HELMHOLTZ GEMEINSCHAFT OPEN ACCESS

Wer hat Angst vor Open Access?

hier: Zugang zu Forschungsdaten

Hans Pfeiffenberger

Alfred-Wegener-Institute for Polar and Marine Research,
Helmholtz Association - Germany

FFG-KOWI Workshop, 2013-06-04, Wien

AWI

HELMHOLTZ GEMEINSCHAFT OPEN ACCESS

Die kurze Antwort auf die Hauptfrage:

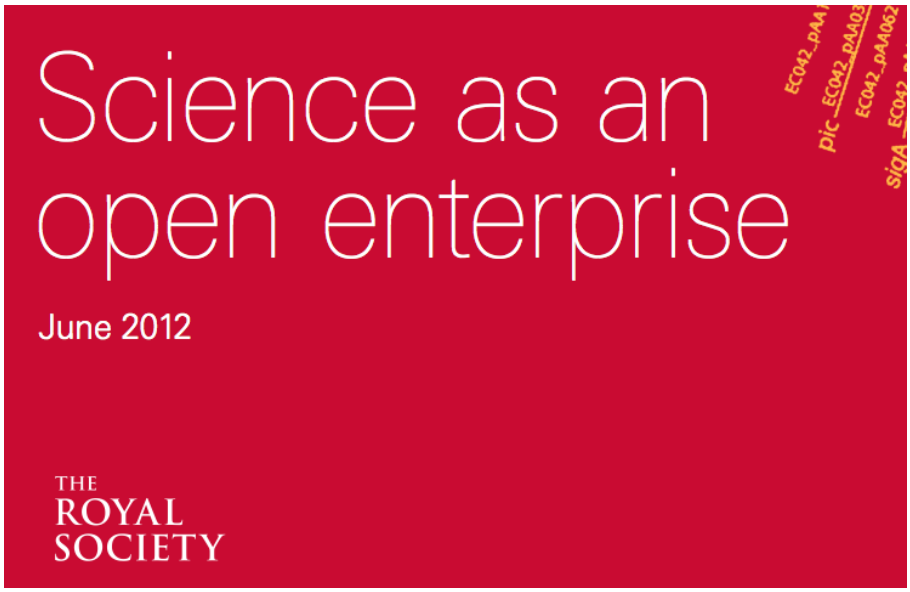
- **Diejenigen, welche ihre Daten**
  - **erfunden (Stapel) haben,**
  - **geschönt haben (gern bei Clinical Trials) oder**
  - **über- oder fehlinterpretiert haben (Reinhart/Rogoff)**

- **Diejenigen, welche ihr Geschäftmodell auf der Monopolisierung von Wissen oder Fakten aufbauen**
  - **Verlage mit Abo-Modell (“article of the future”)**
  - **Datenvermittler (Thomson Reuters, “Web of Knowledge”)**
  - **öffentlich finanzierte Institute (keine Namen, bitte!)**

Royal Society: Science as an Open Enterprise (2012)

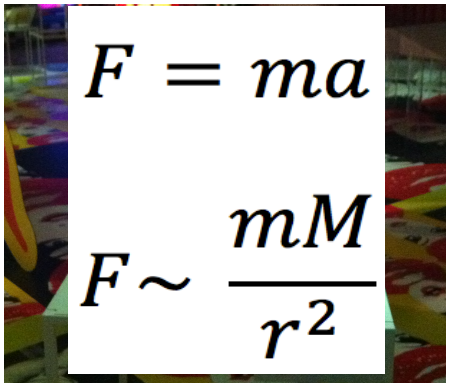
- **Offenheit gehört zum Kern der Wissenschaft.**

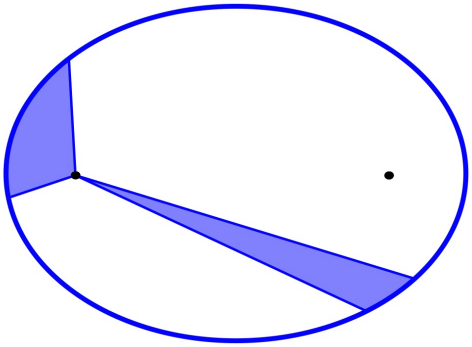
- (Open enquiry has been at the heart of science …)



Modern Science is based on data – since Renaissance!

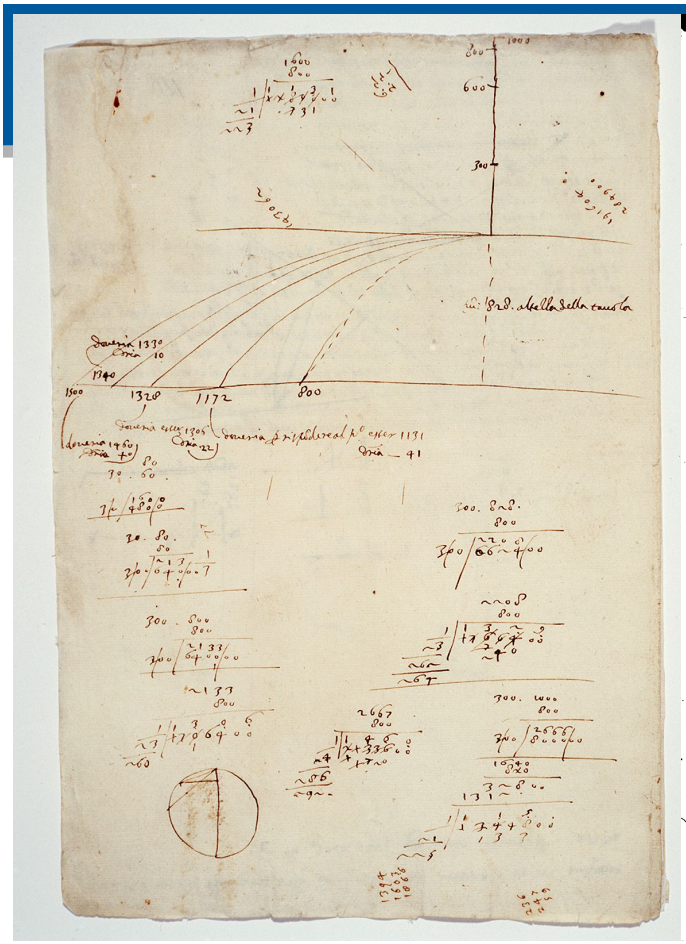
- **1546 - 1601: Tycho Brahe, *quality* data**
- **1606 - 1618: Kepler's Laws**
- **1684 - 1687: Newton De Motu – Principia**
  - **explained (!) Kepler's laws**





Planet	T	d	T^2	d^3	T^2/d^3
Merkur	0,241	0,387	0,058081	0,057960603	1,002077221
Venus	0,615	0,723	0,378225	0,377933067	1,000772446
Erde	1	1	1	1	1
Mars	1,881	1,524	3,538161	3,539605824	0,999591812
Jupiter	11,863	5,203	140,730769	140,8515004	0,999142846
Saturn	29,458	9,555	867,773764	872,3526289	0,994751131

T = siderische Umlaufzeit in trop. Jahren d = große Halbachse in astronomischen Einheiten (Abstand Erde–Sonne)



PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For Anno 1665, and 1666.

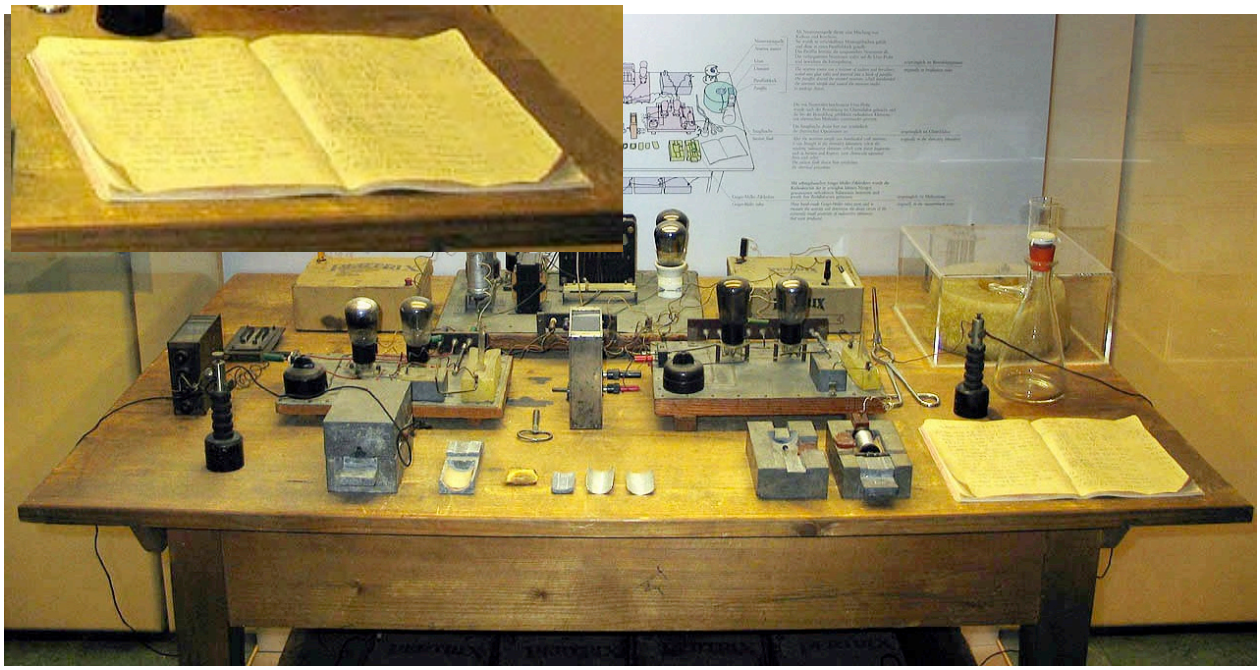
In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane,
Printers to the Royal Society.

Paradigms of modern (17th century) Science

- **Experiment / Observation**
  - **Repeatable, (documented), …**
  - **quantified**
- **Theory**
  - **mathematically expressed**
  - **checked against facts**

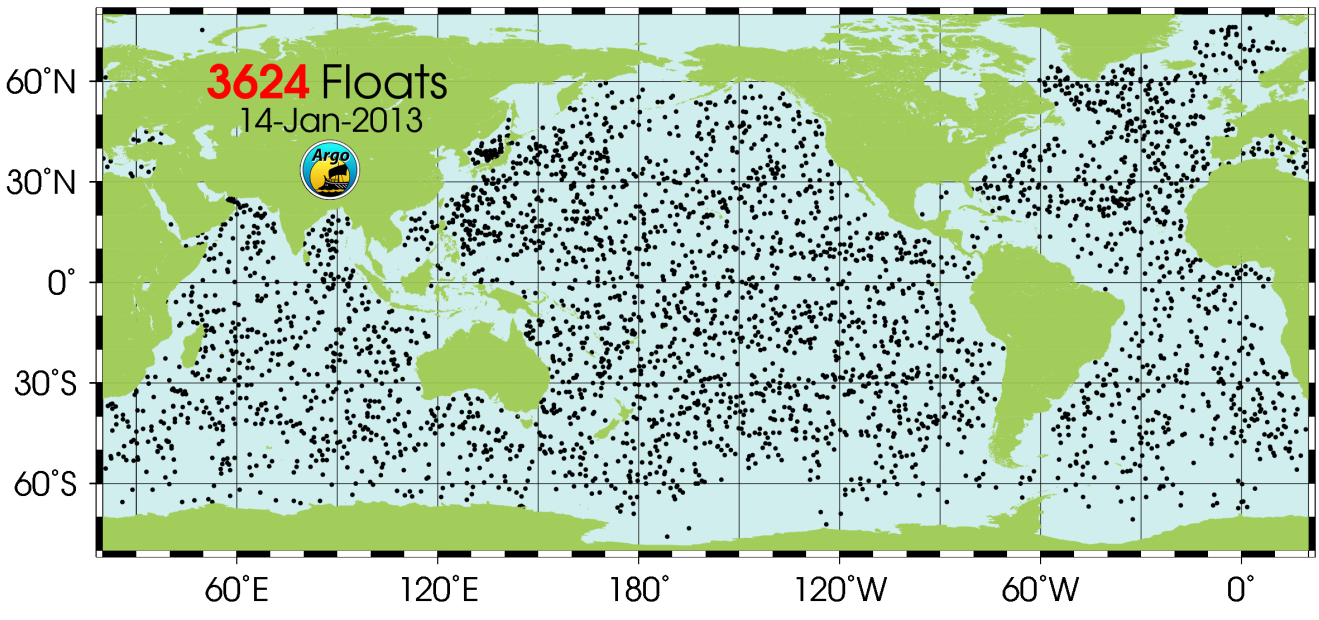
- **Published**
- **Governed by "Rules of good scientific conduct"**

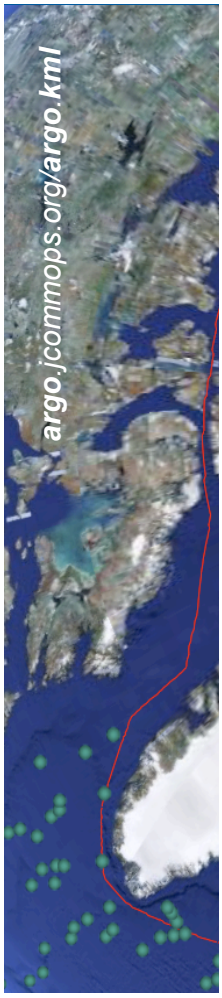
Meitner-Hahn-Strassmann Uran-Experiment, Berlin-Dahlem, 1938



The last big discovery by a „single“ person with a lab notebook ?

The biggest experiment in the world (not at CERN!)





6900499

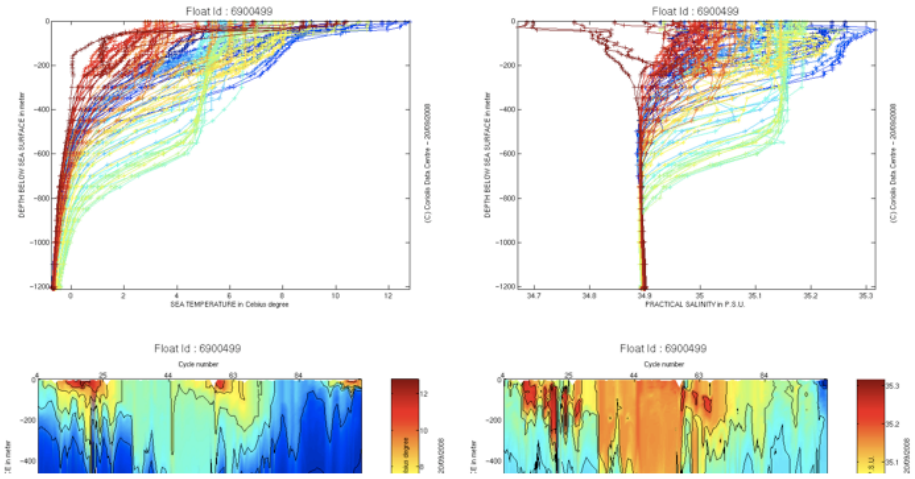
NORWAY (Argo NORWAY) 880 Days
95 profiles at GDACs (origin Coriolis) including 0 DM profiles

Deployment Date: 13/04/2006 Lat : 64.6500 Lon: -.0216
Latest Location Date: 09/09/2008 Lat: 67.0903 Lon: -9.0152
Web Products AIC Coriolis JMA CSIRO MEDS

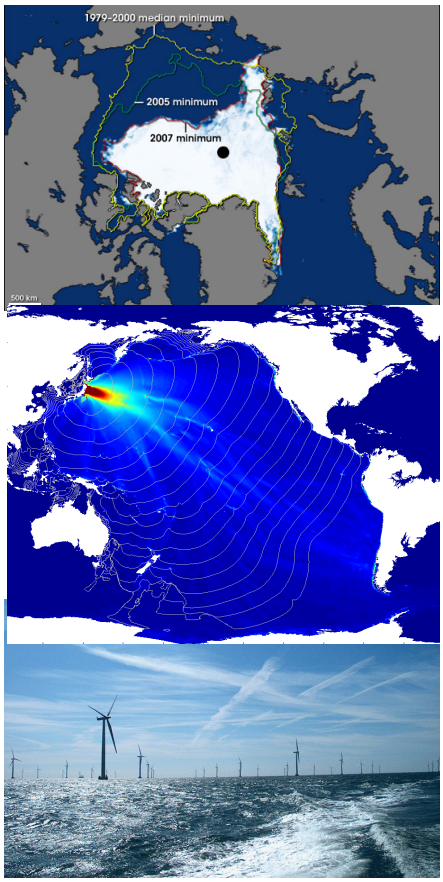
Data (netCDF) Profiles Metadata Trajectory Technical

QC Altimetry QC

Subsurface Temperature - Subsurface Salinity (source IFREMER/Coriolis)



MaNIDA – Enabling Data-Intensive Marine Science



Global Change

- Assessing, understanding, and predicting environmental changes
- Human environmental impact

Hazards

- Risk analysis and support for disaster management
- Understanding environmental factors affecting human health

Resources

- Sustainable ecosystem management
- Energy from the ocean

Royal Society: Science as an Open Enterprise (2012)

- Geoffrey Boulton, **Open your minds and share your results Nature** (486) 441, doi:10.1038/486441a
- Open enquiry has been at the heart of science since the first scientific journals were printed in the seventeenth century. …

  Science's capacity for self-correction comes from this openness to scrutiny and challenge.

- « Intelligent Openness »

www.nature.com/nature

Data's sham

Research cannot flourish if data

ture

o. 7261 | 10 September 2009

d must act accordingly.

More and more often these days, a measured not just by the publica the data it makes available to the ing archives such as GenBank have demo such legacy data sets can be for generati cially when data are combined from man in ways that the original researchers coul

All but a handful of disciplines still lac and cultural frameworks required to sup (see pages 168 and 171) — leading to a sharing of data by researchers (see page 16 needs to be addressed by funders, unive themselves.

Research funding agencies need to rec and access to digital data are central to be supported accordingly. Organizatio for instance, have made a good start. The joint information systems responsibility for preserving digital data and making them accessible

igators to do this. One impor- software: tools that streamline a with a description of what the d them, which algorithms have — information that is essential effectively.

a when data can be mixed and software that can keep track of m. Such systems are essential if re ever to give credit — as they of

"Data management should be woven into every course in science."

One PICK of a TALE (1)

"[Researchers would prefer] just one point of access to all data, which would be simple to use and 'fool proof'."

But she suspects it is wishful thinking to ask for Google-like simplicity when one looks for "chlorophyll data in the Atlantic at 200 meters depth"

Looks simple! (Isn't)

Karin Lochte
(Alfred Wegener Institute for Polar and Marine Research)

SEVENTH FRAMEWORK PROGRAMME

APA ALLIANCE FOR PERMANENT ACCESS

HELMHOLTZ | ASSOCIATION

Opportunities for Data Exchange

One of *ODE's* HYPOTHESES

"Without the infrastructure that helps scientists manage their data in a convenient and efficient way, no culture of data sharing will evolve."

Stefan Winkler-Nees
Deutsche Forschungs-Gemeinschaft (DFG)

SEVENTH FRAMEWORK PROGRAMME

APA ALLIANCE FOR PERMANENT ACCESS

HELMHOLTZ | ASSOCIATION

Opportunities for Data Exchange

HELMHOLTZ GEMEINSCHAFT OPEN ACCESS

Earth System Science Data
The Data Publishing Journal

Home
Online Library ESSD
Online Library ESSDD
- Papers in Open Discussion
- Volumes and Issues
- Special Issues
- Most Commented Papers
- Full Text Search
- Title and Author Search

Alerts & RSS Feeds
General Information
Submission

Earth Syst. Sci. Data Discuss., 5, 491-520, 2012
www.earth-syst-sci-data-discuss.net/5/491/2012/
doi:10.5194/essdd-5-491-2012

Article | Discussion | Related Articles

Global marine plankton functional type biomass distributions: coccolithophores

C. J. O'Brien, J. A. Peloquin, M. Vogt, M. Heinle, N. Gruber, P. Ajani, H. Andruleit, J. Arístegui, L. Beaufort, M. Estrada, D. Karentz, E. Kopczyńska, R. Lee, T. Pritchard, and C. Widdicombe

Interactive Discussion — **Status: Open (indefinitely extended)**

AC: Author Comment | **RC**: Referee Comment | **SC**: Short Comment | **EC**: Editor Comment

[Post a Comment] [Subscribe to Comment Alert] - Printer-friendly Version - Supplement

Reviewer: „no effort appears to have been made to engage the specialist scientists who have spent months or years at sea collecting such data. “ - not knowing that:

Authors asked 164 potential contributors – got answer from 13!

HELMHOLTZ GEMEINSCHAFT OPEN ACCESS

Was tun?

Glauben Sie niemandem, der einfache Lösungen propagiert!

Schlüssel - Ziele

- **Nützlichen** Zugang zu Daten schaffen
  - Zugangs-Portale, Data-Mining, …
  - *Sicherung der Rechte*
- **Einfache** Datenablieferung und –dokumentation
  - Werkzeuge und Infrastruktur
  - *Sicherung der Rechte*
- **Änderung** von Bewusstsein und Verhalten
  - Anreiz(?): Evaluierung
  - Druck: Förderbedingung
  - “Kultur” – “So macht man gute Wissenschaft”

Sicherung der Rechte

- **Daten (Fakten) unterliegen nicht dem Urheberrecht!**
- Ihre Nutzung kann (theoretisch) durch Lizenzverträge beschränkt oder frei gegeben werden
  - **Wissenschafts-adäquat: CC0+gute wiss. Praxis**
  - **“Licenses for Europe” behindern die Wissenschaft**
- Der **Zugang kann trotzdem de facto durch kommerzielle Erschließung monopolisiert** und dabei durch eine Vielzahl von Anbietern “balkanisiert” werden.
  - auch Metadaten, Kontext im Blick behalten
  - **Volle Dateninfrastruktur in öffentlicher Verfügung**

Intelligent, qualified Openness (™ Royal Soc.)

- **Accessible** (such that it can readily be found)
- **Intelligible** (to those wishing to understand or scrutinise them)
- **Assessable** (to be able to make some judgment)
- **Usable** (data should be able to be reused)

- There are **legitimate boundaries of openness** which must be maintained in order to **protect commercial value, privacy, safety and security**.

(siehe Science as an Open Enterprise)

Informations-/Dateninfrastruktur 1

- **An den Anforderungen der Disziplinen ausgerichtet**
  - siehe Nutzer-Anforderungen
- **Ermöglicht aber auch interdisziplinäres Forschen**
  - so könnte der größte Nutzen entstehen
- **Internationale Kompatibilität "vom ersten Spatenstich"**
  - gibt es nationale Eigenheiten in der Forschung?
- **Ausbildung von Forschern, "Data Scientists"**
  - zunächst Graduierte, dann Erstsemester

(siehe Grundsätze der dt. Allianz-Initiative demnächst Science Europe?)

Informations-/Dateninfrastruktur 2

- **Kosten, Preise**
  - **ca. 1-2% der Forschungskosten**

    (in D: 200 – 400 Mio € / Jahr)
- **Nachhaltige "Geschäftsmodelle"**
  - **z.B. 70% der Kosten aus Entgelt für Datenabgabe**
  - **müssen bei Antragstellung kalkulierbar seien**

    (zwingt zur Überlegung: was ist gute wiss. Praxis?)

(siehe z.B. Beagrie, Knowledge Exchange)

Sozio-kulturelle Veränderung durch ... „Druck" 1

- **NSF Proposal Preparation Instructions (Jan 2013) Proposals müssen enthalten:**
  - **"A list of: (i) up to five products most closely related to the proposed project; …**
    Acceptable products must be citable and accessible including but not limited to publications, data sets, software, patents, and copyrights."
  - **"Plans for data management and sharing of the products of research. … no more than two pages".**
- **s.a. San Francisco declaration … DFG Qualität statt Quantität**

(www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gpg_2.jsp#IIC2fic)

Sozio-kulturelle Veränderung durch ... „Druck“ 2

- NSF Post Award Requirements
  - Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, …
  - in a form that protects the privacy of individuals and subjects involved. General adjustments and, where essential, exceptions to this sharing expectation may be specified by the funding NSF Program or Division/ Office for a particular field or discipline …

(http://www.nsf.gov/bfa/dias/policy/dmp.jsp)

Erkennbare Problemfelder

- Zugang über “Save Heaven” (medizinische Daten)
  - Bsp.: Helmholtz-Kohorte: 200 M€ Invest – optimale Nutzung vs. Patientenvertrauen
- Embargo-Zeiten: Ein viel weiteres Feld als in der Debatte um 6 vs. 12 Monate (Natur- vs. Geisteswiss.)
  - Satelliten-/Astronomiedaten sofort / nach 1 Jahr
  - Wiss. In anderen Feldern “kämpfen” um 5 Jahre
- “Offen” auch für kommerzielle Nutzung?
  - bester Umgang mit / Effekt von Steuermitteln?
  - US/UK: Ja, weil bester volkswirtschaftl. Effekt; Royal Soc. Report, ESA Sentinel (Erdbeobachtung)

Sozio-kulturelle Veränderung ... durch Einsicht

- **DIESES Verhalten (Offenheit) ist gute Wissenschaft, bringt die Wissenschaft voran**
  (und jenes ist schlechtes Verhalten)
- **Gemeinsame Auffassung und Einsicht IST die einzig nachhaltige Lösung!**
- **Wo formen sich solche Auffassungen?**
  Wissenschaftliche Gesellschaften, Editorial Boards
- **Rolle der Forschungsförderer?**
  Forum für die Communities organisieren - Moderation – Ermittlung und Verbreitung von Best Practise

Was tun? Versuch einer Hilfe

- **Das jetzt Machbare+Wirkungsvolle+Vorbildliche fördern**
  - **Gesamtprozess wird Dekaden dauern!**
- **Nicht das Rad wieder erfinden (lassen)!**
  - **Geförderte Projekt, Infrastrukturen**
  - **“Eigene” Policies**
- **“Best Policies”: NSF, ERC, EU, …**
  - **Gemeinsam den Stand ermitteln, weiter entwicklen!?**
- **Foren:**
  - **Science Europe (GRC?)**
  - **Research Data Alliance (Implementierbarkeit!)**

Nützliche Links

- „Grundsätze zum Umgang mit Forschungsdaten“; www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsaetze
- “intelligent openness”, siehe “Science as an Open Enterprise”, Report der Royal Society, Juni 2012; royalsociety.org/policy/projects/science-public-enterprise/report/
- “NSF Proposal Preparation Instructions”, Jan 2013; www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gpg_2.jsp#IIC2fic
- “NSF Dissemination and Sharing of Research Results”; www.nsf.gov/bfa/dias/policy/dmp.jsp
- “NSF Division of Ocean Sciences Sample and Data Policy”, 2011; www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf11060
- Beagrie/JISC, Keeping Research Data Safe“; www.beagrie.com/krds.php
- Knowledge Exchange, Daten Arbeitsgruppe; www.knowledge-exchange.info/Default.aspx?ID=284
- ODE Projekt, „Briefing Sheets“ und „Ten Tales of Drivers & Barriers in Data Sharing“, 2012; www.ode-project.eu/ode-outputs