

„Data Librarianship“ - Rollen, Aufgaben, Kompetenzen

98. Deutscher Bibliothekartag

Erfurt, 05.06.2009

Heinz Pampel, Helmholtz-Gemeinschaft

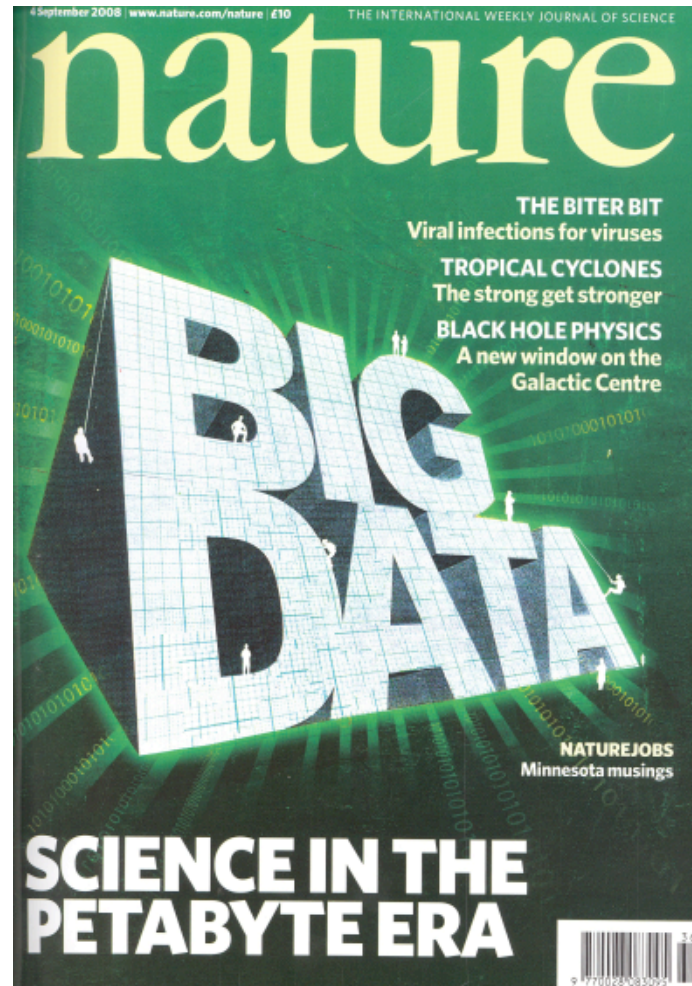
Roland Bertelmann, Deutsches GeoForschungsZentrum – GFZ

Prof. Dr. Hans-Christoph Hobohm, Fachhochschule Potsdam

Agenda

- Relevanz
- Forschungsdatenmanagement
- „Data Librarianship“
- Anforderungen und notwendige Kompetenzen
- Internationales Curriculummodell

Forschungsdaten – Thema für Bibliotheken?



[1]

Relevanz: Helmholtz-Gemeinschaft



Helmholtz Open Access Projekt [2]

Relevanz: Helmholtz-Gemeinschaft

The screenshot shows the website for the Study Group for Data Preservation and Long Term Analysis in High Energy Physics (DPHEP). The page features a navigation menu with links for Home, People, Committees, Subgroups, Workshops, Documents, and Work Space. The main content area is titled "Study Group on Data Preservation and Long Term Analysis in High Energy Physics" and includes a paragraph explaining the group's purpose and objectives. Below the text is the DPHEP logo, which consists of the letters "DPHEP" in a stylized font above a red house-like shape, and the SLAC logo (National Accelerator Laboratory) below it. A notice at the bottom of the page states: "2nd Workshop on Data Preservation and Long Term Analysis in High Energy Physics will be held at SLAC, May 26-28, 2009." On the right side of the page, there is a search bar and a "Working Groups" section with a list of four groups: WG1: Physics Case, WG2: Models, WG3: Governance, and WG4: Technologies, each with a checked checkbox.

DPHEP Study Group for Data Preservation and Long Term Analysis in High Energy Physics

Home People Committees Subgroups Workshops Documents Work Space


DPHEP

Study Group on Data Preservation and Long Term Analysis in High Energy Physics

High Energy Physics experiments initiate with this **Study Group** a common reflection on **data persistency and long term analysis** in order to get a common vision on these issues and create a multi-experiment dynamics for further reference. It is foreseen that the workshop is a starting point for intense consultations and common work on defined subject. A second follow-up workshop will take place mid 2009.

The objectives of the workshop are:

- *Review and document the physics objectives of the data persistency in HEP.
- *Exchange information concerning the analysis model: abstraction, software, documentation etc. and identify coherence points.
- *Address the hardware and software persistency status
- *Review possible fundings programs and other related international initiatives.
- *Converge to a common set of specifications in a document that will constitute the basis for future collaborations.



2nd Workshop on Data Preservation and Long Term Analysis in High Energy Physics will be held at SLAC, May 26-28, 2009.

Working Groups

- WG1: Physics Case
- WG2: Models
- WG3: Governance
- WG4: Technologies

Deutsches Elektronen-Synchrotron DESY [3]



Relevanz: Helmholtz-Gemeinschaft

The screenshot shows the website of the Helmholtz-Zentrum Berlin. The header includes the logo and navigation links like 'Intranet', 'Mein Intranet', 'Sitemap', 'Kontakt', and 'Impressum'. A search bar is visible with the text 'Erweiterte Suche' and 'Suche: In allen Bereichen'. The main content area is titled 'Archivierung von Forschungsdaten' and contains text about data archiving practices, including a section 'Archivieren oder Speichern?' and 'Verlauf und Stand der gegenwärtigen Diskussion'. On the right side, there are two sidebars: 'Kontakt Lesesaalpult' with contact information (phone, fax, email) and 'Ansprechpartner' featuring a photo and contact details for Dr. Wolfgang Fritsch.

Helmholtz-Zentrum Berlin für Materialien und Energie [4]



Relevanz: Helmholtz-Gemeinschaft

Publication and Citation of Scientific Primary Data

Home

- + About STD-DOI
- + Documents
- + Contact

About STD-DOI

- Demonstration of the DOI resolver
- STD-DOI Publication Agents
- Citation of Data
- Publication of Data

supported by
Deutsche
Forschungsgemeinschaft
DFG

STD-DOI - Creating access to scientific Data

"Publication and Citation of Scientific Primary Data" (STD-DOI) is a project funded by the German Science Foundation. Its aim is to make primary scientific data citeable as publications. In this system, a data set would be attributed to its investigators as authors like it would be done for a work in the conventional scientific literature. Thus, scientific primary data should not exclusively be understood as part of a scientific publication, but may have its own identity. [Download and view a short MS Powerpoint presentation on the idea behind STD-DOI: [Deutsch / English](#)]

How can I publish data and make them citeable? Follow this link to find out more about [publication of data](#).

How do I cite data publications? Follow this link to find out more about [citation of data](#).

How can I participate as a data publication agent? Follow this link to find out more about becoming a [data publication agent](#).

```
graph TD; IF[International DOI Foundation] --> GHS[Global Handle System]; IF --> DNB[German National Library (DNB) (URN Node)]; IF --> TIB[TIB Hannover Registration Agency]; TIB --> GFZ[GFZ Potsdam Geophysics]; TIB --> MPIM[M+D / MPIM Climate Models]; TIB --> MARUM[AWI / MARUM Env. Observ.]; TIB --> DLR[DLR-DFD Satellite Data];
```

Deutsches GeoForschungsZentrum - GFZ ^[5]



Relevanz: Helmholtz-Bibliotheken



ALBERT

All Library Books, journals and Electronic Records Telegrafenberg

Simple Search

Advanced Search

Journals A-Z

Saved Results (0)

Info

Bathymetry of Lake Baikal

GO

1 hit(s) - Search time 0.0s

Results: Books (0) | Journal Titles (0) | [Open Access](#) (1) | More (0) | **All Sources** (1)

Search history: [Data](#) | [lake](#) | [Bathymetry of Lake Baikal](#)

Sort by: **Relevance** | [Year ascending](#) | [Year descending](#) | [Title ascending](#) | [Title descending](#)

Hits per page: [5](#) | **10** | [20](#) | [50](#) | [100](#)

Selection: [All](#) | [None](#) | [Reverse](#) | [Save](#) | [Mail](#) | [Export](#)

Go to no.:



	Year	Title	Author
1. <input type="checkbox"/>	2007	[DATA] Bathymetry of Lake Baikal 10.1594/GFZ.SDDDB.1204 S·F·X	Marc De Batist Charlet Francois

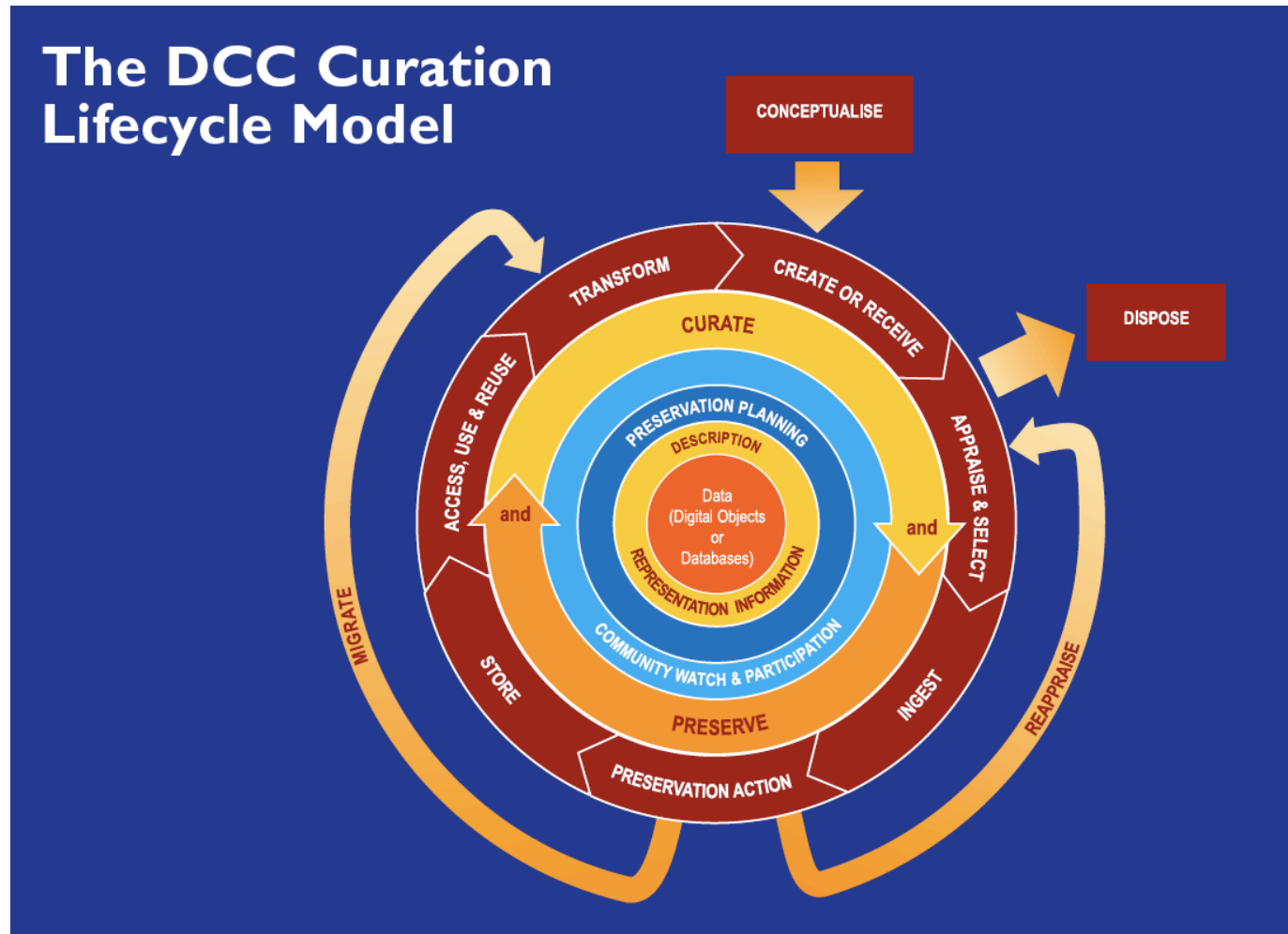
[Top](#)

hosted by
kobv

Deutsches GeoForschungsZentrum - GFZ ^[6]

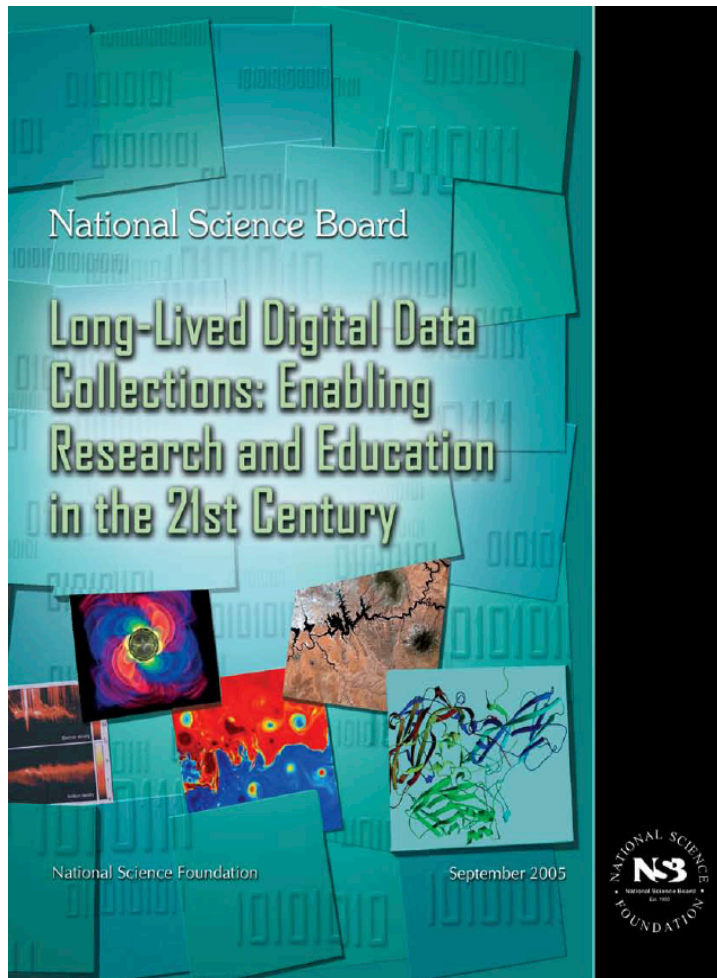


Forschungsdatenmanagement



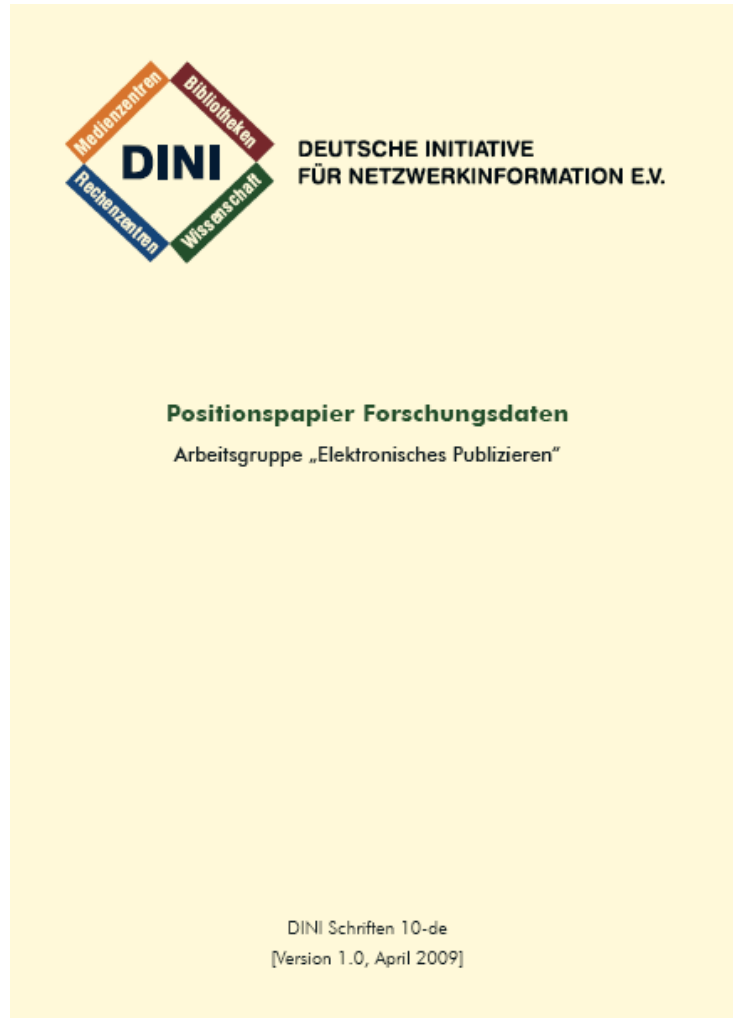
[7]

Forschungsdatenmanagement



- Digital Data Collections^[8]
 - Research data collections
 - Resource or community data collections
 - Reference data collections

Forschungsdatenmanagement



- Fünf Aspekte
 - Disziplinarität
 - Organisation
 - Technik
 - Zugang
 - Qualität

[9]

„Data Librarianship“

IASSIST



**International Association for
Social Science Information Service
and Technology**

About IASSIST
Activities
Membership

IASSIST is an international organization of professionals working with information technology and data services to support research and teaching in the social sciences.

Its 300+ members work in a variety of settings, including data archives, statistical agencies, research centers, libraries, academic departments, government departments, and non-profit organizations.

Objectives

- To encourage and support local and national information centers for social science data.
- To foster international exchange and dissemination of information regarding substantive and technical developments related to social science data.
- To coordinate international programs, projects, and general efforts that provide a forum for discussion of issues relating to social science data.
- To promote the development of standards for social science data.
- To encourage educational experiences for personnel engaged in work related to these objectives.

Annual Conference

A spring annual conference provides social science data professionals with ...

- Unique workshops focused on the profession
- Stimulating panels, poster and plenary sessions
- Networking with peers from around the world
- International conference venues reflect our membership base. See webpage for past and future venues

Committees and Action Groups

- Administrative Committee
- Education Committee
- Membership Committee
- Publications Committee
- International Outreach
- Strategic Plan Action Group
- Technology Action Group

IASSIST welcomes new members. The IASSIST Membership year is based on a July-to-June calendar.

Benefits of Membership include:

- Peer to peer support and involvement in an international network of data professionals
- Access to the IASSIST QUARTERLY publication and the IASSIST Communiqué blog.
- Participation in IASST-L, the organization's electronic list.
- Opportunities to serve on IASSIST committees and action groups.

Regional Secretaries

Regional Secretaries provide support for activities within their designated regions.

They exist for the following:

Africa	Asia-Pacific
Canada	Europe
United States	

For More Information...

More information about membership, past and future conferences and the work of the committees and action groups can be found at the website:

www.iassistdata.org

- „Data Library“
- University of Edinburgh (1983)
- University of Oxford (1988)
- London School of Economics & Political Science (1997)



Y (degrees)	Y_min (degrees)	Y_max (degrees)	Y_min/Y_max (%)	Y (degrees)	Y_min (degrees)	Y_max (degrees)	Y_min/Y_max (%)	Y (degrees)	Y_min (degrees)	Y_max (degrees)	Y_min/Y_max (%)
11.6	4.406	1.983	-0.063	19.9	5.091	2.913	-0.052	31.0	20.00	1.961	-0.019
14.7	40225	1.884	-0.066	20.0	50049	2.9016	-0.053	31.5	19706	1.907	-0.015
14.8	39967	1.893	-0.065	20.2	29708	1.8745	-0.053	32.0	19492	1.883	-0.017
14.9	39713	1.897	-0.065	20.4	28893	1.8470	-0.054	32.5	19220	1.895	-0.012
15.0	39462	1.910	-0.063	20.6	28254	1.8219	-0.052	33.0	18968	1.816	-0.014
15.1	39211	1.785	-0.060	20.8	28859	1.7964	-0.050	33.5	18719	1.8076	-0.010
15.2	38960	1.7629	-0.064	21.0	28700	1.7713	-0.051	34.0	18470	1.7845	-0.008
15.3	38712	1.7293	-0.065	21.2	28448	1.7467	-0.052	34.5	18224	1.7617	-0.008
15.4	38468	1.7163	-0.064	21.4	28196	1.7220	-0.050	35.0	18018	1.7399	-0.008
15.5	38222	1.6937	-0.065	21.6	27951	1.6970	-0.047	35.5	17778	1.7186	-0.005
15.6	38020	1.6713	-0.063	21.8	27710	1.6737	-0.049	36.0	17585	1.6980	-0.005
15.7	37790	1.6491	-0.063	22.0	27474	1.6529	-0.049	36.5	17318	1.6780	-0.006
15.8	37561	1.6271	-0.060	22.2	27242	1.6305	-0.048	37.0	17117	1.6586	-0.006
15.9	37339	1.6055	-0.063	22.4	27014	1.6086	-0.044	37.5	16902	1.6398	-0.003
16.0	37117	1.5844	-0.064	22.6	26791	1.5893	-0.044	38.0	16703	1.6215	-0.003
16.1	36896	1.5630	-0.063	22.8	26571	1.5688	-0.043	38.5	16509	1.6036	-0.001
16.2	36682	1.5421	-0.063	23.0	26355	1.5489	-0.041	39.0	16340	1.5865	-0.001
16.3	36469	1.5215	-0.063	23.2	26143	1.5295	-0.041	39.5	16186	1.5696	-0.001
16.4	36258	1.5011	-0.063	23.4	25933	1.5104	-0.040	40.0	16038	1.5535	-0.006
16.5	36049	1.4810	-0.062	23.6	25730	1.4916	-0.040	40.5	15926	1.5376	-0.006
16.6	35843	1.4611	-0.062	23.8	25539	1.4631	-0.042	41.0	15764	1.5222	-0.006
16.7	35639	1.4414	-0.063	24.0	25351	1.4460	-0.041	41.5	15608	1.5072	-0.007
16.8	35438	1.4220	-0.062	24.2	25166	1.4292	-0.041	42.0	15457	1.4926	-0.008
16.9	35239	1.4028	-0.061	24.4	24982	1.4038	-0.037	42.5	15310	1.4784	-0.009
17.0	35043	1.3838	-0.062	24.6	24797	1.3796	-0.038	43.0	15167	1.4645	-0.009
17.1	34848	1.3650	-0.062	24.8	24622	1.3557	-0.039	43.5	15028	1.4511	-0.011
17.2	34656	1.3465	-0.062	25.0	24450	1.3322	-0.035	44.0	14892	1.4380	-0.012
17.3	34466	1.3281	-0.062	25.2	24281	1.3091	-0.035	44.5	14759	1.4252	-0.011
17.4	34278	1.3100	-0.061	25.4	24125	1.2864	-0.034	45.0	14631	1.4128	-0.010
17.5	34093	1.2921	-0.061	25.6	23971	1.2641	-0.033	45.5	14508	1.4006	-0.010
17.6	33909	1.2743	-0.062	25.8	23819	1.2420	-0.036	46.0	14383	1.3888	-0.011
17.7	33728	1.2568	-0.061	26.0	23671	1.2201	-0.035	46.5	14264	1.3774	-0.011
17.8	33548	1.2395	-0.059	26.2	23528	1.2005	-0.034	47.0	14147	1.3661	-0.011
17.9	33371	1.2221	-0.061	26.4	23385	1.1819	-0.030	47.5	14034	1.3552	-0.011
18.0	33195	1.2048	-0.059	26.6	23245	1.1643	-0.031	48.0	13924	1.3445	-0.011
18.1	33021	1.1886	-0.060	26.8	23107	1.1476	-0.028	48.5	13816	1.3341	-0.011
18.2	32850	1.1729	-0.060	27.0	22973	1.1314	-0.028	49.0	13711	1.3240	-0.010

[10]

„Data Librarianship“

■ Skills, Role & Career Structure of Data Scientists & Curators ^[11]

Key Perspectives

Consultants in scholarly information

THE SKILLS, ROLE AND CAREER
STRUCTURE OF DATA SCIENTISTS
AND CURATORS: AN ASSESSMENT OF
CURRENT PRACTICE AND FUTURE
NEEDS

REPORT TO THE JISC

July 2008

Prepared by:
Alma Swan and Sheridan Brown
Key Perspectives Ltd
48 Old Coach Road
Playing Place
Truro
TR3 6ET
UK
+44 1392 879702
aswan@keyperspectives.co.uk
www.keyperspectives.co.uk

- **Data creator:** researchers with domain expertise who produce data. These people may have a high level of expertise in handling, manipulating and using data
- **Data scientist:** people who work where the research is carried out – or, in the case of data centre personnel, in close collaboration with the creators of the data – and may be involved in creative enquiry and analysis, enabling others to work with digital data, and developments in data base technology
- **Data manager:** computer scientists, information technologists or information scientists and who take responsibility for computing facilities, storage, continuing access and preservation of data
- **Data librarian:** people originating from the library community, trained and specialising in the curation, preservation and archiving of data

„Data Librarianship“

■ Skills, Role & Career Structure of Data Scientists & Curators ^[11]

Key Perspectives

Consultants in scholarly information

THE SKILLS, ROLE AND CAREER
STRUCTURE OF DATA SCIENTISTS
AND CURATORS: AN ASSESSMENT OF
CURRENT PRACTICE AND FUTURE
NEEDS

REPORT TO THE JISC

July 2008

Prepared by:
Alma Swan and Sheridan Brown
Key Perspectives Ltd
48 Old Coach Road
Playing Place
Truro
TR3 6ET
UK
+44 1392 879702
aswan@keyperspectives.co.uk
www.keyperspectives.co.uk

- The library and information science community should have an important role to play in the data science arena, particularly in delivering awareness and understanding of data issues and the importance of good data science and data curation
 - Training researchers to be more data-aware
 - Adopting a data archiving and preservation role
 - The training and supply of data librarians

„Data Librarianship“

■ E-Science Talking Points for ARL Deans & Directors^[12]

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, 2006
Association of Research Libraries
<http://www.arl.org/rtl/escience/>

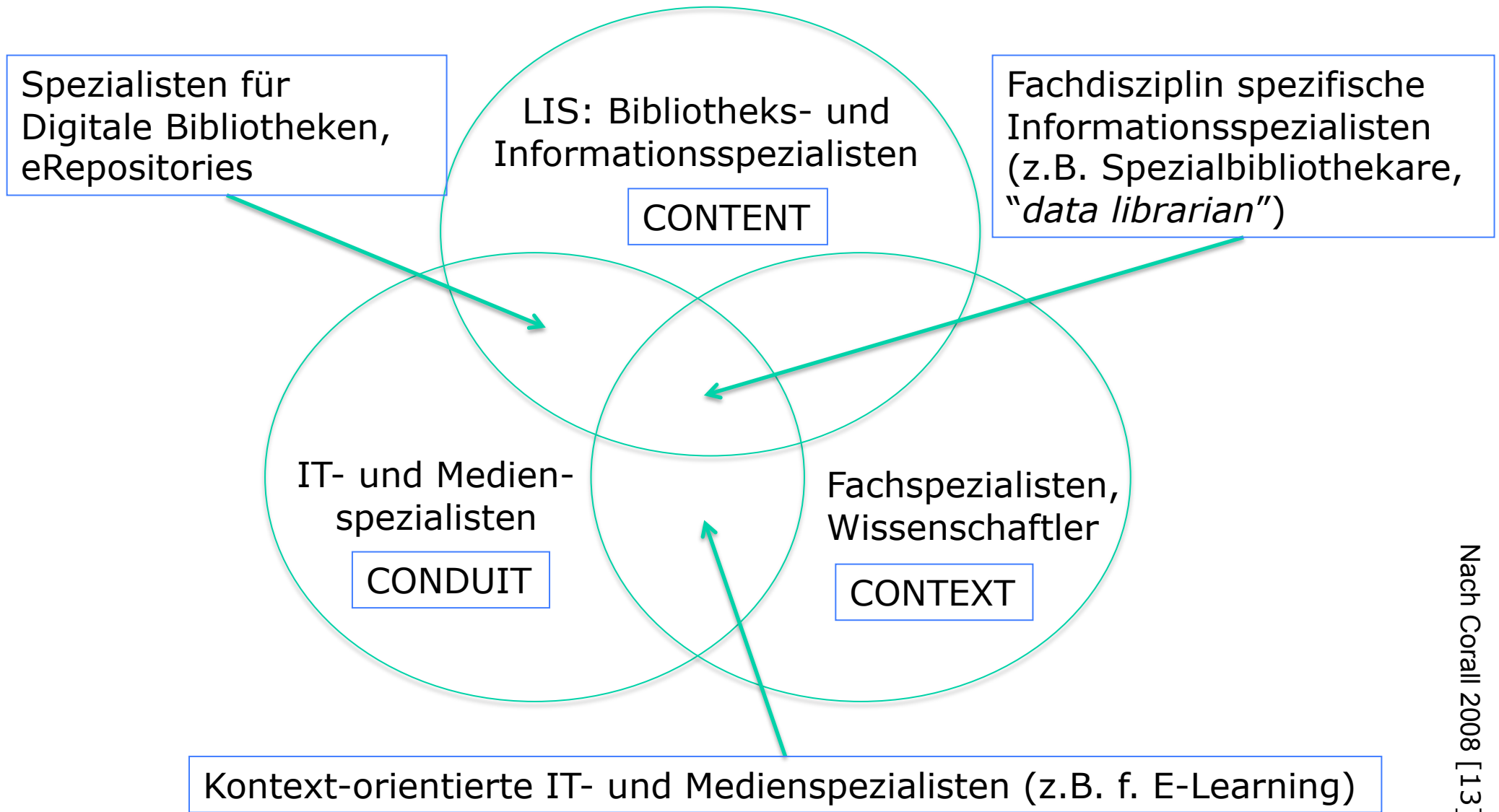


- Data management, including collection, organization, description, curation, archiving, and dissemination
- Creation of new data- and scholarship-based electronic resources for university and/or public use
- Development of new models, standards, and architectures for various aspects of data management, 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

„Data Librarianship“

- Handlungsfelder für Bibliotheken und BibliothekarInnen
 - Sensibilisierung
 - Bibliothek als Ansprechpartner
 - Kooperation
 - Wissenschaftsmanagement, Fachwissenschaft, Datenzentren, Rechenzentren
 - Infrastruktur
 - Forschungsdaten-Repositoryn (Standards, Interoperabilität, Open Access und Langzeitarchivierung)
 - Services
 - z.B. Einwerbung, Erschließung und Nachweis, Unterstützung bei der Qualitätssicherung, Publikation, Rechtliche Aspekte
 - Aus- und Weiterbildung

„Informationsspezialisten in hybrider Umgebung“

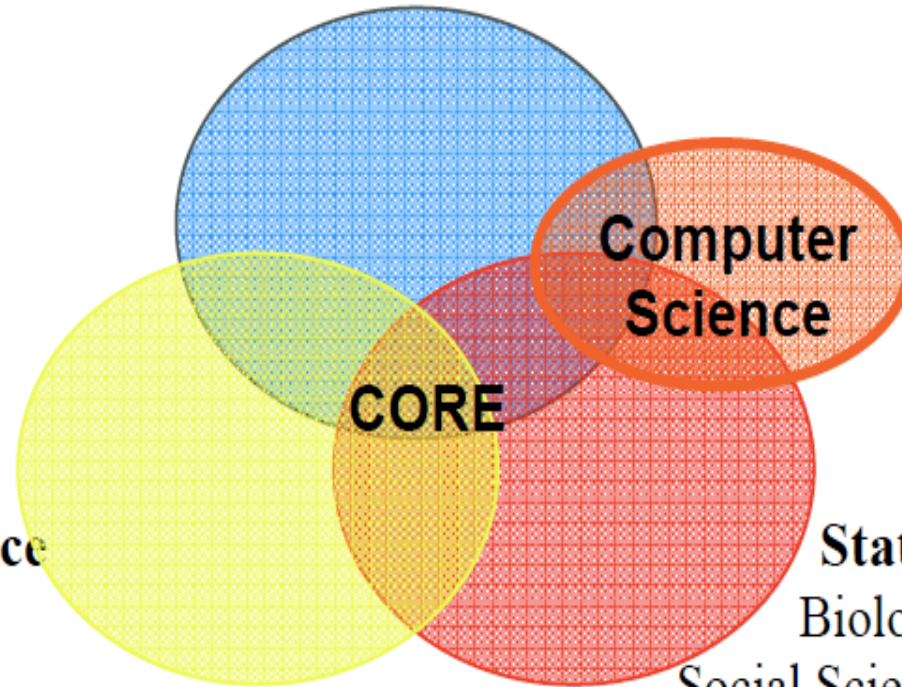


Nach Corraï 2008 [13]

e-IRG ETTF: domänenspezifische Kompetenzen! [14]

Numerical Models

Physics, Engineering, Earth Systems, Chemistry, Materials Science



Epistemology and Provenance

Arts, Languages, Humanities

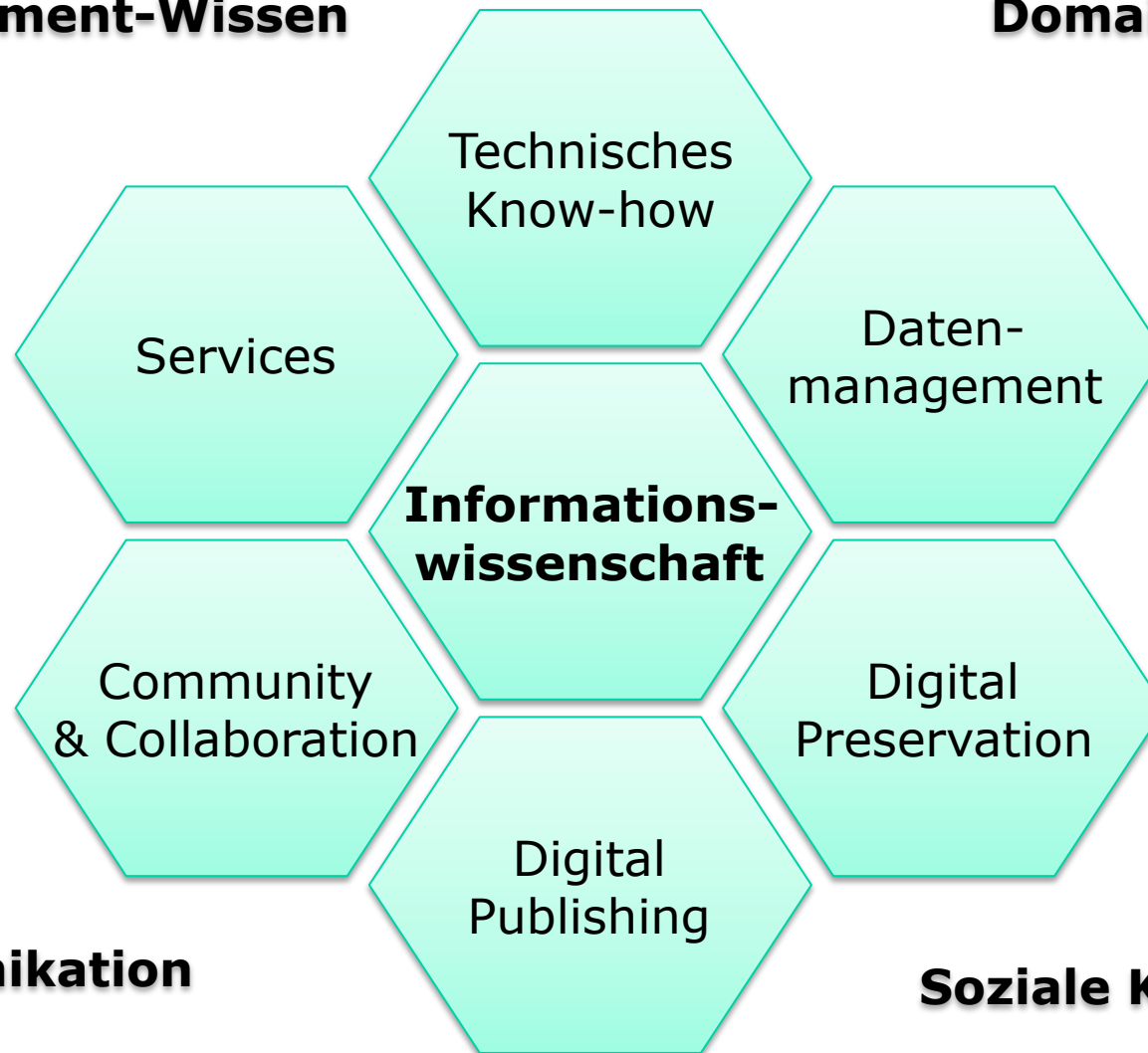
Statistical Models

Biology, Medicine,
Social Sciences, Economics

Informationswissenschaft als Kern

Management-Wissen

Domain-Wissen



Kommunikation

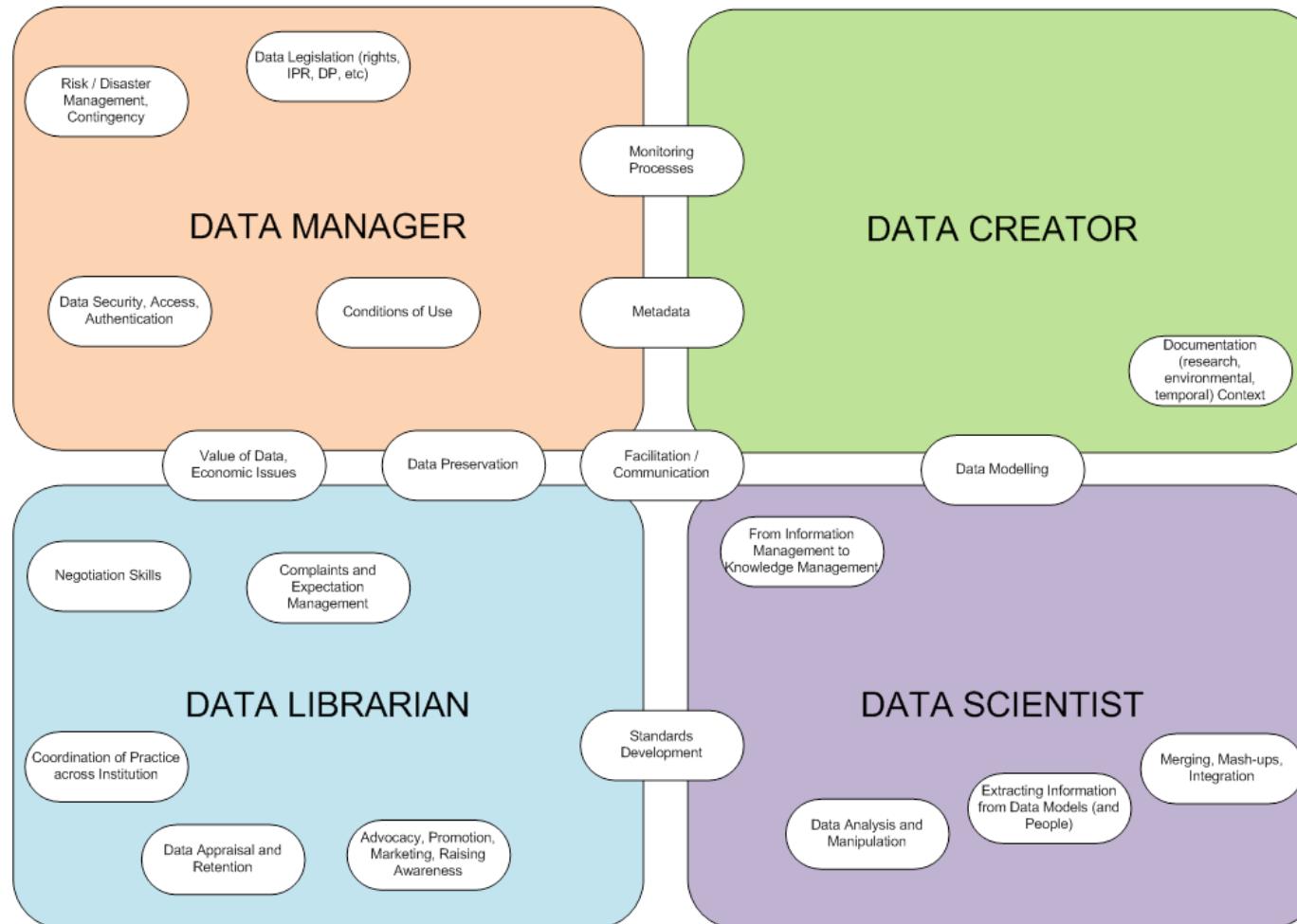
Soziale Kompetenz

Silke Grosamann

Rollen und ihre Kompetenzen [15]

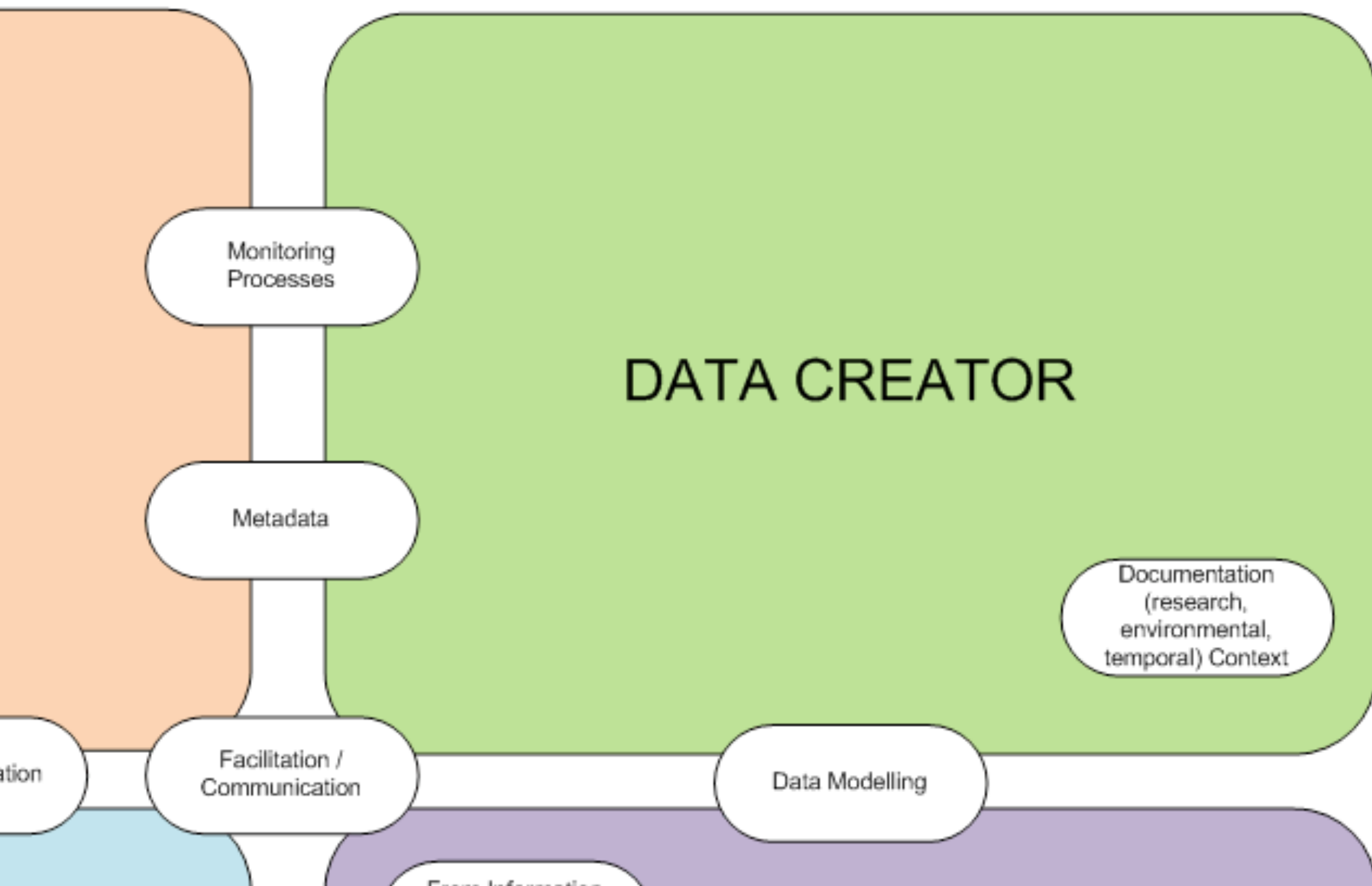
CORE SKILLS FOR DATA MANAGEMENT

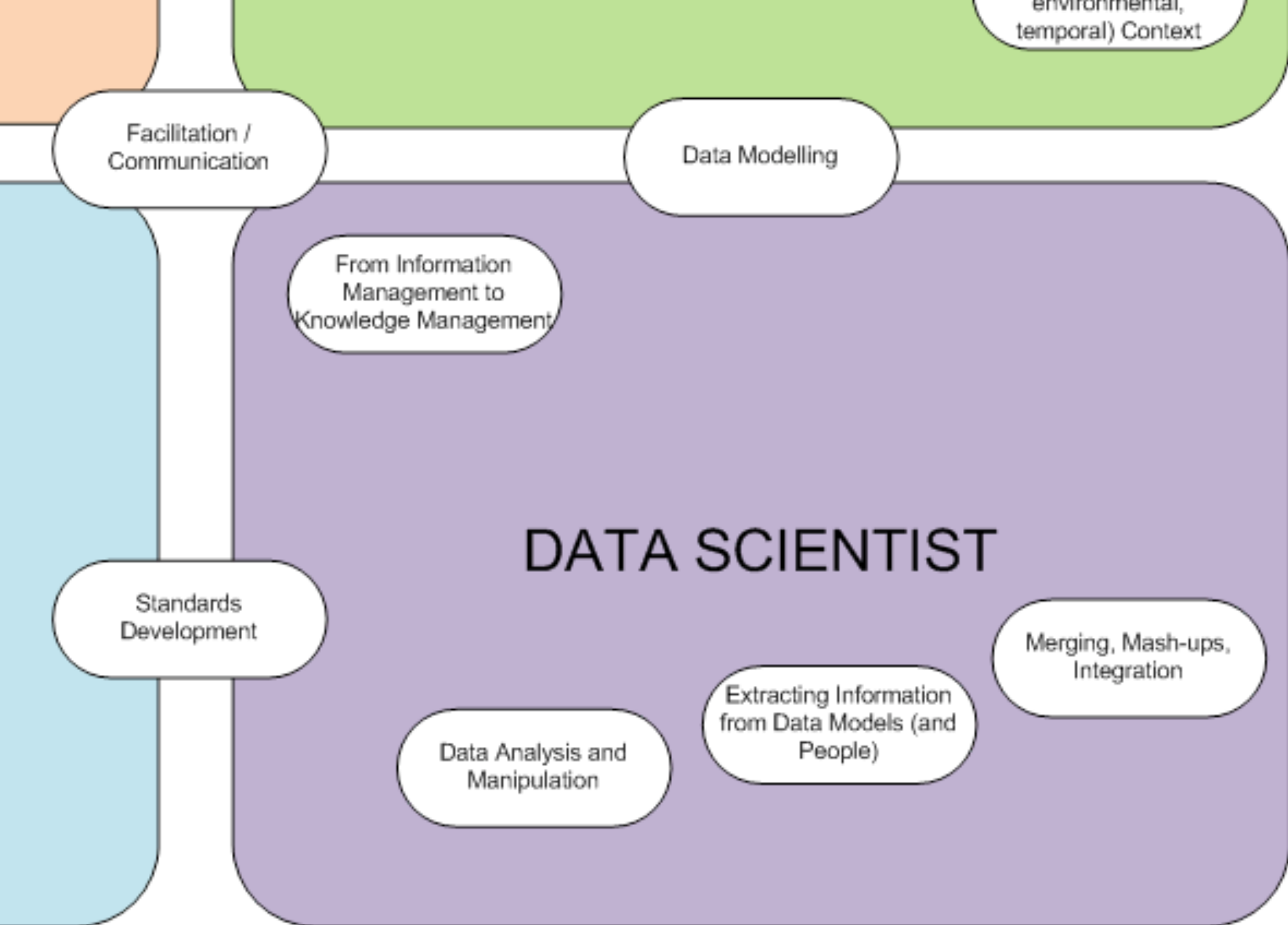
A follow-up from the second DCC Research Data Management Forum (November 2008)



TOOLS FOR DATA MANAGEMENT

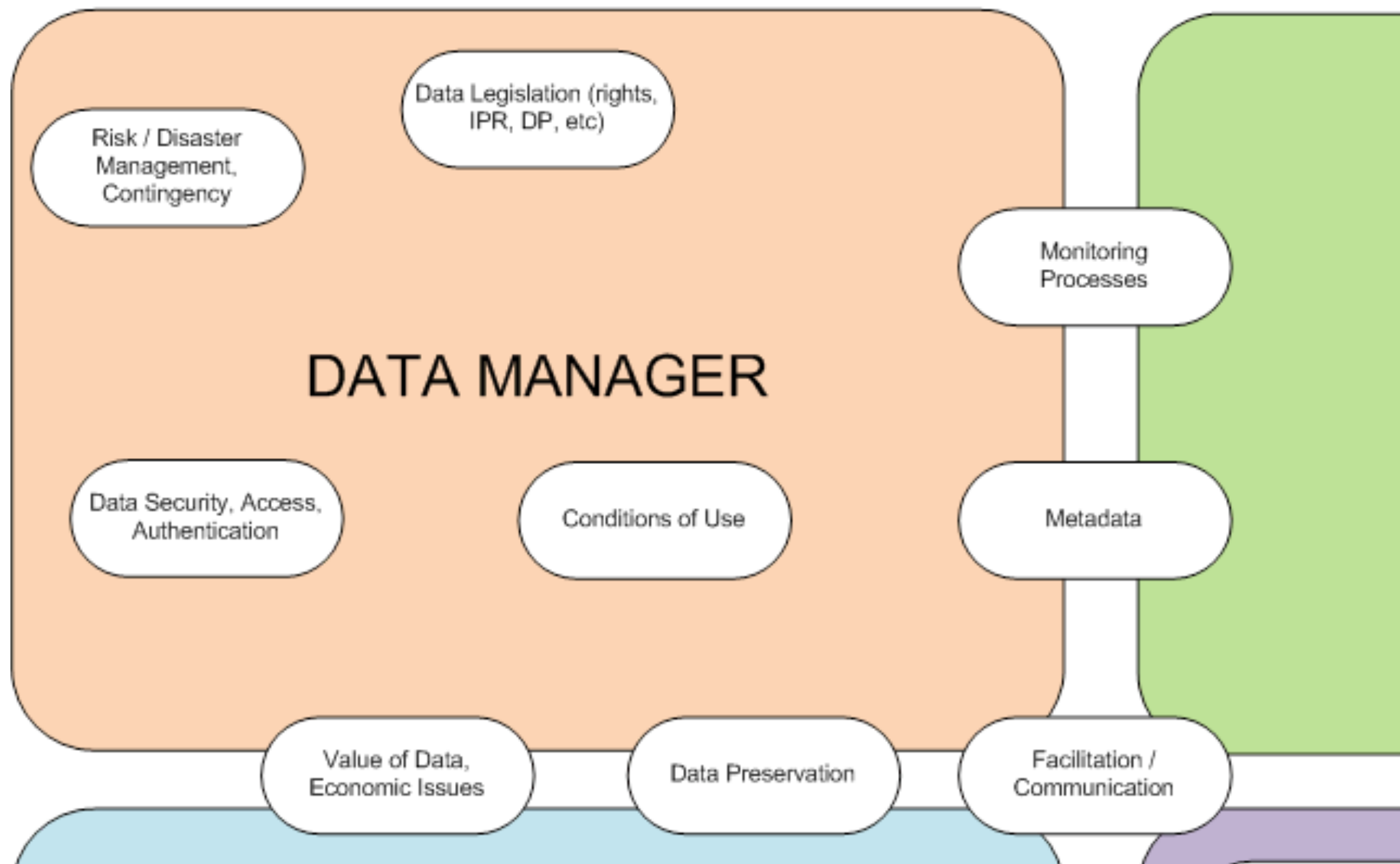
DCC Research Data Management Forum (November 2008)





CORE SKILLS FOR DATA MAN

A follow-up from the second DCC Research Data Management



Value of Data,
Economic Issues

Data Preservation

Facilitation /
Communication

Negotiation Skills

Complaints and
Expectation
Management

From
Man
Knowled

DATA LIBRARIAN

Coordination of Practice
across Institution

Standards
Development

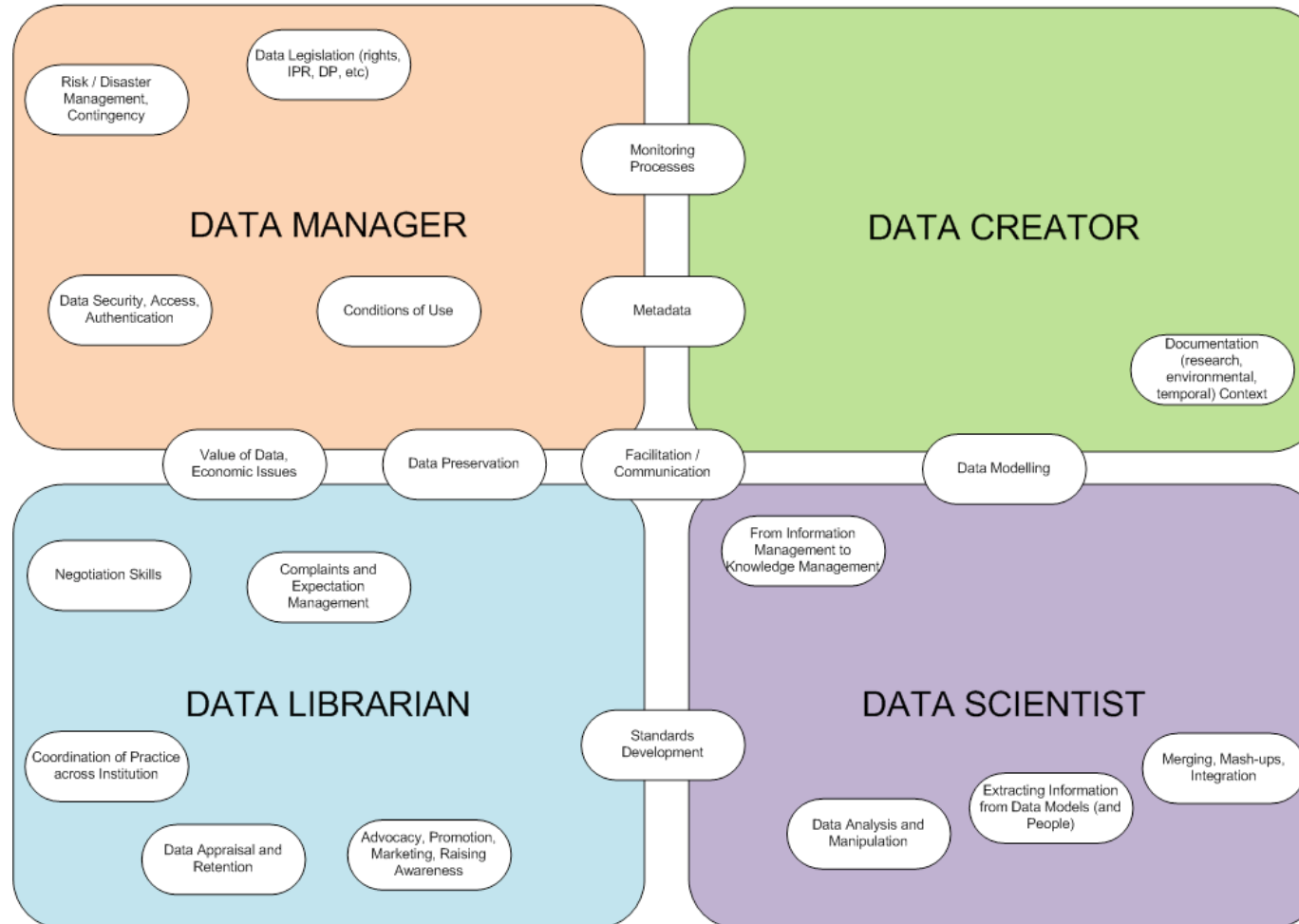
Data Appraisal and
Retention

Advocacy, Promotion,
Marketing, Raising
Awareness

Rollen und ihre Kompetenzen [15]

CORE SKILLS FOR DATA MANAGEMENT

A follow-up from the second DCC Research Data Management Forum (November 2008)



Stellenanzeige: „Data Librarian“ U Washington



Data Librarian

The Institute for Health Metrics and Evaluation (IHME) is a new organization at the University of Washington. Its mission is to monitor global health conditions and health systems, as well as to evaluate interventions, initiatives, and reforms. It also will build an audience of interested policymakers, researchers, donors, and other decision makers in health who are interested in learning more about how to apply, use, and derive better evidence in global health.

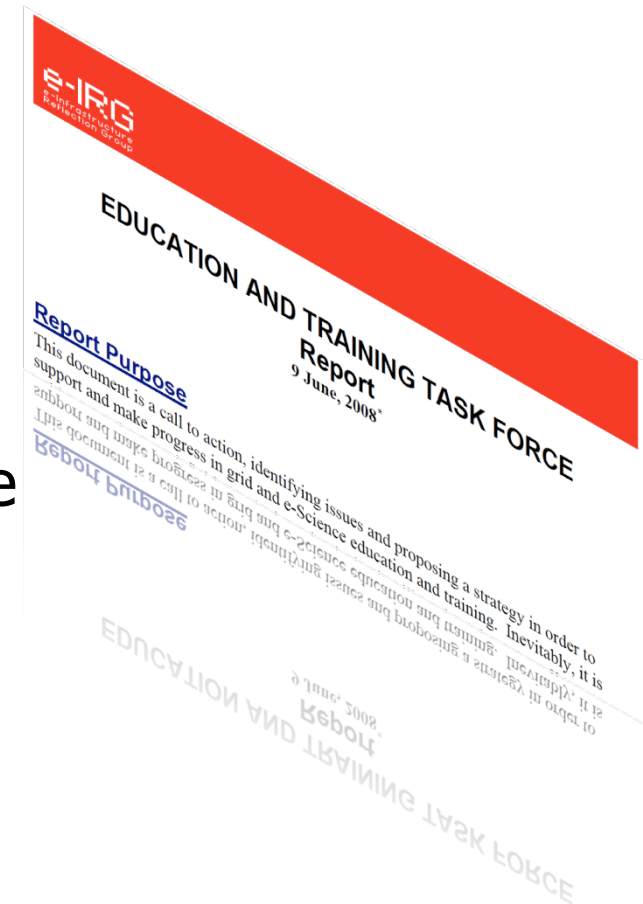
IHME has an outstanding opportunity for a Data Librarian. The Data Librarian will be the central resource within IHME on health data, both for internal availability and the global health data landscape. The Data Librarian will develop and improve processes and standards for data archiving, management of data needs, and facilitating efficient data retrieval, and play a central role in developing the IT infrastructure for data management. The Data Librarian will report to the Global Data Bank Director.

Stellenanzeige: „Data Librarian“ U Washington

- ✓ Bachelor's or higher degree in an appropriate field like Information or Library Science, or equivalent combination of education and experience
- ✓ 3 years experience in online content archiving, data documentation and online research
- ✓ Demonstrated facility with **analytic tasks** and ability to participate productively in **interdisciplinary** research teams
- ✓ **Interest** in global health research
- ✓ Expert user of **MS Excel** and Word required, familiarity with MS Access, database and cataloguing software preferred
- ✓ Strong **quantitative aptitude**
- ✓ High degree of attention to detail
- ✓ Demonstrated organizational skills, self-motivation, flexibility, good communication skills and the ability to work and thrive in a fast-paced, energetic, highly creative, entrepreneurial environment.

Modellcurriculum (OGF-ET CG, eIRG ETTF, ICEAGE) [17]

- Niveau 6 “Bachelor” (EQF):
- acht Hauptgebiete
- 40 Themen
- + domänenspezifische Werkzeuge
- Obligatorische Basisfächer:
 - Critical thinking
 - Data management
 - Use of models
 - Responsibility, legal, ethical & social issues
 - Trust: security, privacy and integrity.



Quellen

- [1] Nature: Community cleverness required. Nature 455 (2008), Nr. 7209. S. 1. DOI:10.1038/455001a
- [2] Helmholtz Open Access Projekt: Offener Helmholtz Workshop. Umgang mit Forschungsdaten. Bremen, 18.09.2008.
URL: <http://oa.helmholtz.de/index.php?id=239>
- [3] Study Group on Data Preservation and Long Term Analysis in High Energy Physics. URL: <http://www.dphep.org>
- [4] Helmholtz-Zentrum Berlin für Materialien und Energie: Archivierung von Forschungsdaten. URL: <http://tinyurl.com/l3w2vr>
- [5] Publication and Citation of Scientific Primary Data (STD-DOI). URL: <http://www.std-doi.de>
- [6] Bibliothek des Wissenschaftsparks Albert Einstein. URL: <http://bib.gfz-potsdam.de>
- [7] Digital Curation Centre: DCC Curation Lifecycle Model. URL: <http://www.dcc.ac.uk/docs/publications/DCCLifecycle.pdf>
- [8] National Science Board: Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century. Arlington, 2005. URL: <http://www.nsf.gov/pubs/2005/nsb0540/start.jsp>
- [9] Deutsche Initiative für Netzwerkinformation, Arbeitsgruppe Elektronisches Publizieren: Positionspapier Forschungsdaten. Göttingen, 2009. In Bearbeitung.
- [10] International Association for Social Science Information Service and Technology: IASSIST at a Glance. URL: <http://tinyurl.com/nc6h2o>
- [11] Swan, A.; Brown, S.: Skills, Role & Career Structure of Data Scientists & Curators. Assessment of Current Practice & Future Needs. Report to the JISC. Truro, 2008. URL: <http://tinyurl.com/lv83lw>
- [12] Jones, E.: E-Science Talking Points for ARL Deans and Directors. Washington, 2009. URL: <http://tinyurl.com/lv83lw>
- [13] Corral, S.: Research Data Management. Professional Education and Training Perspectives. Research Data Management Forum. Manchester, 2008. URL: <http://www.dcc.ac.uk/events/data-forum-2008-november/presentations/07.pdf>
- [14] Atkinson, M. et al. In: Computing Now (2009), March 2009 URL: <http://www2.computer.org/portal/web/computingnow/0309/education>
- [15] Donnelly, M.: RDMF2. Core Skills Diagram. DDC Research Data Management Forum: <http://data-forum.blogspot.com>
- [16] Job opening 24.4.2009: http://www.healthmetricsandevaluation.org/print/IHME_Data_Librarian.pdf
- [17] Atkinson, M. et.al.: Education and Training Task Force. Long Report. Lugano, 2008. Version 6, final, 09.06.2008.
URL: http://www.e-irg.eu/images/stories/publ/task_force_reports/ettf_long_report_final_july08.pdf

Vielen Dank für Ihre Aufmerksamkeit!

Roland Bertelmann
rab@gfz-potsdam.de
<http://www.gfz-potsdam.de>

Prof. Dr. Hans-Christoph Hobohm
hobohm@fh-potsdam.de
<http://www.fh-potsdam.de>

Heinz Pampel
pampel@gfz-potsdam.de
<http://oa.helmholtz.de>