Research data "publishing" – models, roles and responsibilities

Sünje Dallmeier-Tiessen

CERN | Scientific Information Service

Humboldt Universität, Berlin

Outline

Research data "publishing"

- Introduction: Research data

- Models
  - Text-centered occurrence
  - Data-centered occurrence
  - Object-centric principle

- Roles
  - Motivation, Competences, Responsibilities

- Conclusions

Research data

What is research data?

How does it look like?

How can this be integrated into modern scholarly communication?

Foto: AWI Bremerhaven, Gutt
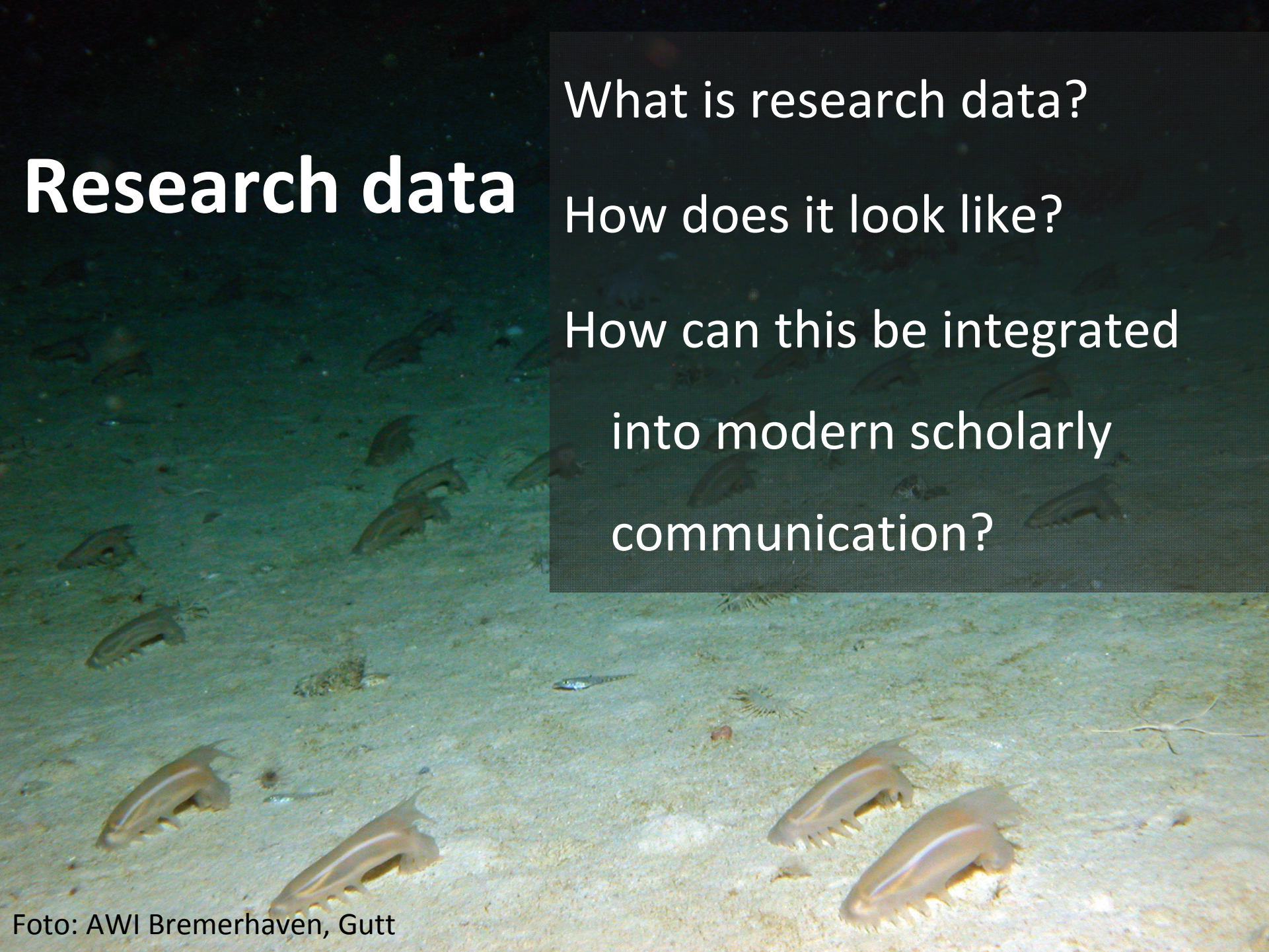
Research data "publishing" beyond disciplines?

Discipline specific characteristics of research data, BUT challenges are similar:

- Non existing data publishing culture
  - Missing infrastructure
  - Non established workflows
  - Missing awareness/motivation/submission

    "we need to get this paper out" vs. "we need to publish this dataset"

- Diverse research datasets
  - Complexity: From simple files to complex datasets
  - Size: From small to "huge"
 - Documentation: varying requirements

Models

How does the world of research data publishing look like today?

stand-alone, text centered, data centered
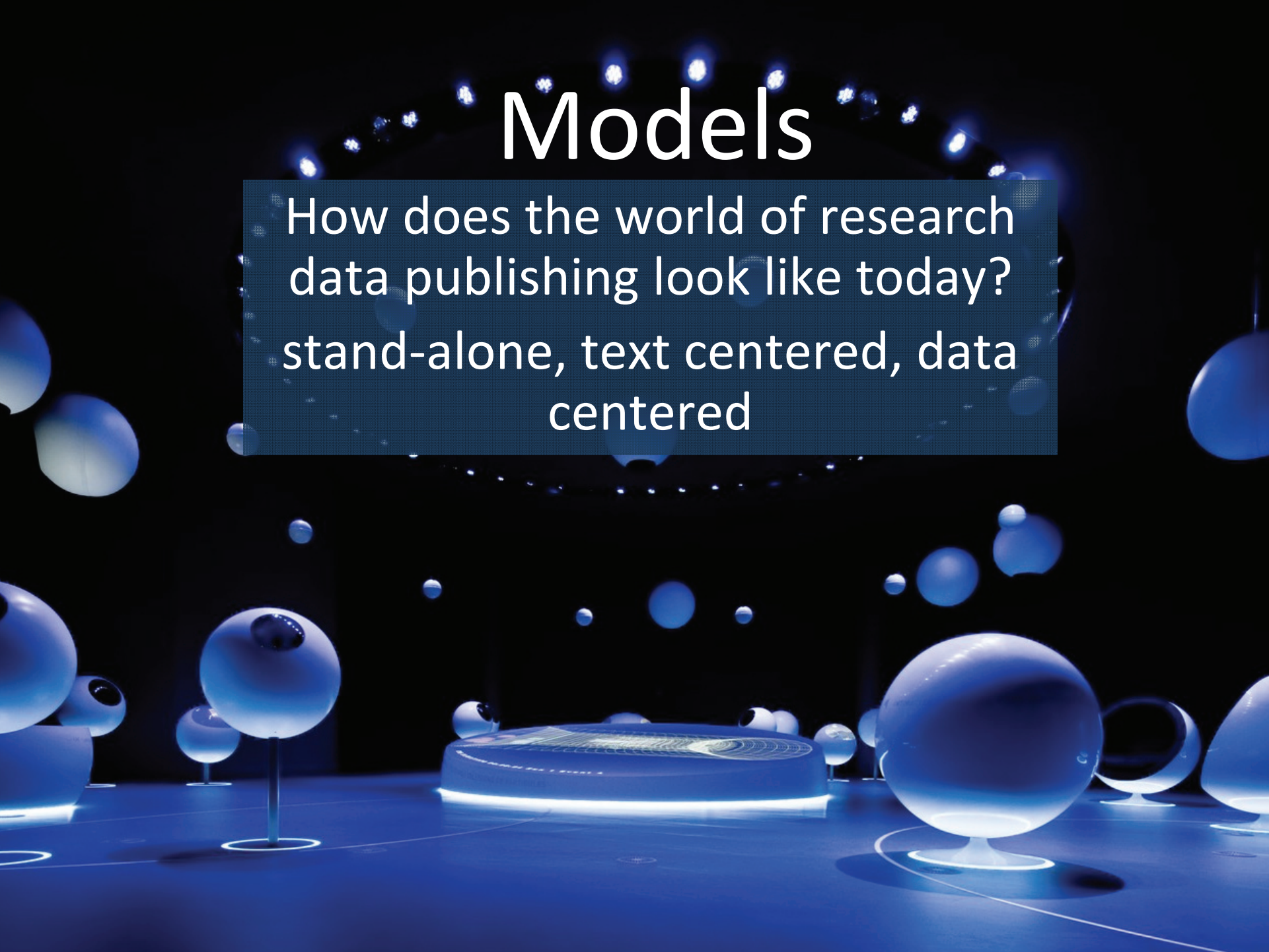
Occurence:
text centered publication model

- Long tradition

- Linked research data in external data respositories, e.g. DataCite

- Attached supplemental material: tables, graphs, audio/video/data files (on same plattform as text publication)

Sunje Dallmeier-Tiessen | CERN
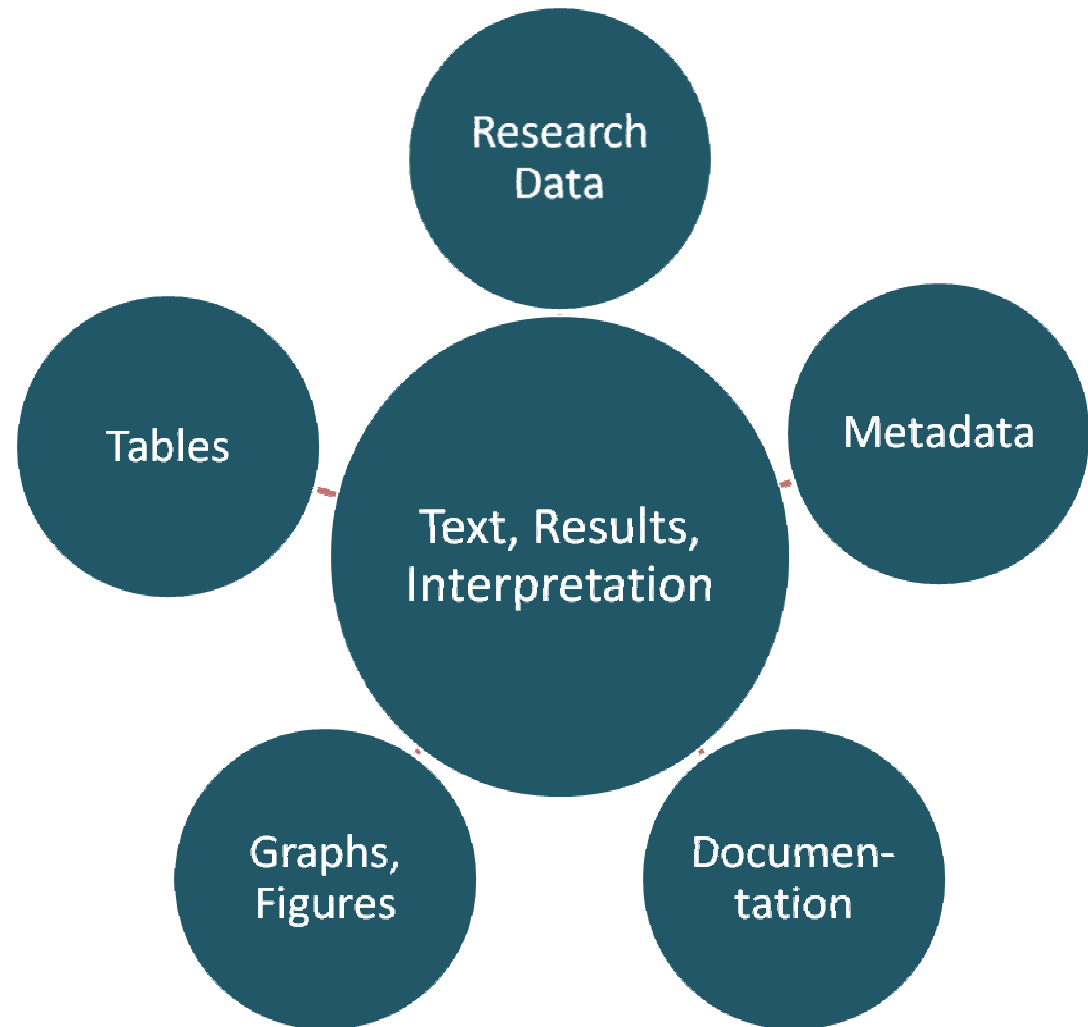
Supplement

Subscribe   Register   Submit manus

nature
International weekly journal of science

Search   This journal

Journal home > Archive > Brief Communications > Full text > Supplementary information

Supplementary information

From the following article:
Language evolution: Semantic combinations in primate calls
Kate Arnold & Klaus Zuberbühler
Nature **441**, 303(18 May 2006)
doi:10.1038/441303a

▾ Download plugins and applications

Supplementary Methods
This file contains Supplementary Figures S1–S5

📕 Download PDF file (402KB)

Audio clip 1
A series of 'pyow' calls: these can function as an alarm in response to a nearby leopard, but are also used in other contexts.
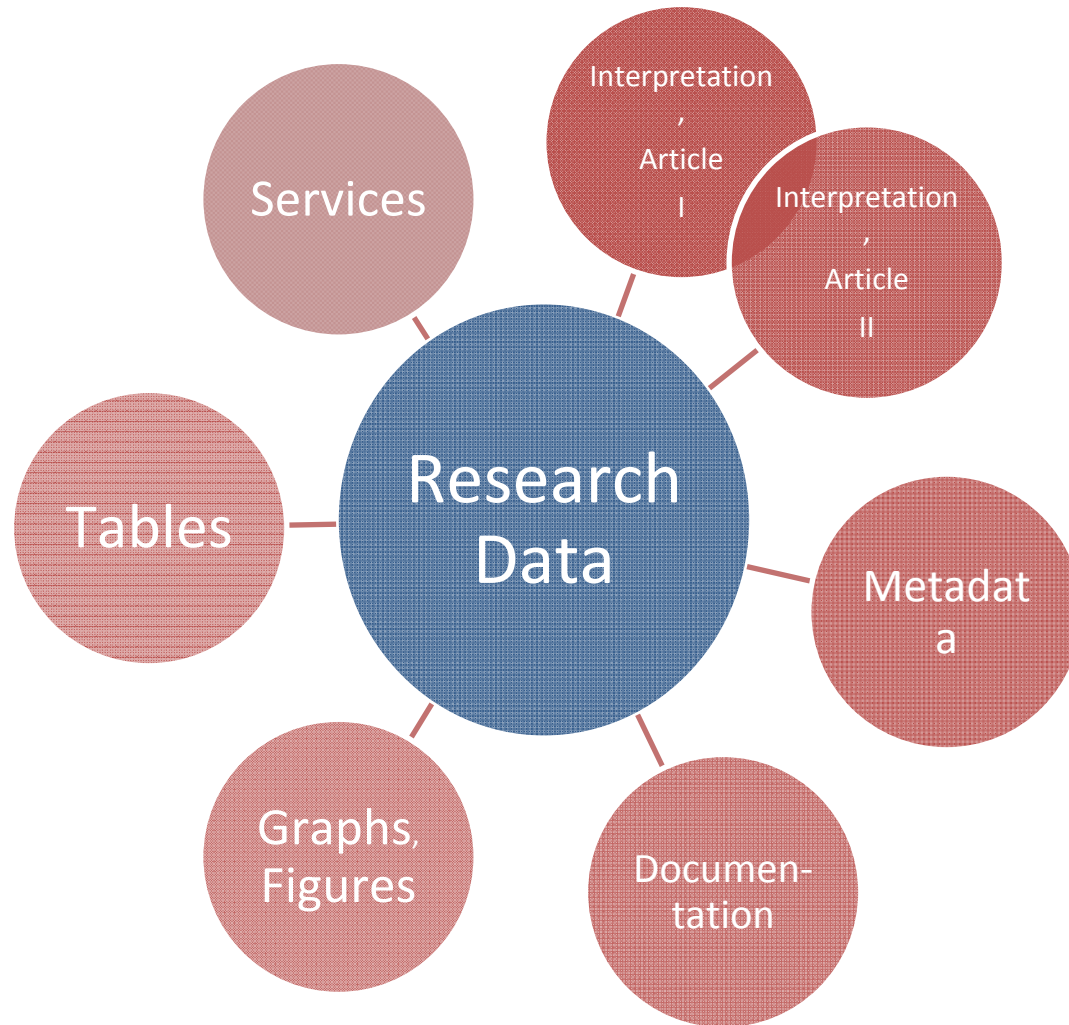
🔈 Download Audio file (2MB)

Audio clip 2
A series of 'hack' calls: mostly functions as an alarm in response to a nearby eagle.

🔈 Download Audio file (385KB)

Source: Nature 2006

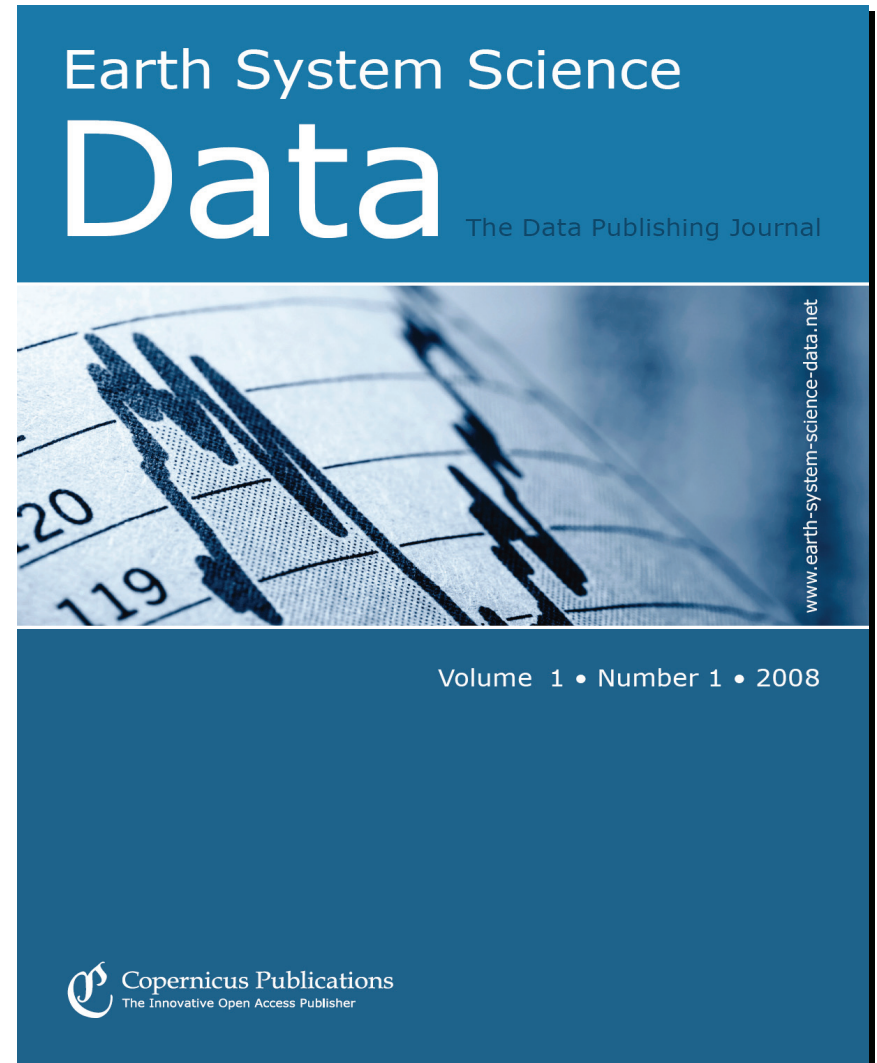# Occurrence: data centered publication model

Data Journal

Data centered publication

- including data documentation

- data access

- quality assurance – peer review

- open access

Incentive: publication for enhanced data provision

Earth System Science **Data** The Data Publishing Journal
www.earth-system-science-data.net
Volume 1 • Number 1 • 2008
Copernicus Publications
The Innovative Open Access Publisher

Direct data access

Abstract

On 22 May 1985 the first balloon-borne ozonesonde was successfully launched by the staff of Georg-Forster-Station (70°46′ S, 11°41′ E). The following weekly ozone soundings mark the beginning of the continuous investigation of Germany to study the vertical ozone distribution in the southern hemisphere.

In 1985 these ozone soundings have been the only record showing the change of vertical ozone distribution in the southern polar stratosphere in September and October. The regular ozone soundings from 1985 until 1992 are a valuable reference data set since the chemical ozone loss became a significant feature in the southern polar stratosphere.

The balloon-borne soundings were performed at the upper air sounding facility of the neighbouring station Novolazarevskaya, just 2 km apart from Georg-Forster-Station. Till 1992, ozone soundings were taken without interruption. Afterwards, the ozone sounding program was moved to Neumayer-Station (70°39′ S, 8°15′ W) 750 km further west.

Data coverage and parameter measured

Repository-Reference: doi:10.1594/PANGAEA.547983
Available at: http://dx.doi.org/10.1594/PANGAEA.547983
Coverage: East: 11.6800, South: -70.7700
Location Name: Georg-Forster-Station, Antarctica
Date/Time Start: 1985-05-22T05:19:00
Date/Time End: 1992-01-29T01:19:00

ESSDD

1, 1–13, 2008

Antarctic ozonesonde profiles

G. König-Langlo and H. Gernandt

Title Page

Abstract | Instruments

Data Provenance & Structure

Tables | Figures

|◄ | ►|
◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

With modern information technology, many opportunities

… to service the different needs in the research communities:

- When do I want to publish my data?
    - After production?
    - Before or after the "text" publication?
- And with whom would I like to share?
    - With my colleagues?
    - With my community?
    - With the entire world?
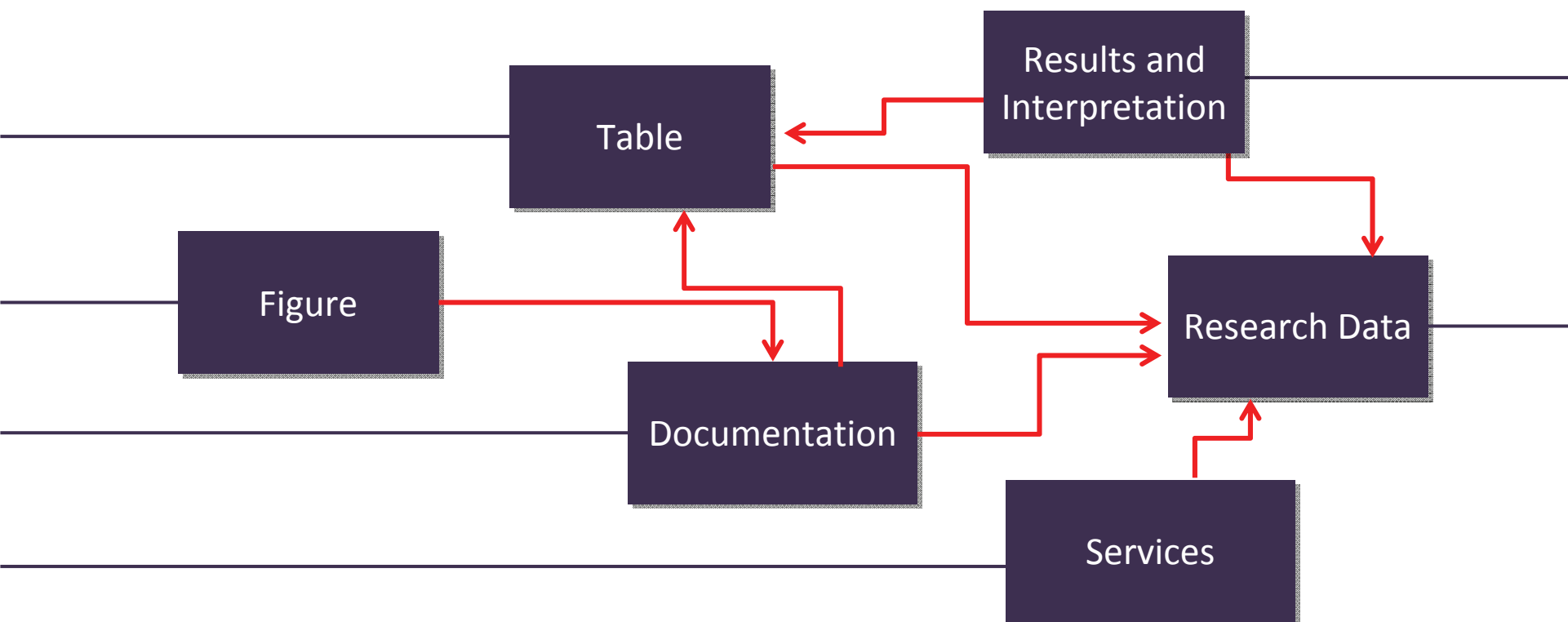   - … or maybe this might change over time?

➔Different models and responsibilities for involved parties

Looking into the future:
Principle – object centered ("mash up")

Materials and services as individually addressed objects
- Persistence of objects/services,
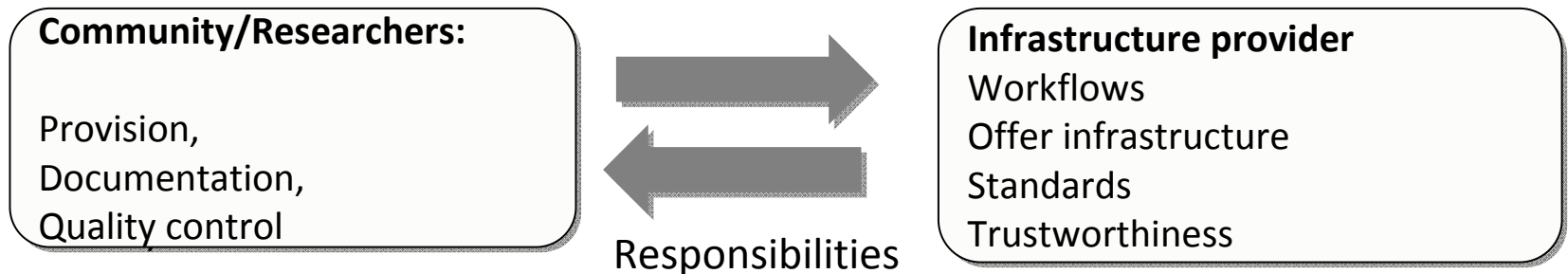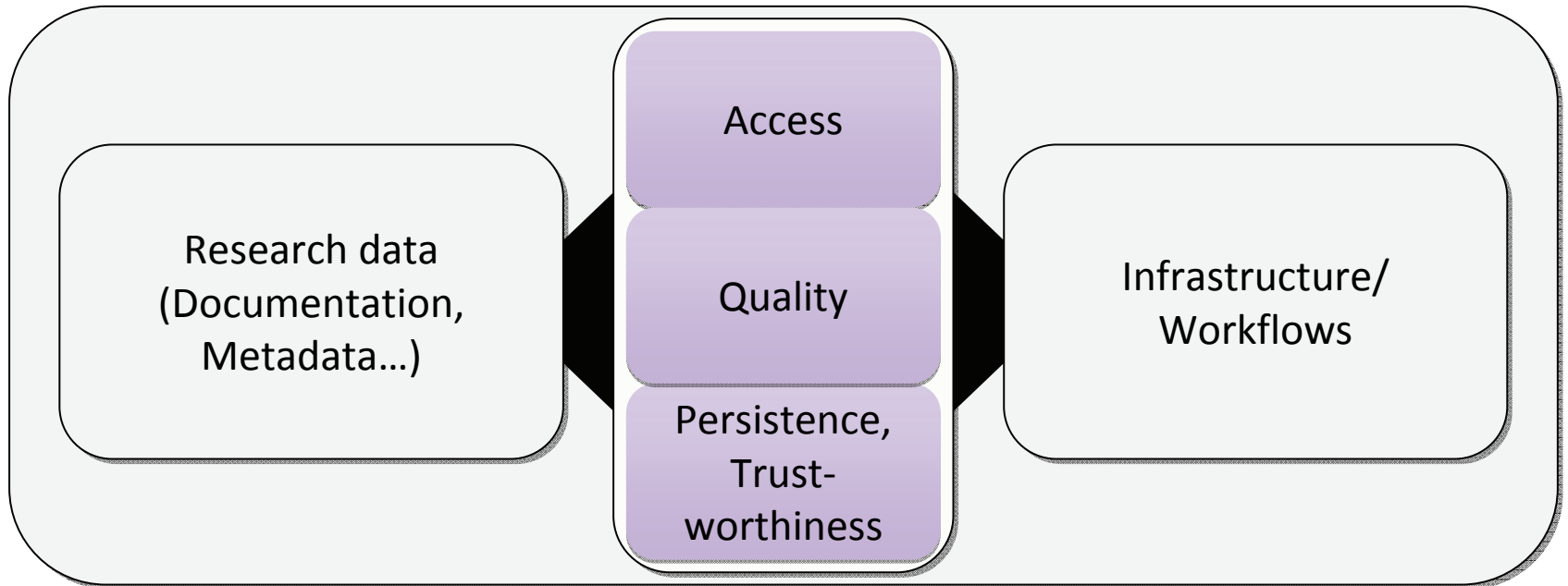- Persistent identifier

Hands-on!?

Information Science/ Publishers

IT/ Computing

Research Community

after Corral, 2008
Foto by katera, flickr

Prerequisites for Research Data Publishing
Cooperation of researchers and service providers

Research data (Documentation, Metadata…)

Access

Quality

Persistence, Trust-worthiness

Infrastructure/ Workflows

Community/Researchers:

Provision, Documentation, Quality control

Responsibilities

Infrastructure provider
Workflows
Offer infrastructure
Standards
Trustworthiness

Role of researchers

Foto by G. F. Wicke, flickr

… as data providers
[data production]

- Data documentation
- Data submission (preparation of publishable dataset)
- Data quality assurance/peer review

… as data users

- Citation of dataset, acknowledgement

Source: Nature 2006

Role of information scientists

Service

- Management of the research material remix
  - Facilitating new workflows
  - Integration of new features: e.g. data citation
  - Preservation, catalogue
- Interface between IT and researcher
  - Make publishing platforms usable
    - Submission
    - Search functionality

foto: flickr

Conclusions

- Missing data publishing culture in many communities

- New publishing models allow tailored services, developed with the community

- Significant role of community in research data preservation and publishing
  - Implications for content recruitment

QUESTIONS

sunje.dallmeier-tiessen@cern.ch