Weierstrass Institute for
Applied Analysis and Stochastics

Leibniz Association

How WIAS handles mathematical research data

Thomas Koprucki, Karsten Tabelow, and Ilka Kleinod

Situation of research data today

- tremendous growth in data generation (huge to massive data)
- modeling and simulation tools readymade to users from applied disciplines
- multi-dimensional, multi-resolution, multi-modal data
- from experiment, simulation or data analysis
- large, unstructured long tail
- plethora of specific data formats and technical standards

Drawbacks:

- letterpress-oriented publication culture (graphs not data)
- paper-oriented recognition of achievements

Emerging:

- Data publications
- Data repositories
- Data sharing culture

Data journals (launched 2-3 years ago)

journals.elsevier.com

Journals – Forschun... | Journal - Elsevier | ScienceDirect.com | Geoscience Data Jo... | Home : Scientific Data | Home : Scientific Data

ELSEVIER

Home › Books & Journals › Data in Brief

Data in Brief

Editor-in-Chief: Hao-Ran Wang

View full editorial board

JOURNAL TOOLS

Get New Content Alerts

Get RSS feed

Save to My Profile

JOURNAL MENU

Journal Home

FIND ISSUES

Current Issue

All Issues

FIND ARTICLES

Early View

RMetS Royal Meteorological Society

Geoscience Data Journal

Wiley

Geoscience Data Journal

Edited By: Dr Rob Allan, Met Office, UK

Online ISSN: 2049-6060

SCIENTIFIC DATA

Search | Go

Advanced search

Nature Publishing Group

Home | Archive | About ▾ | For Authors ▾ | For Referees | Data Policies ▾ | Collections ▾

Featured Data Descriptor

Plant traits, productivity, biomass and soil properties from forest sites in the Pacific Northwest, 1999-2014

Berner & Law | 19th January 2016

Computational models are widely used to explore the impacts of climate change, and to assess potential responses. But, such models often rely

About *Scientific Data*

Scientific Data is an open-access, peer-reviewed journal for descriptions of scientifically valuable datasets. Our primary article-type, the **Data Descriptor**, is designed to make your data more discoverable, interpretable and reusable.

Browse publications by date ▸

Browse publications by subject ▸

Research Data Repositories for MMS

datashare.is.ed.ac.uk

Edinburgh DataShare

THE UNIVERSITY of EDINBURGH

INFORMATION SERVICES

Contact us

DSpace Startseite / College of Science & Engineering / School of Mathematics

School of Mathematics

Volltextsuche:

Los

Research in the School of Mathematics is undertaken in a wide variety of areas of the mathematical sciences, including pure, applied, statistics, operational research as well as mathematical physics. The research environment is enhanced by the International Centre for Mathematical Sciences and by close collaboration with Heriot-Watt University through the Maxwell Institute.

Research groups include:

Suche

DSpace Suche

In diesem Bereich

MEIN BENUTZERKONTO

Einloggen

Registrieren

STÖBERN

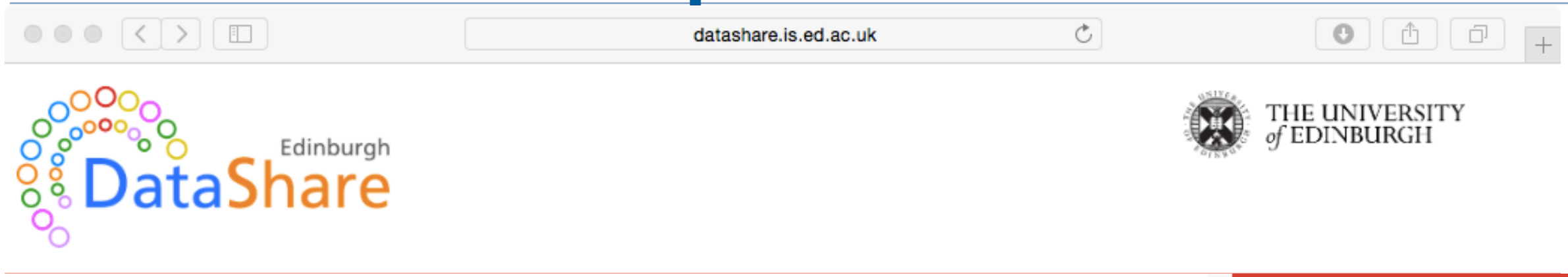
Gesamter Bestand

Bereiche & Sammlungen

Diesen Bereich

Titeln

Research Data Repositories for MMS



Usual way of data sharing

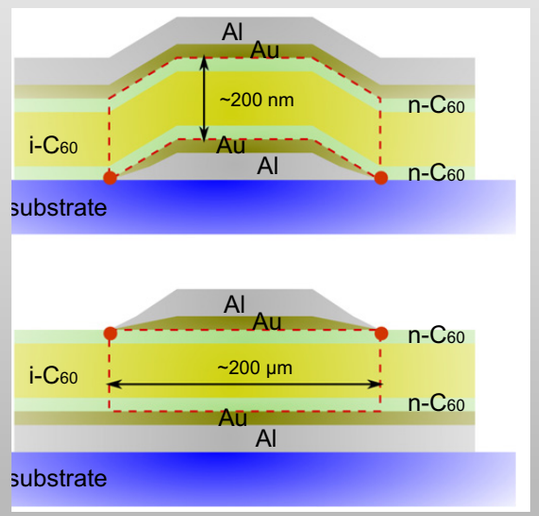
Modeling

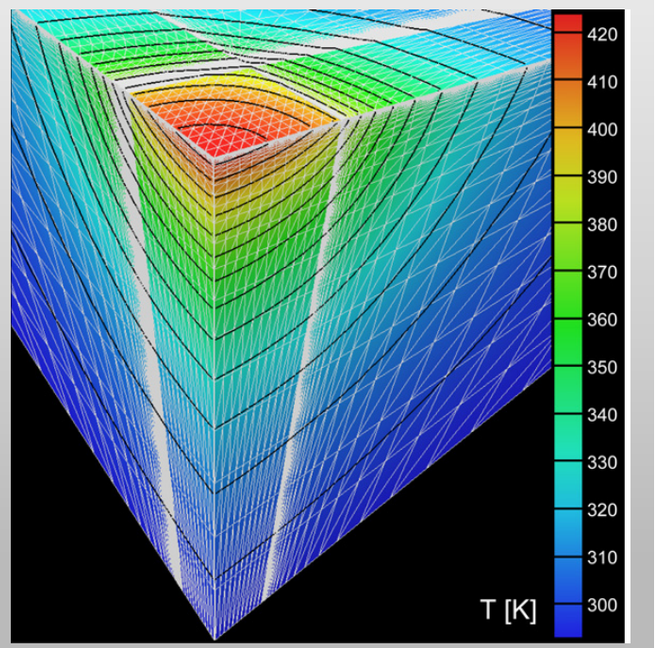
Simulation

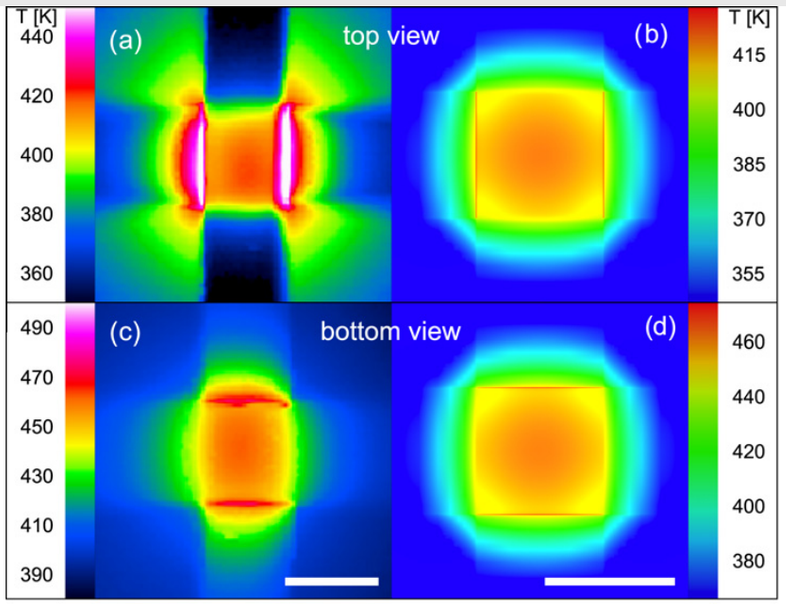
Data Analysis

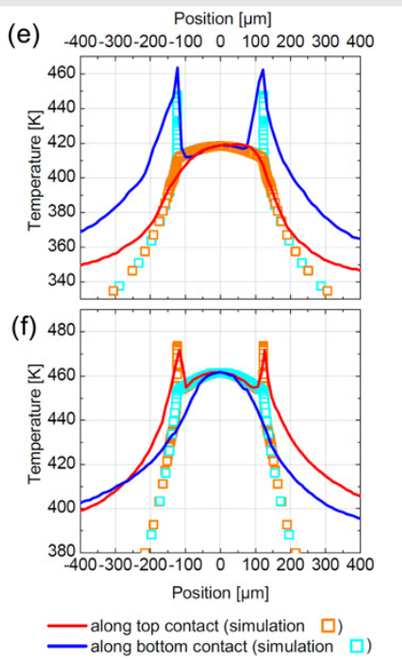
iapp

Application Partner

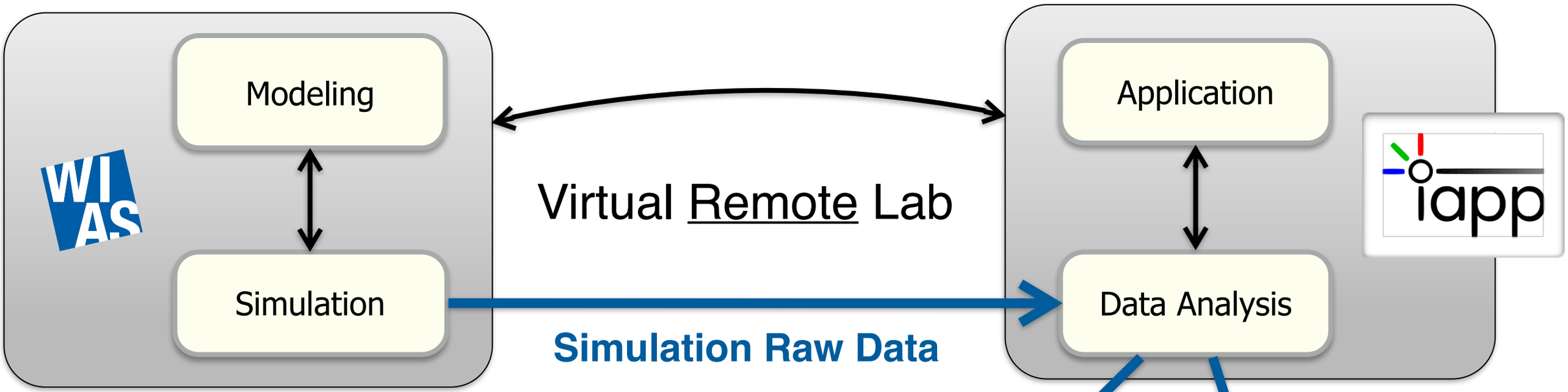


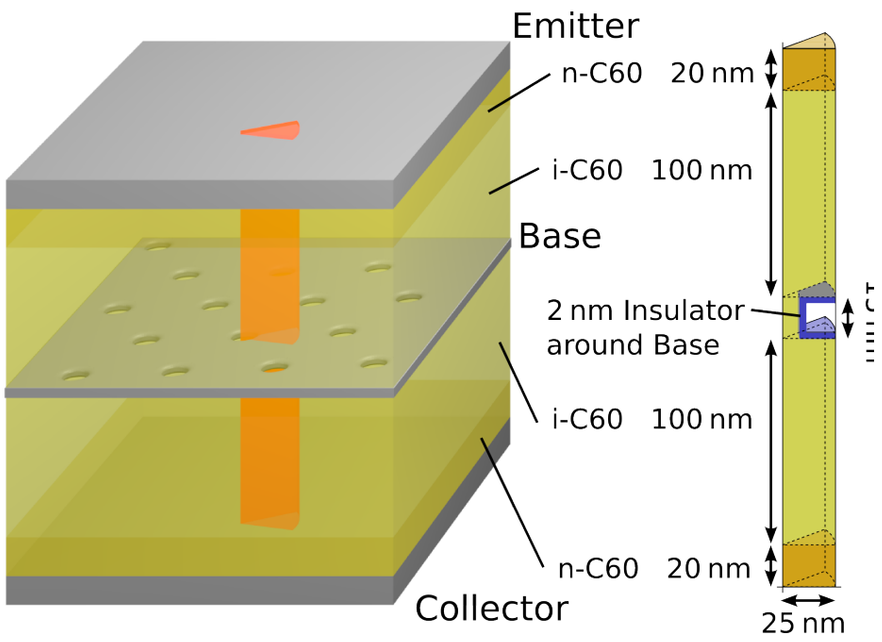


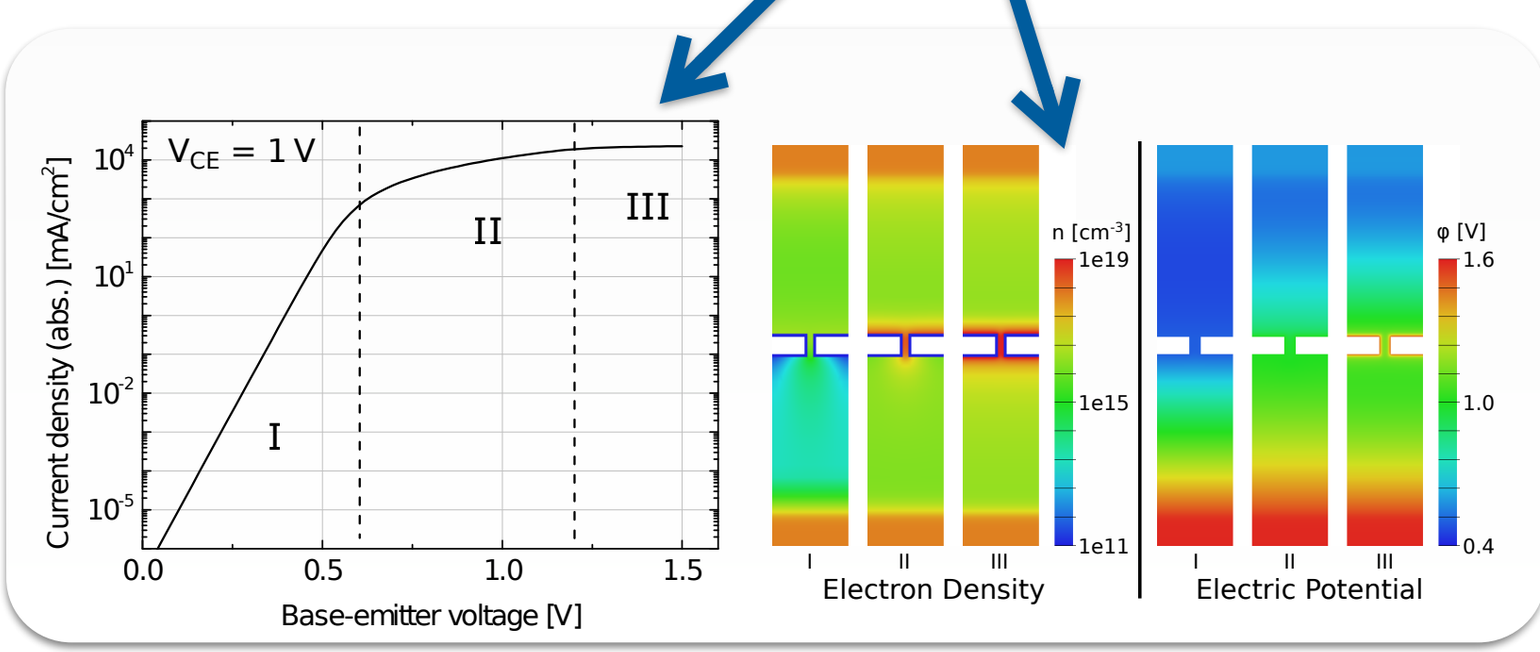




Added value of data sharing

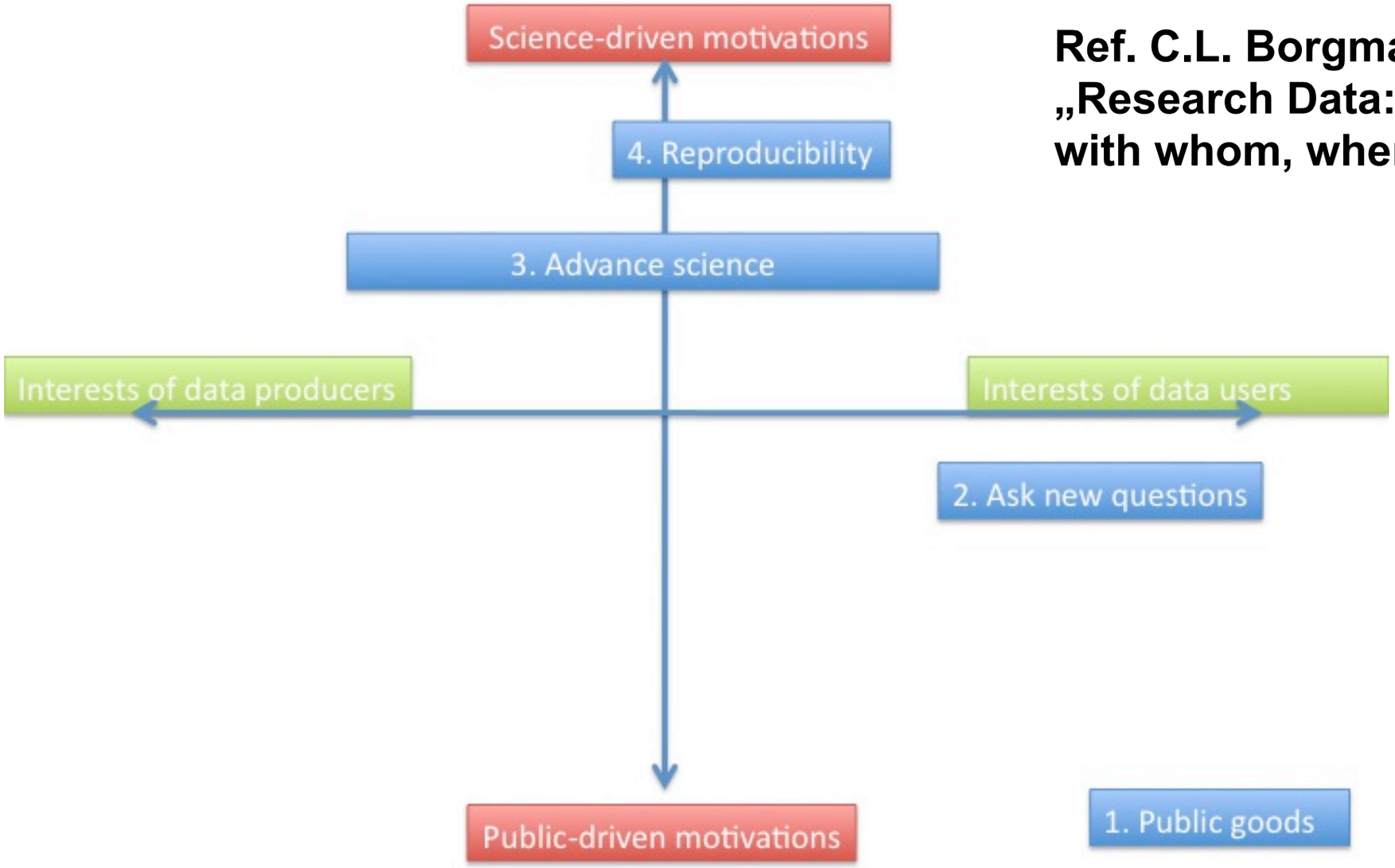






better science by full exploitation of data

Motivation: Data Sharing

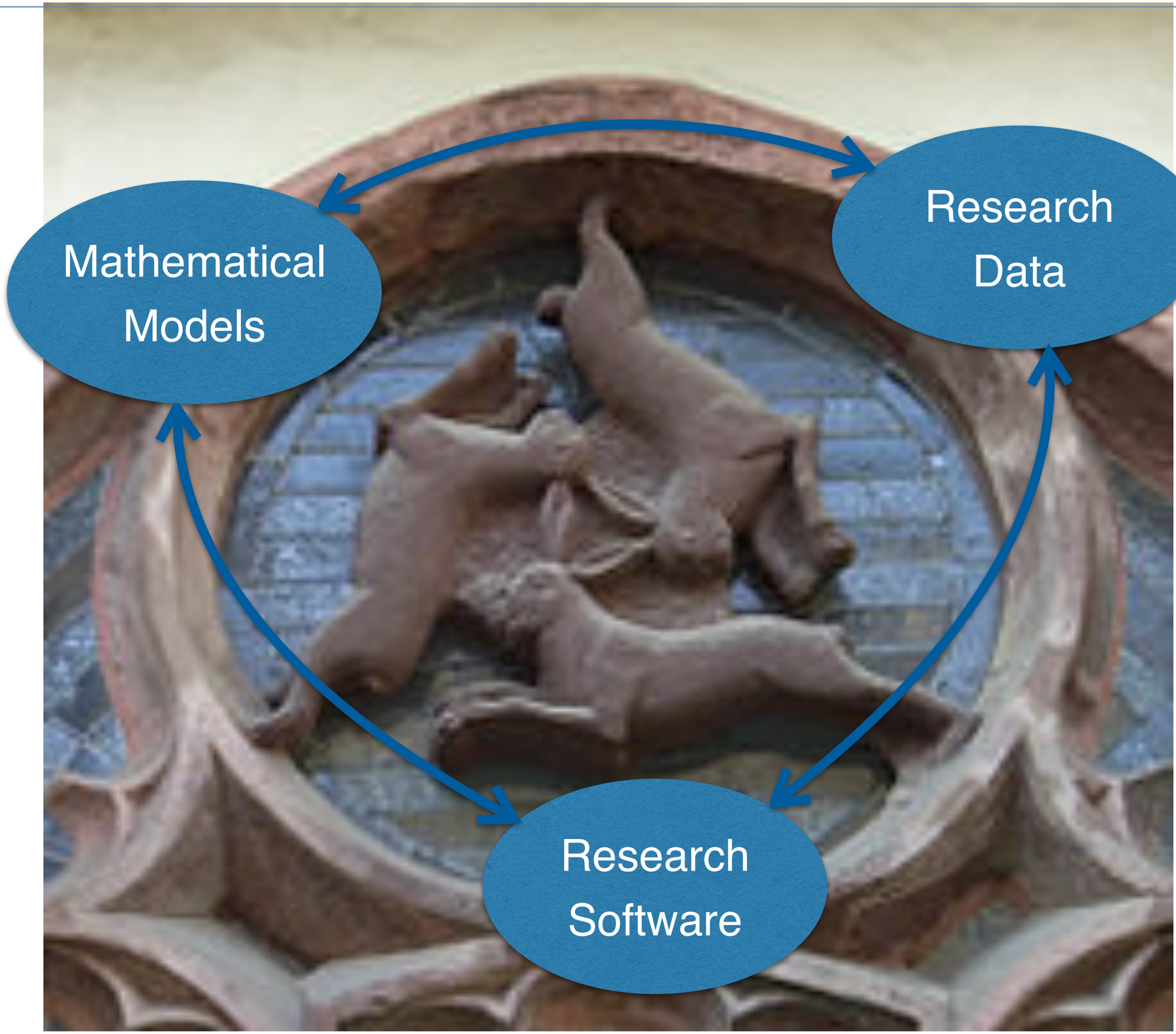


Ref. C.L. Borgmann (UCLA)
„Research Data: Who will share what, with whom, when, and why?“, 2010

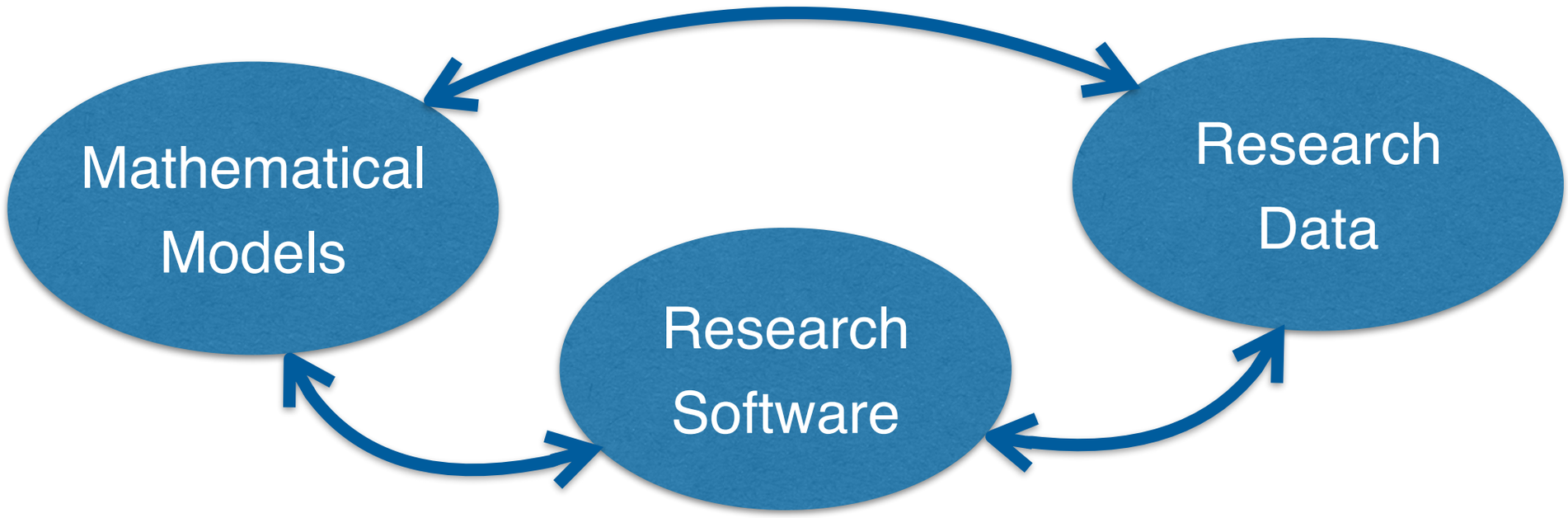
EU Commission:

[..] **optimise the circulation, access to and transfer of scientific knowledge**. [..]

Three rabbits



The ideal world



- Unique Identifier (DOI) for Models, Data and Software
- Smart Data Repositories (like Dropbox, ownCloud)
- Community specific but widely accepted standards and tools
- Digital Mathematical Models and Algorithms; Repositories
- Data-aware publication and recognition culture
- ePubs and Smart Services for interrelation and reuse

better data, better research

Institutions and Funding Organizations

OpenData Pilot in H2020 (for 20% of projects)
incl. Data Management Plan (DMP, mandatory) etc.

Institutions and Funding Organizations

Deutsche Forschungsgemeinschaft

DFG Guidelines on the Handling of Research Data

Research data is an essential foundation for scientific work. The diversity of this data reflects the wide range of different scientific disciplines, research interests and research methods. 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 should therefore also be categorised as research data. The long-term archiving and accessibility of research data contributes to the traceability and quality of scientific work and enables researchers to carry on work begun by others. The Alliance of Science Organisations in Germany voiced its support for the long-term archiving of research data, open access to it and compliance with the conventions of individual disciplines in the "Principles for the Handling of Research Data", adopted in 2010[1]. The "Guidelines on the Handling of Research Data" put the framework stipulated by the Principles into a concrete form in the DFG's funding arrangements.

The following general guidelines apply for applicants submitting proposals to the DFG:

1. Project planning and submission of proposals
Applicants should consider during the planning stage whether and how much of the research data resulting from a project could be relevant for other research contexts and how this data can be made available to other researchers for reuse. Applicants should therefore detail in the proposal what research data will be generated or evaluated during a scientific research project. Concepts and considerations appropriate to the specific discipline for quality assurance and the handling and long-term archiving of research data should be taken as a basis. The relevant explanations must contain information about data types, discipline-specific standards (if applicable) and the choice of suitable repositories, if these are available for a given research area or particular data types. Details should also be provided on any third-party rights affected and preliminary planning for the data publication schedule.

2. Accessibility
Assuming that the publication of research data from a DFG-funded project does not conflict with the rights of third parties (in particular data protection or copyright), research data should be made available as soon as possible. Data should be made accessible at a stage of processing that allows it to be usefully reused by third parties (raw data or structured data). To make sure this is the case, it must be ensured that access to the data is still guaranteed when, through publication, the rights of use relating to research data are transferred to a third party, usually a publishing house.

3. Long-term archiving
In accordance with the rules of good scientific practice, research data should be archived in the researcher's own institution or an appropriate nationwide infrastructure for at least 10 years.

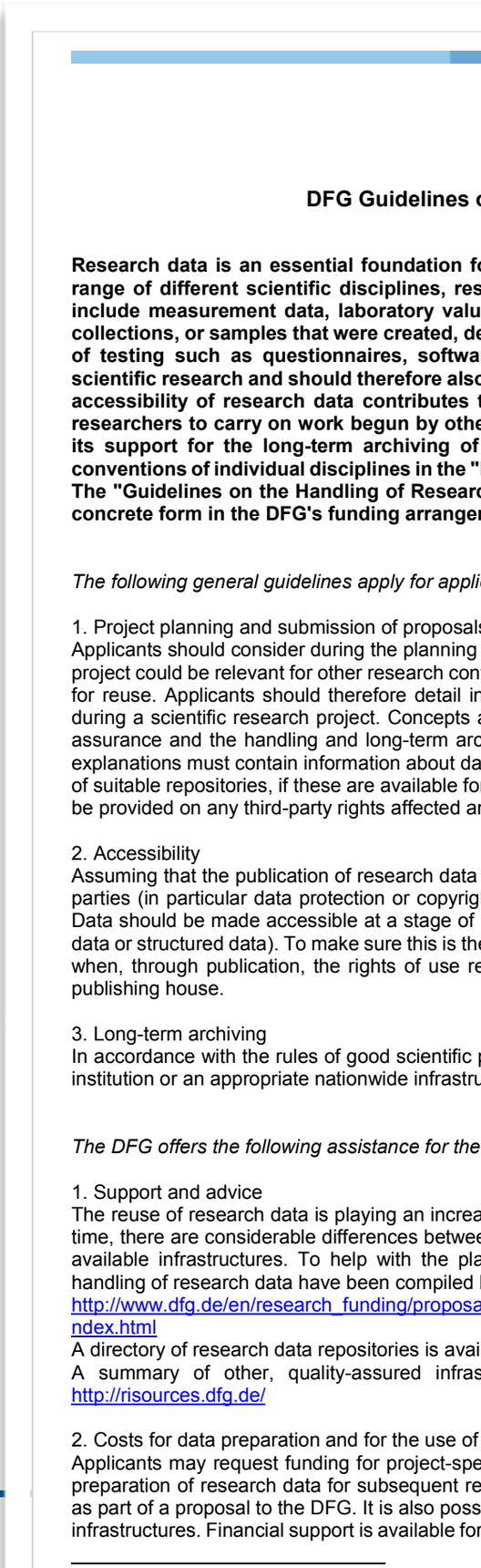
The DFG offers the following assistance for the implementation of the guidelines:

1. Support and advice
The reuse of research data is playing an increasingly important role in nearly all scientific disciplines. At the same time, there are considerable differences between disciplines with regard to the current state of discussion and the available infrastructures. To help with the planning of research projects, information and suggestions for the handling of research data have been compiled here:
http://www.dfg.de/en/research_funding/proposal_review_decision/applicants/submitting_proposal/research_data/i

Institutions and Funding Organizations

From preamble of “DFG Guidelines on Handling of Research Data” (Adopted by the Senate of the DFG at Sep. 30, 2015):

- “Research data is an essential foundation for scientific work.”
- ”The diversity of this data reflects the wide range of different scientific disciplines, research interests and research methods.”
- “Research data might include … audiovisual information, texts, …”
- “… **software and simulations may also produce important results for scientific research and should therefore also be categorised as research data.**“
- “The long-term archiving and accessibility of research data contributes to the traceability and quality of scientific work and enables researchers to carry on work begun by others.”



DFG Guidelines

Research data is an essential foundation fo range of different scientific disciplines, res include measurement data, laboratory valu collections, or samples that were created, de of testing such as questionnaires, softwa scientific research and should therefore also accessibility of research data contributes researchers to carry on work begun by othe its support for the long-term archiving of conventions of individual disciplines in the " The "Guidelines on the Handling of Researc concrete form in the DFG's funding arrange

The following general guidelines apply for appli

1. Project planning and submission of proposals
Applicants should consider during the planning project could be relevant for other research con for reuse. Applicants should therefore detail in during a scientific research project. Concepts a assurance and the handling and long-term arc explanations must contain information about da of suitable repositories, if these are available fo be provided on any third-party rights affected ar

2. Accessibility
Assuming that the publication of research data parties (in particular data protection or copyrig Data should be made accessible at a stage of data or structured data). To make sure this is the when, through publication, the rights of use re publishing house.

3. Long-term archiving
In accordance with the rules of good scientific institution or an appropriate nationwide infrastru

The DFG offers the following assistance for the

1. Support and advice
The reuse of research data is playing an increa time, there are considerable differences betwee available infrastructures. To help with the pla handling of research data have been compiled http://www.dfg.de/en/research_funding/proposa ndex.html
A directory of research data repositories is avai A summary of other, quality-assured infras http://risources.dfg.de/

2. Costs for data preparation and for the use of
Applicants may request funding for project-spe preparation of research data for subsequent re as part of a proposal to the DFG. It is also poss infrastructures. Financial support is available for

[1] http://www.allianzinitiative.de/en/core-activities/rese

Institutions and Funding Organizations

Abgeordnetenhaus BERLIN

Drucksache 17/2512
21.10.2015

17. Wahlperiode

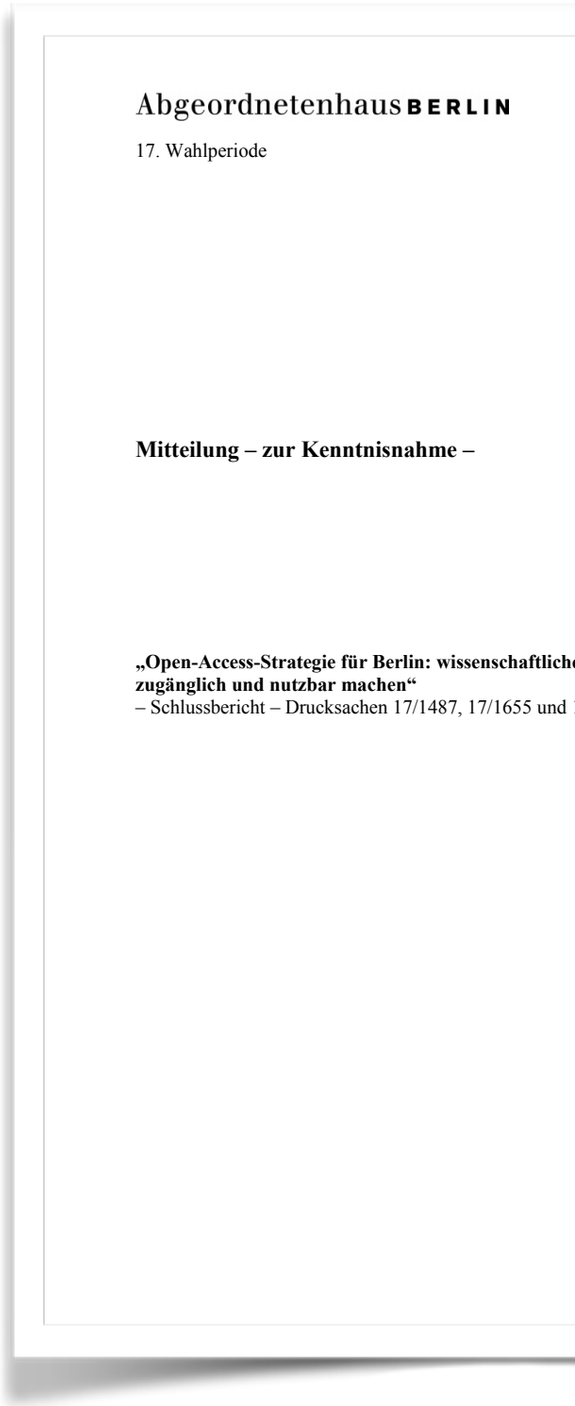
Mitteilung – zur Kenntnisnahme –

„Open-Access-Strategie für Berlin: wissenschaftliche Publikationen für jedermann zugänglich und nutzbar machen“
– Schlussbericht – Drucksachen 17/1487, 17/1655 und 17/2024

Institutions and Funding Organizations

From “Drucksache 17/2512, 21.10.2015 des Abgeordnetenhaus Berlin”

- “Forschungsdaten: **International weniger weit entwickelt ist das Handlungsfeld „Open Access zu Forschungsdaten“**. Die Einrichtungen des Landes Berlin bilden hier keine Ausnahme.”
- “Der **offene Zugang zu Forschungsdaten** und deren **umfassende Nachnutzung** sollen gewährleistet werden. Erforderliche Beiträge des Landes Berlin können nur im Rahmen von national und international abgestimmten Strategien geleistet werden, die noch formuliert werden müssen.”
- “Die Open-Access-Strategien der **wissenschaftlichen Einrichtungen in Berlin sollen eine Forschungsdaten-Policy enthalten**; die Wissenschaftlerinnen und Wissenschaftler sollen ausdrücklich ermutigt werden, die **Daten zitierfähig zur Verfügung zu stellen**. “
- “Für den so genannten *Long-Tail* der Forschungsdaten – also Datensätze mit geringem Volumen, die in **verschiedenen Datenformaten** vorliegen und **schwer standardisierbar** sind – **fehlen in vielen Fachgebieten angemessene Infrastrukturen**.”


Abgeordnetenhaus BERLIN

17. Wahlperiode

Mitteilung – zur Kenntnisnahme –

„Open-Access-Strategie für Berlin: wissenschaftliche zugänglich und nutzbar machen“
– Schlussbericht – Drucksachen 17/1487, 17/1655 und

Steps at WIAS

- Research software and data (and mathematical models) play an important role for the transfer of our research to industry and other disciplines
- Workgroup to develop an Research Data Policy (since end 2014)
- Based on experiences of formulation of a software policy (adpoted at WIAS in May 2015)
- discussion on the topic at WIAS
  - definition for math (!)
  - best practices and guidelines
  - technical infrastructure