average coverage: 2.00
GTGTGTGTGTGTGTCTCTACGCCCTCGTCTTCACTGTACTCGCTGCAGTGCT
CGACGGGGGTGCTCAGGCATCTGGGGTGTGCCCTCTGGGGTTCAGCCCTGGACGAGC
TTGGGGCTGCTGATTGGATGGCTGTGGCCTGGGCTCTTCCTCATTGGTCTGGCCC
TCAGGGCTCCGGCCCTTCCCTGGGCTTGGGGAGGCTCAGTCTCTGCAGCTGGCTTGCC
TCTGGCTCAGGCTCGGGCTCCGGCTACTGTTGCCCTCCGCACCTGACGAGGAGGAGGG
AGCAACTGCTGCTGCTGCTTGTGCTGCTGCATCATGTCATGTCATGTCATGTTG
TACACCAGGTTGGCCATTCCGGCATGTTGATGCCCTGGGCTGGGACGAGCGA
TCTCCAATAGGCTGGGTAGTACTCATTAAGTCGATATGTTGACTGCAATAGGGCGGCC
GTGAGC
>ConsensusfromContig2 Average coverage: 1.64
TCGAGGAGCTGACCGGATGGCTCTCGAGAACGAGGACTGGTCGAGACCCCTGCGCTG
CAGCTCGAGATCGGGCAGCAAGCAGGCGCGGGCACGATGACGACGACGGGGGGGCC
AAGGAGGACGGCTACCCGACGGCCAGAGGGCGCTGTTGTGCGGGTGAAGGAGGAGGCC
CCCGAGCTGTAACCCGGACGGCTACCCGCTGCTCTCCCCGGCGTGAGGCCATC
CGCTCACCGTGAAGAGGAAGGGACGGAGGAGGAGGAGGAGGAGGAGGAGAG
GAGATGGACGACGTGATGGCGGAGACGAAACGACTGCGATGTCG
>ConsensusfromContig3 Average coverage: 3.34
CTGGCTATCAGGTGTAATTTCCAGGCACTGTTCTGGCTTGGCTCAGTTGTC
TTTAGTCTGTTCCAGTTCAGCCTCATTTGAGAAAATAAGTTGATGGCAGGGTGCAC
```

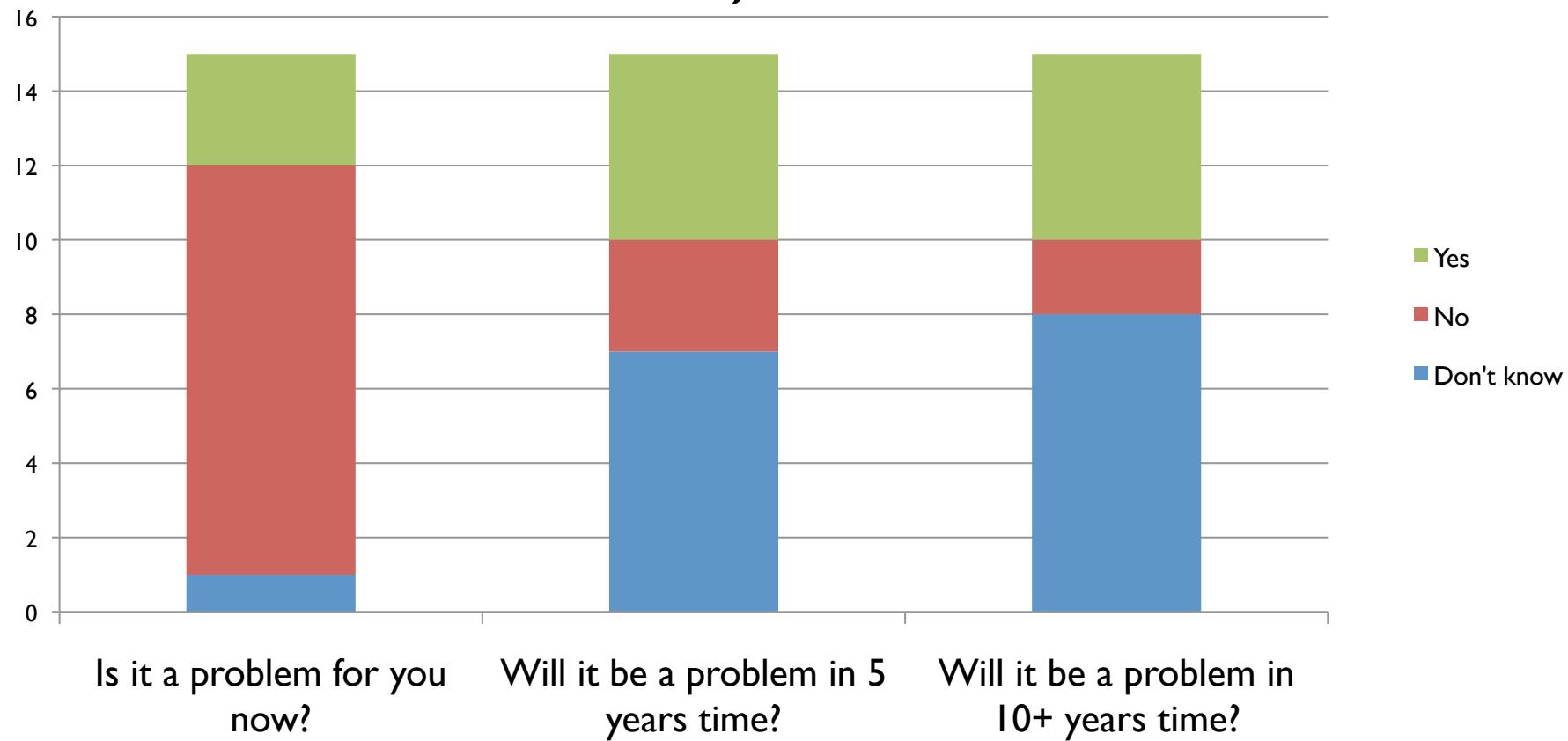
Figshare, <http://figshare.com>

Beispiel: Biomedizin

- 200 Repositorien
- 100 Institutionen
- 350 Personen
- > 10.000 Nutzerinnen und Nutzer
- 60 Millionen Zugriffe per Monat
- 35 Millionen EUR jährliche Kosten

ELIXIR, n.d.

Funding of the repositories - now and in the future, n = 15



re3data.org: Mission

- Aufbau eines “Registry of Research Data Repositories”
- Ziel ist ein weltweites, disziplinübergreifendes, web-basiertes Verzeichnis von Forschungsdaten-Repositorien
- Orientierung für Wissenschaftler, Förderer, Verlage, Infrastruktur
- Vernetzung und Standardisierung von Daten-Repositorien
- Förderung der Sichtbarkeit von Daten
- Schaffung einer “cultur of sharing”

re3data.org: Partner

- Deutsches GeoForschungsZentrum GFZ, Bibliothek
- Humboldt-Universität zu Berlin, IBI
- Karlsruher Institut für Technologie (KIT), KIT Library



- Förderung durch die DFG (2012-2014)



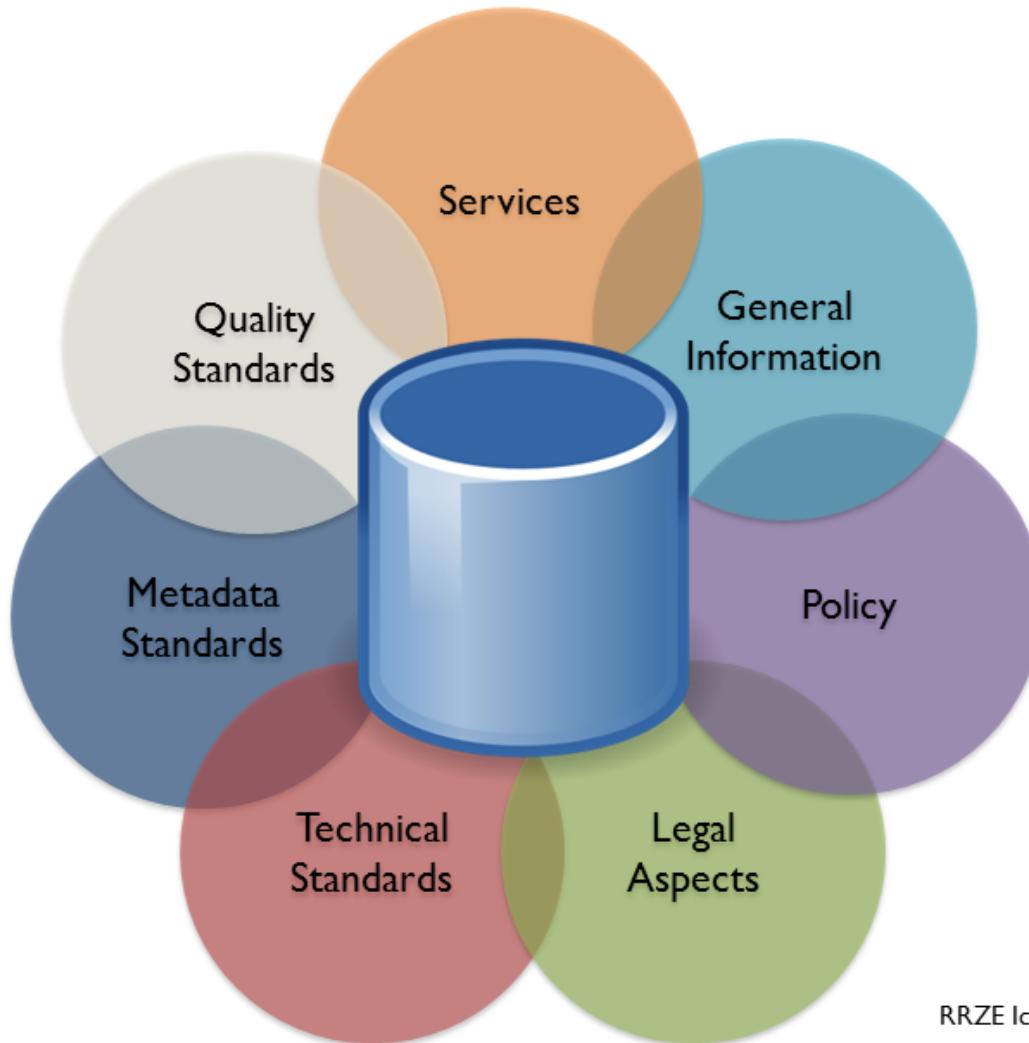
- Vernetzung im Rahmen von DINI
- MoU mit DataCite



re3data.org: Status quo

- Aktivitäten in der ersten Projektphase:
 - Konzeption des Registries
 - Entwicklung eines Metadaten-Schemas
 - Formulierung von Auswahlkriterien
 - Technische Entwicklung des Registries
 - Vernetzung mit relevanten Initiativen
- Onlinegang für Herbst 2012 geplant

re3data.org: Aspekte



RRZE Icon Set (CC: BY-SA)

Perspektive

- Weiterentwicklung der Repositorien-Landschaft
- Forschung und Entwicklung nötig („Erfolgsfaktoren“)
- Anforderung: „Open Infrastructures for Open Science“
- Zentrale Herausforderung:
 - Standardisierung
 - Vernetzung
 - Finanzierungsmodelle
- Schaffung einer “cultur of sharing”

Danke für die Aufmerksamkeit!

- GFZ:
 - Roland Bertelmann
 - Jens Klump
 - Heinz Pampel
- HU:
 - Peter Schirmbacher
 - Maxi Kindling
 - Paul Vierkant
- KIT:
 - Frank Scholze
 - Hans-Jürgen Goebelbecker
 - Jens Gundlach
 - Stefka Manova

The screenshot shows the homepage of re3data.org. At the top, the logo 're3data.org' is displayed with the subtitle 'REGISTRY OF RESEARCH DATA REPOSITORIES'. Below the logo is a navigation bar with links to Home, About, Partners, Contact, and Imprint. A search bar is located on the right side of the navigation bar. The main content area features a title 'Memorandum of Understanding between DataCite and re3data.org' with a speech bubble icon. Below the title, it says 'Posted on April 13, 2012 by re3data.org team'. The text explains the partnership between DataCite and re3data.org to enhance accessibility and visibility of data sets. To the right of the text, there are logos for GFZ (Helmholtz Centre Potsdam), Humboldt University Berlin, and KIT (Karlsruhe Institute of Technology).

info@re3data.org

<http://re3data.org>



With the exception of all photos and graphics, this slides are licensed under the
“Creative Commons Attribution 3.0 Germany (CC BY 3.0)“ Licence.



Literatur

- American Geophysical Union. (1996). Policy on Referencing Data in and Archiving Data for AGU Publications. Retrieved from http://www.agu.org/pubs/authors/policies/data_policy.shtml
- Dallmeier-Tiessen, S. (2011). Strategien bei der Veröffentlichung von Forschungsdaten. Berlin. Retrieved from http://www.ratswd.de/download/RatSWD_WP_2011/RatSWD_WP_173.pdf
- Deutsche Forschungsgemeinschaft. (2012). Leitfaden für die Antragstellung. DFG-Vordruck 54.01 - I/12. Retrieved from http://www.dfg.de/formulare/54_01/54_01_de.pdf
- ELIXIR. (n.d.). The ELIXIR Strategy for Data Resources. Draft Report from Workpackage 2.The ELIXIR Preparatory Phase. Retrieved from <http://www.elixir-europe.org/bcms/elixir/Documents/reports/ELIXIR%20Strategy%20for%20Data%20Resources%20report.pdf>
- Enquete-Kommission Internet und digitale Gesellschaft. (2012). Projektgruppe Bildung und Forschung. Handlungsempfehlungen. Ausschussdrucksache 17(24)052. Retrieved from http://www.bundestag.de/internetenquete/dokumentation/Sitzungen/20120625/A-Drs_17_24_052_-PG_Bildung_und_Forschung_Handlungsempfehlungen.pdf
- European Commission. (2009). ICT infrastructures for e-science. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2009) 108 final. Retrieved from <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0108:FIN:EN:PDF>
- European Commission. (n.d.). Main references to open access in the European Commission's proposals for Horizon 2020. Retrieved from http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-in-horizon-2020_en.pdf
- High Level Expert Group on Scientific Data. (2010). Riding the wave. How Europe can gain from the rising tide of scientific data. Retrieved from <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>
- National Institutes of Health. (2003, April 12). Final NIH Statement on Sharing Research Data. Retrieved from <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>
- Nature. (2011). Guide to Publication Policies of the Nature Journals. Retrieved from <http://www.nature.com/authors/gta.pdf>
- Pampel, H., & Bertelmann, R. (2011). „Data Policies“ im Spannungsfeld zwischen Empfehlung und Verpflichtung. In S. Büttner, H.-C. Hobohm, & L. Müller (Eds.), Handbuch Forschungsdatenmanagement (pp. 49-61). Bad Honnef: Bock + Herchen. Retrieved from <http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2287>
- Schäfer, A., Pampel, H., Pfeiffenberger, H., Dallmeier-Tiessen, S., Tissari, S., Darby, R., Giaretta, K., et al. (2011). Baseline Report on Drivers and Barriers in Data Sharing. Retrieved from http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2011/11/ODE-WP3-DEL-0002-I_0_public_final.pdf
- The Royal Society. (2012). Science as an open enterprise. The Royal Society Science Policy Centre report 02/12. Retrieved from http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf
- Van der Graaf, M., & Waaijers, L. (2011). A Surfboard for Riding the Wave. Towards a four country action programme on research data. Retrieved from http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=/Files/Filer/downloads/Primary+Research+Data/Surfboard+for+Riding+the+Wave/KE_Surfboard_Riding_the_Wave_Screen.pdf