GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

UNIVERSITÄTSMEDIZIN
GÖTTINGEN : UMG

The Development of National Road Maps for Digital Research Infrastructures in the International Context

KIT, Karlsruhe, 05.10.2016

Otto Rienhoff

Institut für Medizinische Informatik

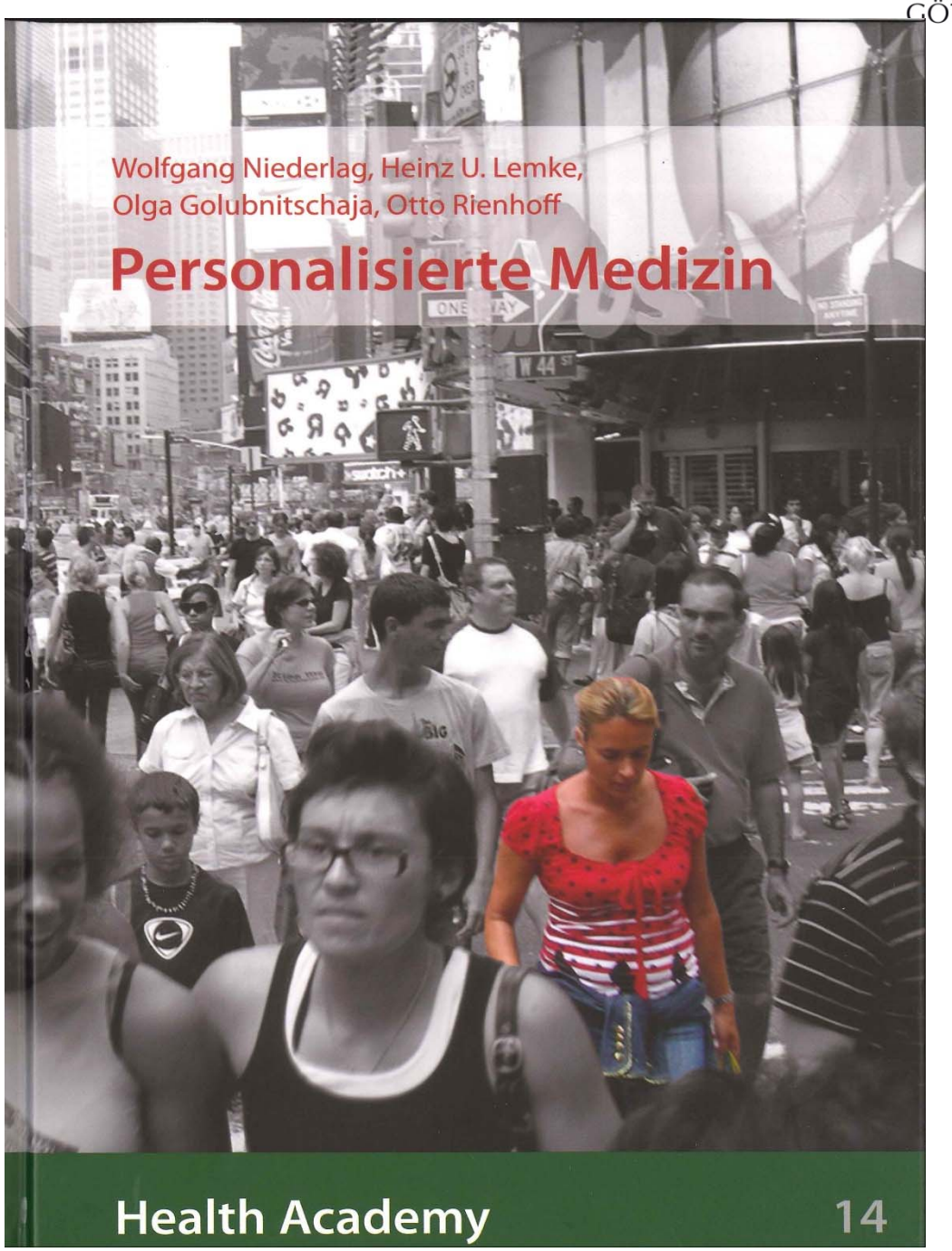
Abstract

Since the introduction of Hollerith machines there have been ideas for the evaluation of time wise massive data collections. However, the main developments started in the sixties based on main-frame computing. Already at that time many dreams were published which took decades to come through. As the delivery of promises and hopes was much less successful than expected many early starters and their institutions were shut down. Driving countries in this competitive game were not only the US and other industrialized countries but also highly funded national institution in many other countries. The complexity of the technology and the dynamics of software development forced most developments out of the competitive game. The US approaches became very dominant due to several "cultural" factors like massive funding resources and well funded disruptive approaches. However, the development of digital research infrastructures in different scientific disciplines has been driven by very specific national factors, of which available IT-power is only one of many im-pact factors. International research competition is more and more based on the management of digital research infrastructures. Easy answers and simple explanations of how to move ahead are not available.


Big Data und Forschungsinfrastruktur -
Perspektiven für die Medizin
Deutsche Gesellschaft für Medizinische Informatik,
Biometrie und Epidemiologie e.V.
59. Jahrestagung
gmds
Göttingen, 7. bis 10. September 2014
www.gmds2014.de

Our way from the past

- Development of IT-Infrastructures started in the last century – with a strong push during WW2.
- The development in different countries was often driven by nationally specific conditions or political targets (e.g. Germany, US, England, Sweden, Israel, Brazil, Cuba).
- Some countries experienced rather continuous developments – others ups and downs in waves (e.g. Denmark, Germany, Greece, South Africa).
- International organizations tend to try to manage the process by overarching funding schemes (e.g. EU) – which only works in internationally established research communities (e.g. astronomy).


Wolfgang Niederlag, Heinz U. Lemke,
Olga Golubnitschaja, Otto Rienhoff
Personalisierte Medizin
Health Academy
14


DER SPIEGEL
Nr. 50/5.12.2015
Deutschland €4,60
Total vermessen
Wir werden gläserne Patienten – und hoffen auf ewige Gesundheit
Geschlecht: weiblich
Alter: 29
Gewicht: 54 kg
Blutdruck: 120/70 mm Hg

The analyses of the German Council for Research Information Infrastructures

- Focus on „research data“ – addressing material collections as well digital data resources.
- Both are seen in closest connection and interdependence.
- It is perceived that both types of research infrastructures have to be promoted and substantially developed together.

- There is no simple recipe how to do it.
- Currently several working groups active (more infos at the end of the talk).

Types of `Research data`

Collections

- archive documents
- books
- bio specimen
- drilling probes
- archaeological artefacts
- …

Digital data

- raw data measurements
- metadata
- data utilization process documentation
- published data
- …

Database, 2016, 1–11 doi: 10.1093/database/baw125
Original article

Original article

The Global Genome Biodiversity Network (GGBN) Data Standard specification

G. Droege[1,*], K. Barker[2], O. Seberg[3], J. Coddington[2], E. Benson[4], W. G. Berendsohn[1], B. Bunk[5], C. Butler[2], E. M. Cawsey[6], J. Deck[7], M. Do¨ ring[8], P. Flemons[9], B. Gemeinholzer[10], A. Gu¨ ntsch[1], T. Hollowell[2], P. Kelbert[1], I. Kostadinov[11], R. Kottmann[12], R. T. Lawlor[13], C. Lyal[14], J. Mackenzie-Dodds[14], C. Meyer[2], D. Mulcahy[2], S. Y. Nussbeck[15], E'. O'Tuama[8], T. Orrell[2], G. Petersen[3], T. Robertson[8], C. So¨ hngen[5], J. Whitacre[2], J. Wieczorek[16], P. Yilmaz[12], H. Zetzsche[17], Y. Zhang[18] and X. Zhou[18]

[1]Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universit€at Berlin, Ko¨ nigin-Luise-Str. 6-8, Berlin 14195, Germany, [2]National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA, [3]Natural History Museum of Denmark, University of Copenhagen, Sølvgade 83, opg. S, Copenhagen DK-1307, Denmark, [4]Damar Research Scientists, Damar, Drum Road, Cuparmuir, Fife KY15 5RJ, UK, [5]Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Inhoffenstr. 7B, Braunschweig 38124, Germany, [6]Australian National Wildlife Collection, CSIRO National Research Collections Australia, Canberra, Australia, [7]Berkeley Natural History Museums, University of California at Berkeley, Berkeley, CA 94720, USA, [8]Global Biodiversity Information Facility Secretariat, Universitetsparken 15, Copenhagen DK-2100, Denmark, [9]Australian Museum, Sydney 2010, NSW, Australia, [10]Systematic Botany, Justus Liebig University, Giessen 35392, Germany, [11]Department of Life Sciences & Chemistry, Jacobs University Bremen gGmbH, Campus Ring 1, Bremen 28759, Germany, [12]Microbial Genomics and Bioinformatics Research Group, Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, Bremen 28359, Germany, [13]ARC-Net Applied Research on Cancer Centre, Department of Pathology and Diagnostics, University of Verona, Verona 37134, Italy, [14]Natural History Museum, Cromwell Road, London SW7 5BD, UK, [15]Department of Medical Informatics and UMG Biobank, University Medical Center Go¨ ttingen, Robert-Koch-Str. 40, Go¨ ttingen 37075, Germany, [16]Museum of Vertebrate Zoology, University of California at Berkeley, Berkeley, CA 94720, USA, [17]Julius Kuehn-Institute (JKI), Federal Research Centre for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Erwin-Baur-Str. 27, Quedlinburg 06484, Germany and, [18]China National GeneBank, BGI-Shenzhen, Shenzhen, Guangdong 518083, China

Disruption as principle of advance ?

- Simple messages pay off in the political discussion.
- The belief in the advantages of „disruption“ follows the idea of transplanting success-stories in other countries into the own environment (e.g. Silicon Valley to Berlin).
- Historical examples support such argumentations (e.g. US Liberty Sip Production).
- However, they need changes in the political set-up to work (e.g. Singapore).
- Societal areas which are highly legally defined will be difficult to disrupt (e.g. health care in Germany).

Résumé: still a long way to go
or: the complexity of social change

- Strategic planning of research data-infrastructures needs societal developments to move consequently ahead.
- Leapfrogging by disruption may only help if tipping-points are reached.
- Thus, it is extremely difficult to foresee the dynamics of the future development.
- However, more data without reliable scientific results may lead to negative political reactions.
- Strategic planning has to manage this kind of development and focus on social aspects rather than just extending IT-hardware resources.

The one and only key: social/legal development by education and training

- Countries have to identify their typical national route ahead.
- There will be massive changes for social groups on that route. Leading to fights for resources, social recognition, and power.
- Countries will manage these distribution fights differently.
- These effects are the stronger the more the sector/discipline is entangled in national traditions.
- They will have a strong impact on big data IT-infrastructures.
- Driving factors in the societal process will be national interests (e.g. Iceland, Israel, Estonia) in respect to the management of research data and research material collections.

Nationaler IT Gipfel

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

Rat für Informations Infrastrukturen

Position Paper 2016

ENHANCING RESEARCH DATA MANAGEMENT: PERFORMANCE THROUGH DIVERSITY

Recommendations regarding structures, processes, and financing for research data management in Germany

CURRENT SITUATION:
Research Data Management (RDM) in Germany

- Dramatic increase in the quantity/heterogeneity of research data in Germany
- Numerous informative statements regarding RDM, but there is still an "implementation deficit"
- Substantial financial and personnel requirements
- Numerous hurdles inhibit the transformation
  - Heterogeneous funding landscape, a lack of coordination, orientation towards technology instead of processes, cultural diversity among the disciplines, unclear quality assurance and reputation mechanisms, as well as uncertainty of the stakeholders in terms of strategic investments

RECOMMENDATIONS

… for policymakers and science

- Changing funding policy – sustainability through the orderly transition of project-based initiatives into suitable sponsoring organisation
- Ensuring efficiency and coordination based on a (distributed) national research data infrastructure (NFDI)
- Roadmaps as a tool for coordination and financial planning of the NFDI
- Linking the NFDI with monitoring and quality assurance systems
- Promoting various aspects of a "research data culture"
- Progressing the development of expertise/personnel at all levels
- Promoting international networking (in the EU and globally)

RECOMMENDATIONS

… for scientific organisations

- Promoting evaluation concepts for RD infrastructures
- Developing guidelines for long-term archiving
- Disseminating standards and quality criteria (and in doing so: working towards fulfilling the needs of users)
- Establishing binding rules of conduct
- Designing attractive training and education programmes for new occupations and career paths
- Establishing dedicated data publication services in science
- Providing guidance during the establishment of the NFDI

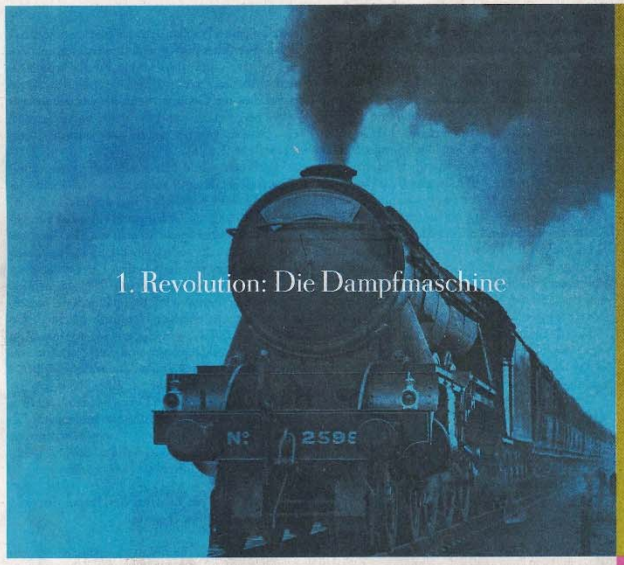
RECOMMENDATIONS

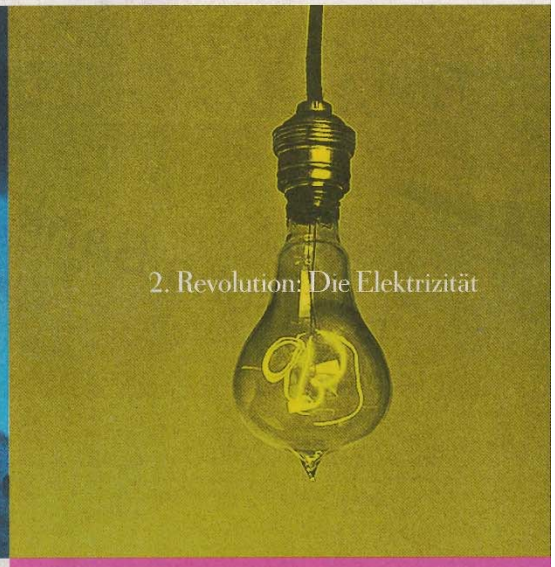
… for researchers and their scientific societies + communities

- Developing RD standards appropriate to the discipline while considering establishing connections to existing services (i.e. ensuring interoperability)
- Planning how to handle data in good time, triggering a scientific discourse regarding archiving periods
- Coordinating the interests in the triangle formed by data users, data producers, and infrastructure centres
- Recognising digitality as a field of expertise of researchers + management, actively initiating a discussion regarding critical points
- Leading a discourse on responsibility for data and data security
- Contributing locally from the bottom up to the establishment of an NFDI

29. SEPTEMBER 2016 DIE ZEIT N° 41

WIRTSCHAFT





Ist der Computer überschätzt?

Ökonomen kritisieren das Internet: Viel Hype, wenig Nettogewinn! THOMAS FISCHERMANN glaubt, dass sie die aktuellen Revolutionen schlecht verstehen