

Research data in library context

Dr Jan Bräse, Head of R&D

10th Anniversary IGeLU Conference in Budapest, Hungary. September 2nd



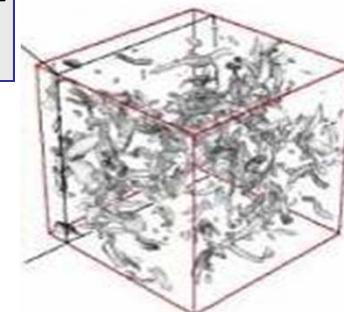
GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Science Paradigms

- Thousand years ago:
science was **empirical**
describing natural phenomena
- Last few hundred years:
theoretical branch
using models, generalizations
- Last few decades:
a **computational** branch
simulating complex phenomena
- Today:
data exploration (eScience)
unify theory, experiment, and simulation



$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K \frac{c^2}{a^2}$$



• Jim Gray, eScience Group, Microsoft Research



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

Consequences for Libraries

- Scientific Information is more than a journal article or a book
- Libraries should open their catalogues to any kind of information
- The catalogue of the future is NOT ONLY a window to the library's holding, but
- A portal in a net of trusted providers of scientific content



We do not have it

BUT

We know where you can find

And here is the link to it!



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

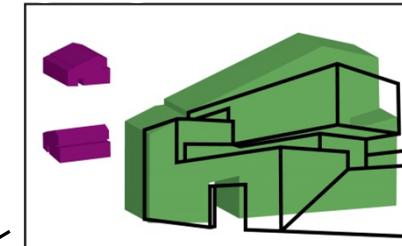
SUB

```
-<state>
<language>de<
<position/>
<length/>4<
<name>peter</
-<favourites>
-<favourite>
```

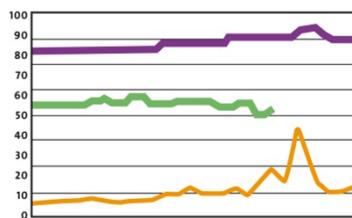
Software



Scientific Films



3D Objects



Research Data



TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	
TexteTexte	

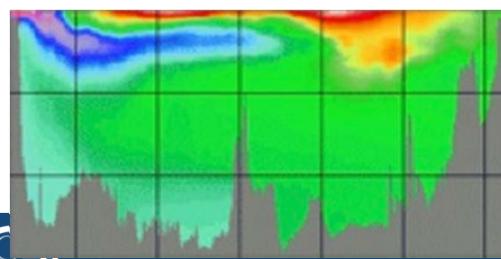
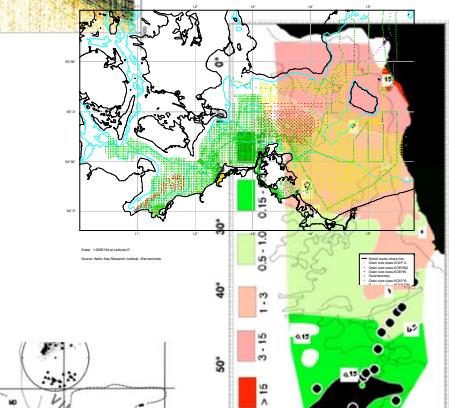
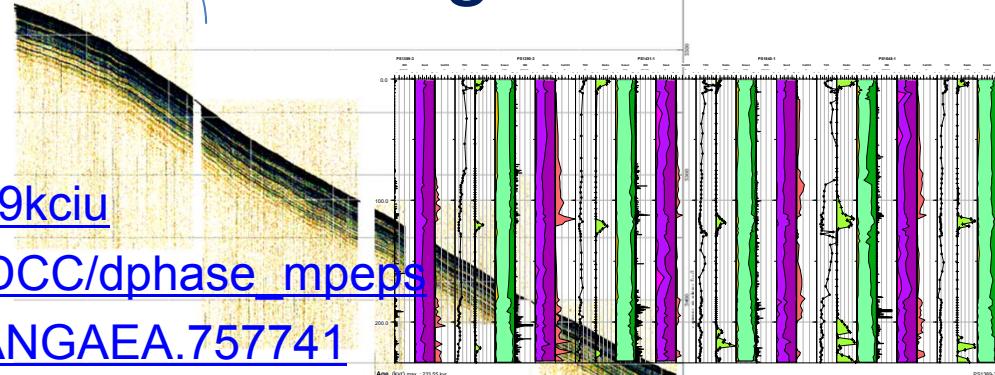
Why is this a role for libraries?

- Libraries have a history in bringing information to the public
- Libraries have a tendency to be persistent
 - A project will be forgotten in 40 years, the library will very likely still exist then
- Library are very trustworthy organisations



What type of data are we talking about?

- Earth quake events => [doi:10.1594/GFZ.GEOFON.gfz2009kciu](https://doi.org/10.1594/GFZ.GEOFON.gfz2009kciu)
- Climate models => [doi:10.1594/WDCC/dphase_mpeps](https://doi.org/10.1594/WDCC/dphase_mpeps)
- Sea bed photos => [doi:10.1594/PANGAEA.757741](https://doi.org/10.1594/PANGAEA.757741)
- Distributes samples => [doi:10.1594/PANGAEA.51749](https://doi.org/10.1594/PANGAEA.51749)
- Medical case studies => [doi:10.1594/eaacinet2007/CR/5-270407](https://doi.org/10.1594/eaacinet2007/CR/5-270407)
- Computational model => [doi:10.4225/02/4E9F69C011BC8](https://doi.org/10.4225/02/4E9F69C011BC8)
- Audio record => [doi:10.1594/PANGAEA.339110](https://doi.org/10.1594/PANGAEA.339110)
- Grey Literature => [doi:10.2314/GBV:489185967](https://doi.org/10.2314/GBV:489185967)
- Videos => [doi:10.3207/2959859860](https://doi.org/10.3207/2959859860)



Examples



German National Library of Science and Technology

The screenshot shows a web page from the GetInfo platform, which is part of the German National Library of Science and Technology (TIB). The left sidebar contains a vertical menu with icons and links: Home, Bestellung ohne Recherche, Suchen und Bestellen, Informationsdienste, Konditionen, Merkliste, MyGetInfo, Registrieren, Über GetInfo, Aktuelles, Newsletter, Kundenservice, AGB, Impressum, and Datenschutz.

The main content area features the GetInfo logo and the TIB Technische Informationsbibliothek logo. The page title is "Detailansicht". The main content is about the "SAFOD borehole trajectory data in absolute coordinates (UTM) and in coordinates relative to drilling platform". It includes a thumbnail image labeled "DATA" and a "Zu den Daten" button. Below this, there is a section titled "Dokumentinformationen" with details such as Format / Umfang: 11331 DataPoints, DOI: 10.1594/GFZ.SDDB.1081, Zitierlink: <http://dx.doi.org/10.1594/GFZ.SDDB.1081>, Dokumenttyp: Forschungsdaten, Dokumentformat: Elektronische Ressource, Sprache: Englisch, and Schlagwörter: San Andreas Fault-Zone Observatory at Depth, Land based, Azimuth, Dog Leg Severity, Easting (Departure), Easting UTM NAD27, Inclination, Northing (Latitude), Northing UTM NAD27, True Vertical Depth, Vertical Section.

The "Abstract" section describes the SAFOD project's motivation and target: "SAFOD is motivated by the need to answer fundamental questions about the physical and chemical processes controlling faulting and earthquake generation within a major plate-bounding fault. SAFOD will drill and instrument an inclined borehole across the San Andreas Fault Zone to a depth of 3.2 km, targeting a repeating microearthquake source. The drill site is located west of the vertical San Andreas Fault on a segment of the fault that moves through a combination of aseismic creep and repeating microearthquakes. It lies at the extreme northern end of the rupture zone of the 1966, Magnitude 6 Parkfield earthquake, the most recent in a series of events that have ruptured the fault five times since 1857. The Parkfield region is the most comprehensively instrumented section of a fault anywhere in the world, and has been the focus of intensive study for the past two decades. This data set contains SAFOD borehole trajectory data."

The "Ähnliche Dokumente" section lists a related document: "MILP-based trajectory generation in Relative Velocity Coordinates" by Di Zu, Jianda Han, and Dalong Tan, published in TIBscholar in 2007.



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Jan Brase – Keynote IgeLu, Budapest

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

GFZ
Helmholtz Centre
POTS DAM

Impressum

HELMHOLTZ CENTRE POTS DAM
**GFZ GERMAN RESEARCH CENTRE
FOR GEOSCIENCES**

 Dataset SAFOD borehole trajectory data in absolute coordinates (UTM) and in coordinates relative to drilling platform  Released

Cite as:
SAFOD (2007): SAFOD borehole trajectory data in absolute coordinates (UTM) and in coordinates relative to drilling platform. Deutsches GeoForschungsZentrum GFZ. <http://dx.doi.org/10.1594/GFZ.SDDB.1081>

Data Files 
[data.csv](#) 117724 Bytes
License: cc-by

Related Work

Find More Research Data
<http://bib.telegrafenberg.de/finden/datenbanken/forschungsdaten/>

Abstract 
SAFOD is motivated by the need to answer fundamental questions about the physical and chemical processes controlling faulting and earthquake generation within a major plate-boundary fault. SAFOD will drill and instrument an inclined borehole across the San Andreas Fault Zone to a depth of 3.2 km, targeting a repeating microearthquake source. The drill site is located west of the vertical San Andreas Fault on a segment of the fault that moves through a combination of aseismic creep and repeating microearthquakes. It lies at the extreme northern end of the rupture zone of the 1966, Magnitude 6 Parkfield earthquake, the most recent in a series of events that have ruptured the fault five times since 1857. The Parkfield region is the most comprehensively instrumented section of a fault anywhere in the world, and has been the focus of intensive study for the past two decades. This data set contains SAFOD borehole trajectory data.

Keywords
Solid Earth, Deep Drilling, Geology, Azimuth, Dog Leg Severity, Easting (Departure), Easting UTM NAD27, Inclination, Land based, Northing (Latitude), Northing UTM NAD27, San Andreas Fault-Zone Observatory at Depth, True Vertical Depth, Vertical Section

GCMD Science Keywords
EARTH SCIENCE > Solid Earth > Deep Drilling > Well-Logging

More Metadata
iso19115: [view inline](#) / [download xml](#)
datacite: [view inline](#) / [download xml](#)
dif: [view inline](#) / [download xml](#)
esidoc: [view inline](#) / [download xml](#)

Location
Northern Latitude: 35.9713 Southern Latitude: 35.9712
Eastern Longitude: -120.5512 Western Longitude: -120.5513

University Library of Bielefeld

The screenshot shows the homepage of the University Library of Bielefeld's digital repository, PUB. At the top, there is a green header bar with the text "PUB" and "Universität | International | Informationen für" followed by a search icon. To the right of the search icon is a link "Uni von A-Z". Below the header, there is a large image of a person's hands typing on a laptop keyboard. To the right of the image is a smaller image of a modern university building.

On the left side, there is a user profile icon. Next to it, the name "Alexander Schwegmann" is displayed, followed by the text "Former Bielefeld University Researcher". To the right of the profile, there is a blue button labeled "PUB Login". Below the login button is a link "Your Marked Publications 0".

Below the profile section, there are two tabs: "Publications" (which is selected) and "Data Publications". To the right of these tabs is a link "Mark all".

The main content area displays a list of publications:

- [3] 2014 | Journal Article | PUB-ID: 2694962 [Mark](#)
Depth information in natural environments derived from optic flow by insect motion detection system: a model analysis
Schwegmann A, Lindemann J, Egelhaaf M (2014)
Frontiers in Computational Neuroscience 8: 83.
[PUB](#) | [PDF](#) | [DOI](#) | [PubMed](#) | [Europe PMC](#)
- [2] 2014 | Journal Article | PUB-ID: 2701544 [Mark](#)
Temporal Statistics of Natural Image Sequences Generated by Movements with Insect Flight Characteristics
Schwegmann A, Lindemann J, Egelhaaf M (2014)
PLoS ONE 9(10): e110386.
[PUB](#) | [PDF](#) | [DOI](#) | [PubMed](#) | [Europe PMC](#)
- [1] 2014 | Dissertation | PUB-ID: 2684681 [Mark](#)
Consequences of self-motion on image statistics and depth perception in natural cluttered environments : a model analysis
Schwegmann A (2014)
Bielefeld
[PUB](#)

On the right side of the page, there are several filter and display options:

- Search Publications: A search input field and a "Go!" button.
- Filter Publications: Options for "Publication Type" (2), "Uni-Bi Co-authors" (2), and "Publishing Year" (1).
- Display & Export Publications: Options for "Citation Style: default" and "Sorting".
- RSS Feed: A link to an RSS feed icon.

University Library of Bielefeld

The screenshot shows the University Library of Bielefeld's digital platform, PUB. At the top, there is a green header bar with the logo "PUB" and links for "Universität | International | Informationen für" and "Uni von A-Z". Below the header, a banner features a person typing on a laptop and a view of a modern university building.

A user profile for "Alexander Schwegmann" is displayed, identifying him as a "Former Bielefeld University Researcher". He has a grey placeholder profile picture. To the right of his name is a "PUB Login" button. Below the profile, there is a link to "Your Marked Publications" with a count of 0.

The main content area is divided into two sections: "Publications" and "Data Publications". The "Data Publications" tab is currently selected, indicated by a blue border. It lists three publications:

- [3] 2014 | Research Data | PUB-ID: 2689637
Panoramic high dynamic range images in diverse environments
Meyer HG, Schwegmann A, Lindemann JP, Egelhaaf M (2014)
Bielefeld University. doi:10.4119/unibi/2689637
PUB | DOI | Files available
- [2] 2014 | Research Data | PUB-ID: 2693180
Matlab .m scripts for processing panoramic hdr images
Schwegmann A, Lindemann JP, Egelhaaf M (2014)
Bielefeld University. doi:10.4119/unibi/2693180
PUB | DOI | Files available
- [1] 2014 | Research Data | PUB-ID: 2689483
Translational sequences of panoramic high dynamic range images in natural environments
Schwegmann A, Lindemann JP, Egelhaaf M (2014)
Bielefeld University. doi:10.4119/unibi/2689483
PUB | DOI | Files available

To the right of the publications, there are several filter and display options:

- Search Data Publications**: Includes a search input field and a "Go!" button.
- Filter Data Publications**: Includes dropdown menus for "Uni-Bi Co-authors (3)" and "Publishing Year (1)".
- Display & Export Publications**: Includes dropdown menus for "Citation Style: default" and "Sorting", and a link to an "RSS Feed".

University Library of Bielefeld

The screenshot shows a publication page on the PUB platform. At the top, there's a green header bar with the text "Universität Bielefeld" and "PUB – Publications at Bielefeld University". Below this is a large yellow banner with the word "PUB" in white. The main content area features a photograph of a person's hands typing on a laptop keyboard. To the right of the image is a smaller photo of a modern glass-enclosed building. Below the images, there are navigation links: "University", "International", "Information for", and "University from A-Z".

Panoramic high dynamic range images in diverse environments

Meyer HG, Schwegmann A, Lindemann JP, Egelhaaf M (2014)
Bielefeld University. doi:[10.4119/unibi/2689637](https://doi.org/10.4119/unibi/2689637)

Download [readme.pdf](#) 80.14 KB
[preview_images.rar](#) 69.92 MB
[data_001-100.rar](#) 214.71 MB 

DOI [10.4119/unibi/2689637](https://doi.org/10.4119/unibi/2689637)

Research Data *Mark*

Details **File Details**

Authors Meyer, Hanno Gerd^{UniBi}; Schwegmann, Alexander^{UniBi}; Lindemann, Jens Peter^{UniBi}; Egelhaaf, Martin^{UniBi}

Department Fakultät für Biologie
Center of Excellence - Cognitive Interaction Technology CITEC
Neurobiologie

Abstract

This database contains 421 panoramic high dynamic range images recorded in diverse environments. The Images are panoramas in full 360° in azimuth and between -58° below and 47° above the horizon in elevation. We used a spectral filter to limit the camera's spectral sensitivity to wavelengths in the range of 480-560 nm (green). This filtering mimics the spectral sensitivity of photoreceptors R1-R6 of the fly that provide the input of the motion vision system. As a consequence, the mapping of colors to gray values in these images is similar to the green color channel in RGB images. The raw images have a resolution of approximately 1 Megapixel (928x928) and 12-bit. The images have high dynamic range covering the entire brightness range encountered in natural environments (excluding the solar disc) After linearization the resulting image values had a dynamic range of 1:23,900 covering 3,955 intensity steps. Note, however, that the pixel brightness values cannot be recalculated to a SI unit like candela, though the values are proportional to luminance in the green spectral range. For more technical details about the recording of the image sequences see Meyer et al. (2014) In addition to the raw camera images,

Export

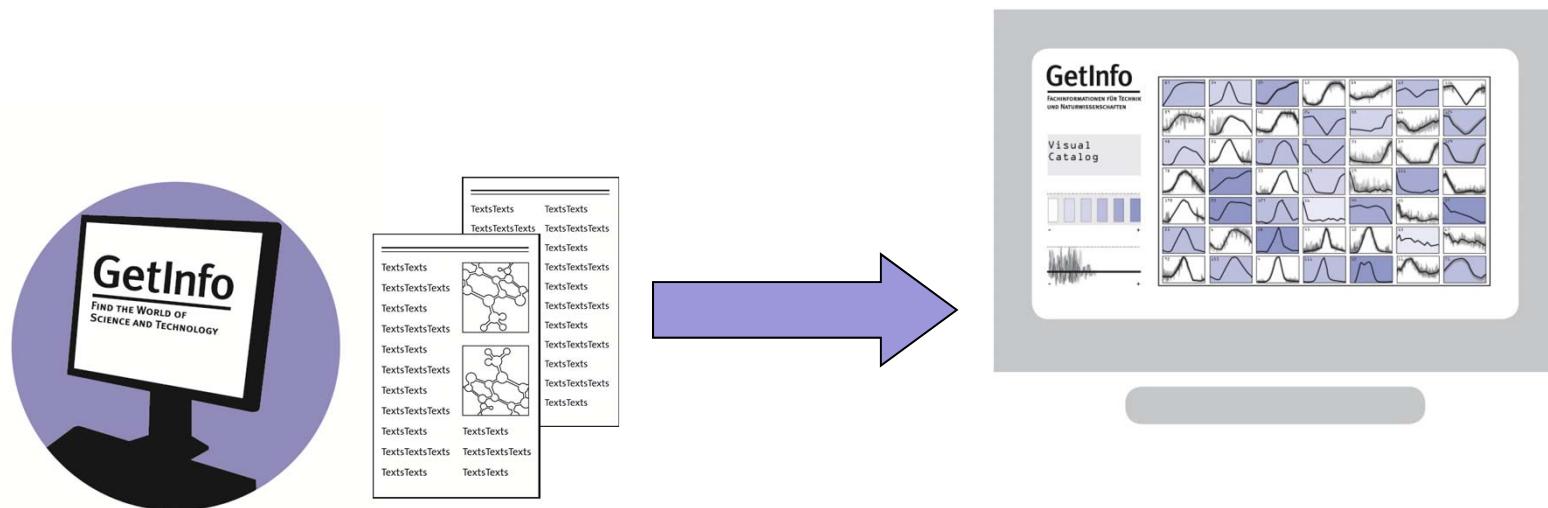
 [Marked Publication](#)
 [Open Data PUB](#)

Search this title in

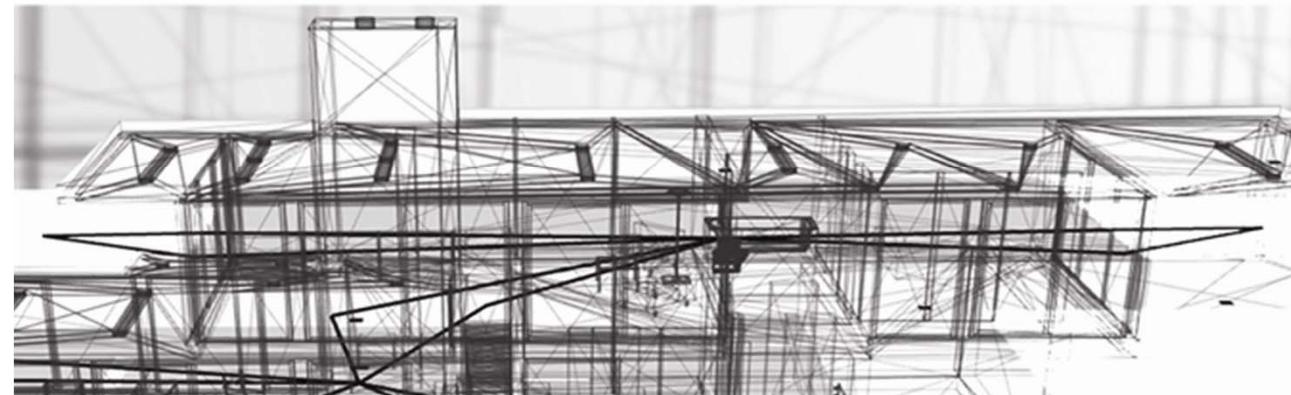
 [Google Scholar](#)
 [BASE](#)
 [Google](#)

Research data in Library context- Tools

- Make more scientific and technical content searchable
 - Develop tools to address each type of scientific and technical Information
- ➔ Present systems are designed to handle text formats



PROBADO | 3D



TECHNISCHE
UNIVERSITÄT
DARMSTADT

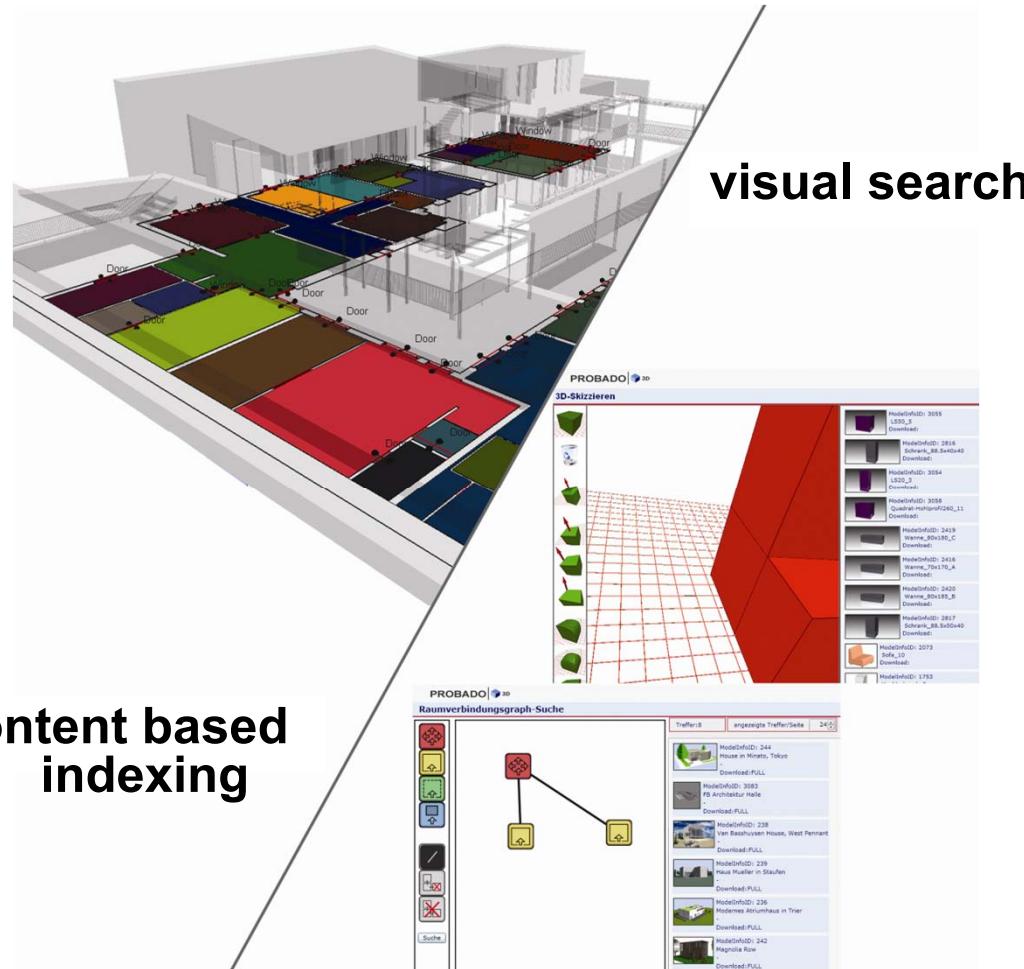


GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

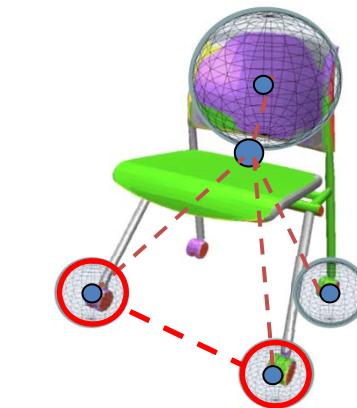
NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

15
SUB

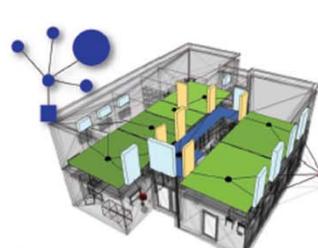
Indexing and search



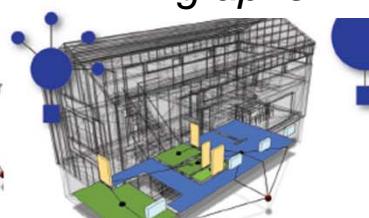
> content based indexing



*segmentation with
form-primitives*



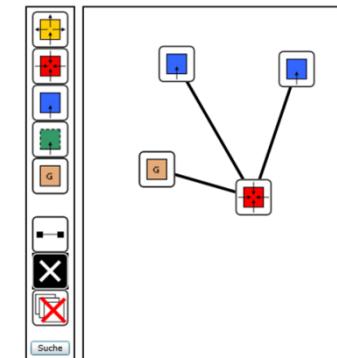
(a) corridor type



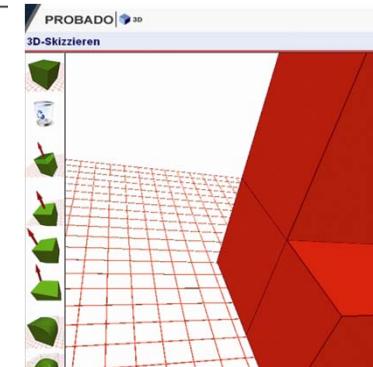
(b) distributor type

*classification of floor types
→ machine learning*

> visual search



attributed graph



3D sketch

result visualization

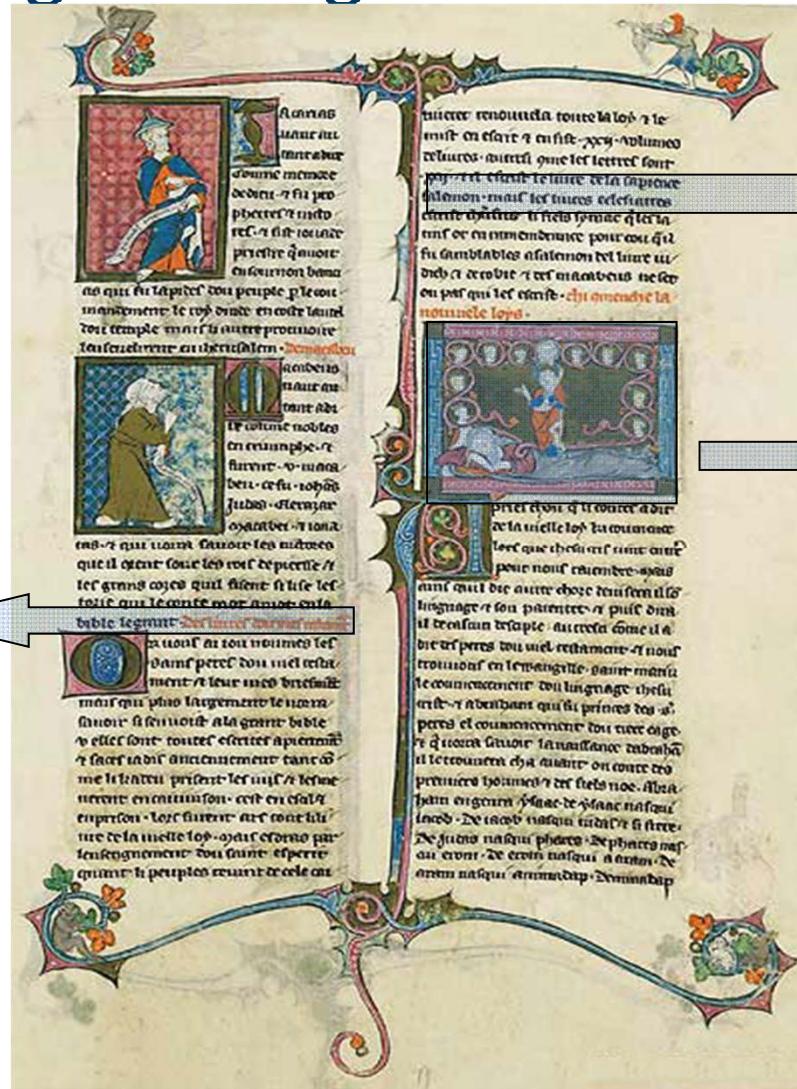
NIEDERLÄNDISCHE STADT UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

SUB Göttingen - Digital Editions



Historical
People

Historical
Places

Links to other
Digital objects



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

Chemical search

The screenshot shows the Chem.de search interface. At the top, there's a navigation bar with links for Home, Kontakt, Disclaimer, Impressum, and Hilfe. Below the navigation is a banner featuring four images: a scientist at a microscope, laboratory glassware, a magnifying glass over a computer screen, and a close-up of a computer keyboard.

The main search area is titled "Struktursuche" (Structure Search). On the left, there's a sidebar with "Suchoptionen" (Search options) including "Maximale Suchzeit: 10 Sek", "Maximale Trefferanzahl: 20", "Suchmethode: substructure", "Similarity-Grenzwert: 0.1", and a "Suche" button. The search results are displayed in a grid format. One result is highlighted:

methyl-4*H*-pyrrolo[2,3-*c*]isoquinoline: two new unnatural D-ring stripped
J. W.; Lemière, Guy L. F.;
Mikhail V.; Utepova, Irina A.;
bio[3,4-*b*][1,3,4]thiadiazines using chitosan as heterogeneous catalyst
nitrones prepared from D-isoascorbic acid with methyl acrylate
onyl acrylonitrile in the synthesis of novel 3-heteroarylindoles:
von: El-Nezhawy, Ahmed O. H.; Radwan, Mohamed A. A.; Ragab, Eman A.; Shaaban, Mohamed R.;
veröffentlicht in: Archive for Organic Chemistry (2009)

On the right side of the search results, there are three panels: "Chemische Entitäten" (Chemical Entities) showing small chemical structures; "Reaktionen" (Reactions) listing reaction types like Michael addition, ribosylations, formylation, deoxylations, and dehydrogenation reactions; and "Autoren" (Authors) listing authors with their publication counts.

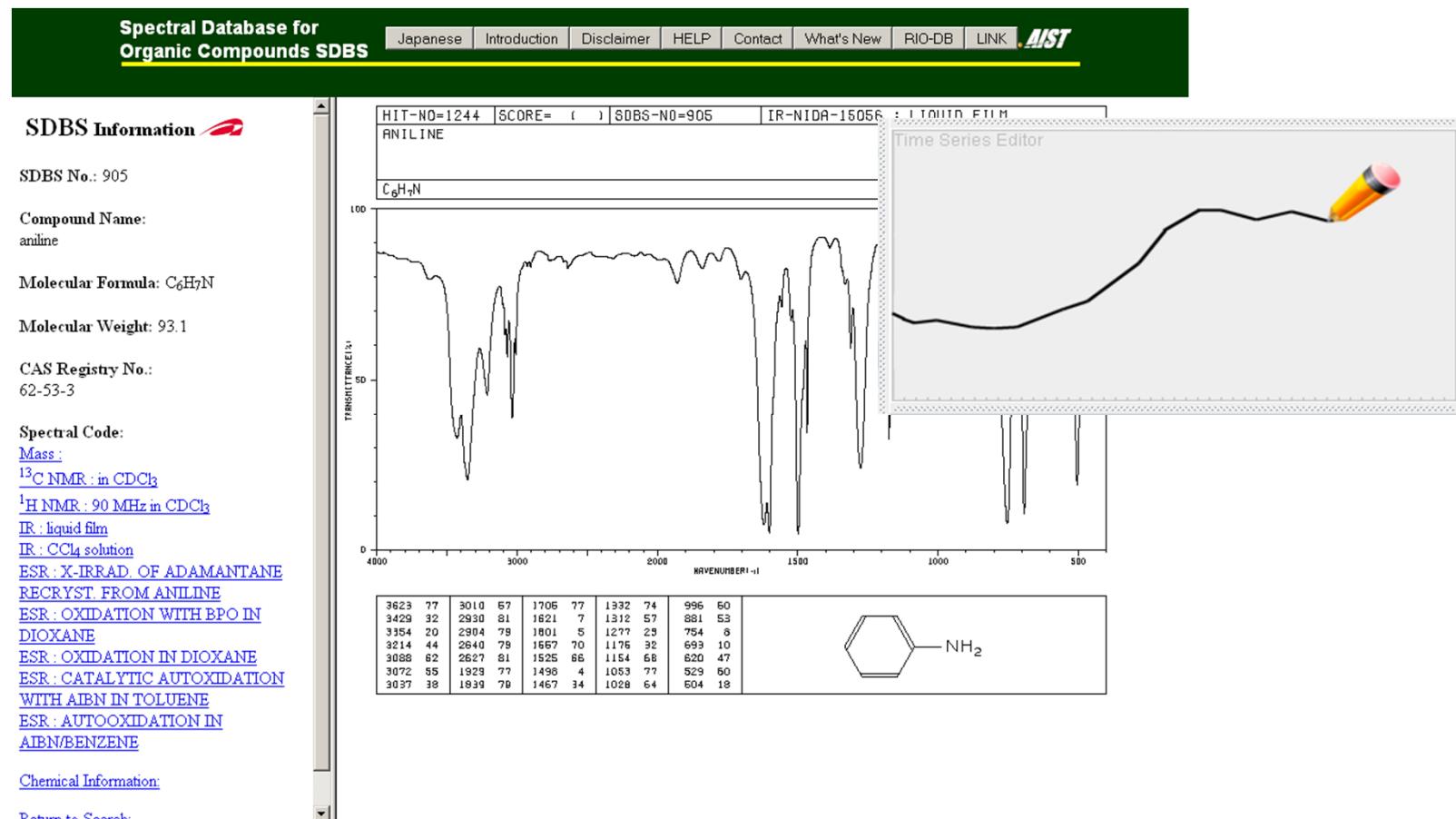


GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

Content based search



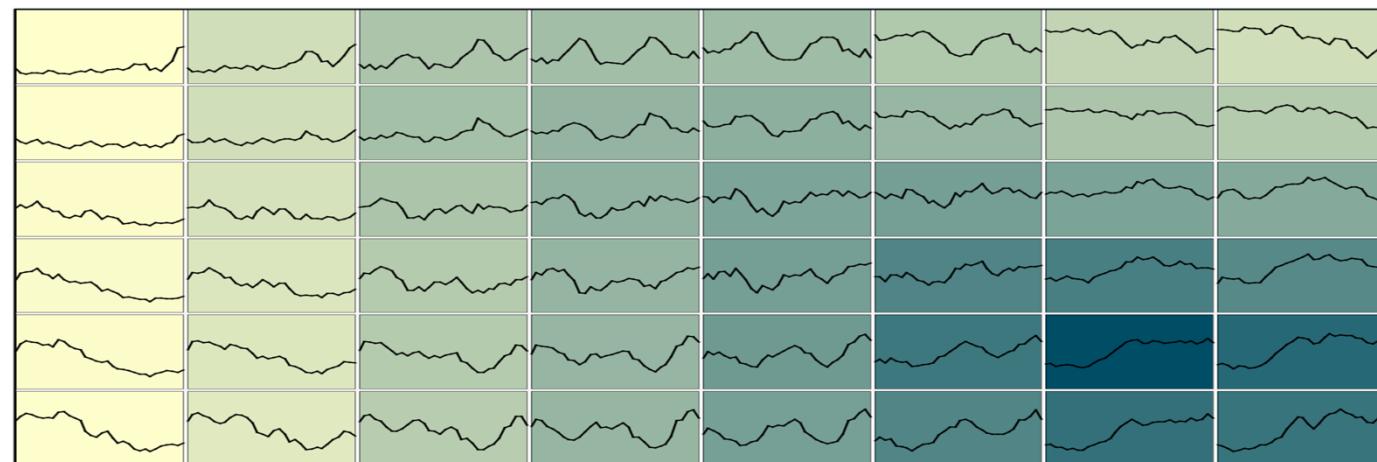
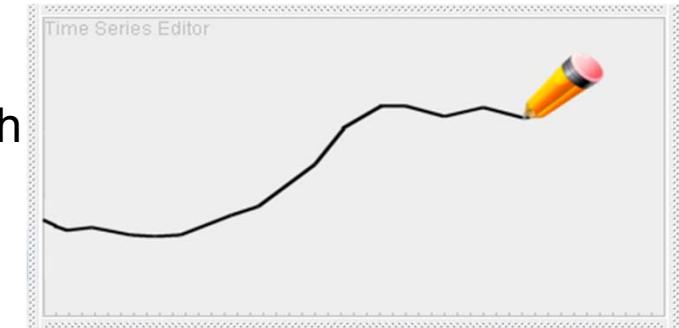
GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

Visual Search in Time series

- Query-by-Example, Query-by-Sketch
- Visual Catalog as result list
- Colormaps for the indication of similarity



Video cataloging with automatic indexing

Home Subjects Publisher About AV-Portal Watchlist Upload Login Register

Search for People, Places, Subjects ...

Follow us: [f](#) [t](#)

Back to shortlist

Email 0

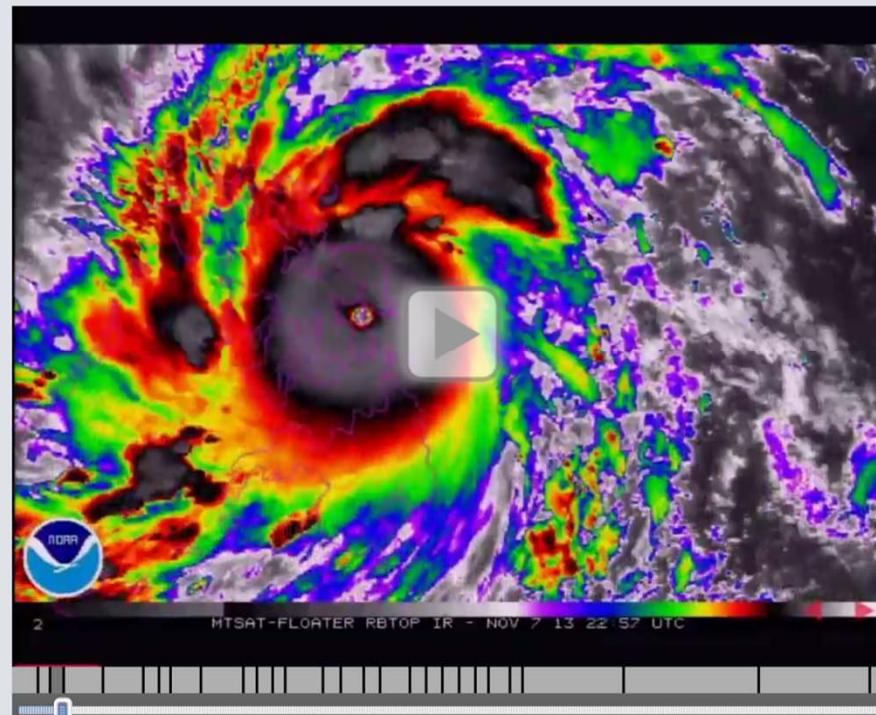
Recommend

Tweet

+1

Video detail page

Mapping for Good



Automated Media Analysis (i) BETA

Recognized Entities Full Audio-Transcript

Search ...

Speech Text in the video Image content

Service (economics) Computer animation Order (biology)

00:13 Computer animation

00:19 Computer animation Frequency

00:25 Computer animation

00:44 Denial-of-service attack Arithmetic progression
Lecture/Conference Meeting/Interview Case modding

01:03 Lecture/Conference Personal digital assistant



DataCite



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN



What if any kind of scientific content would be citable?

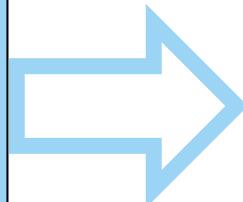
- High visibility of the content
- Easy re-use and verification.
- Scientific reputation for the collection and documentation of content (Citation Index)
- Encouraging the *Brussels declaration on STM publishing*
- Avoiding duplications
- Motivation for new research



DOI names for citations

URLs are not persistent

- (e.g. Wren JD: URL decay in MEDLINE- a 4-year follow-up study. Bioinformatics. 2008, Jun 1;24(11):1381-5).



The page cannot be found

The page you are looking for might have been removed, had its name changed, or is temporarily unavailable.

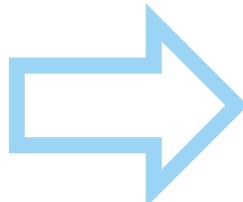
Please try the following:

- If you typed the page address in the Address bar, make sure that it is spelled correctly.
- Open the <http://apache.org> home page, and then look for links to the information you want.
- Click the [Back](#) button to try another link.
- Click [Search](#) to look for information on the Internet.

HTTP 404 - File not found
Internet Explorer

Digital Object Identifiers (DOI names) offer a solution

- Mostly widely used identifier for scientific articles
- Researchers, authors, publishers know how to use them
- Put datasets on the same playing field as articles



Dataset

Yancheva et al (2007). Analyses on sediment of Lake Maar.
PANGAEA.

[doi:10.1594/PANGAEA.587840](https://doi.org/10.1594/PANGAEA.587840)



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

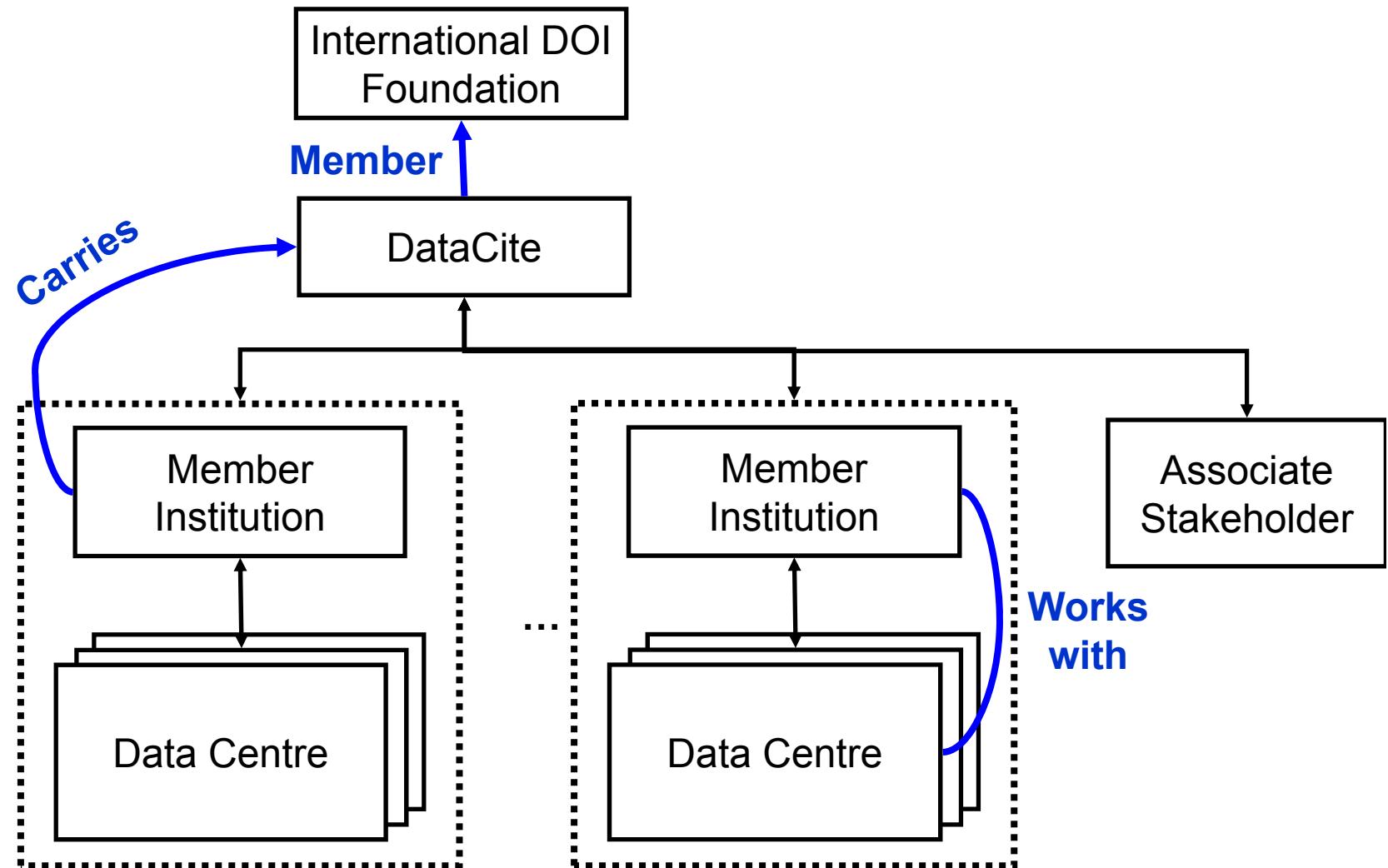
SUB

DataCite

- Global consortium carried by local institutions mostly libraries
- focused on improving the scholarly infrastructure around datasets and other non-textual information
- focused on working with data centres and organisations that hold content
- Providing standards, workflows and best-practice
- Initially, but not exclusively based on the DOI system
- Founded December 1st 2009 in London



DataCite structure



DataCite members

-
1. Technische Informationsbibliothek (TIB), Germany
 2. Göttingen State and University Library (SUB), Germany
 3. ZB MED, Germany
 4. ZBW, Germany
 5. Gesis, Germany
 6. Library of TU Delft, The Netherlands
 7. Technical Information Center of Denmark
 8. The British Library
 9. Library of ETH Zürich
 10. L'Institut de l'Information Scientifique et Technique (INIST), France
 14. Swedish National Data Service (SND)
 15. Australian National Data Service (ANDS)
 16. Conferenza dei Rettori delle Università Italiane (CRUI)
 17. National Research Council of Thailand (NRCT)
 18. The Hungarian Academy of Sciences
 19. University of Tartu, Estonia
 20. Bibsys, Norway
 21. Canada Institute for Scientific and Technical Information (CISTI),
 22. California Digital Library, USA
 23. Purdue University, USA
 24. Office of Scientific and Technical Information (OSTI), USA
 25. Japan Link Center (JaLC)
 26. South African Environmental Observation Network (SAEON)
 27. European Organisation for Nuclear Research (CERN)

Affiliated members:

1. Digital Curation Center (UK)
2. Microsoft Research
3. Interuniversity Consortium for Political and Social Research (ICPSR)
4. Korea Institute of Science and Technology Information (KISTI)
5. Beijing Genomic Institute (BGI)
6. IEEE
7. Harvard University Library
8. GWDG



Example

The dataset:

Storz, D et al. (2009):

Planktic foraminiferal flux and faunal composition of sediment trap L1_K276 in the northeastern Atlantic.

<http://dx.doi.org/10.1594/PANGAEA.724325>

Is supplement to the article:

Storz, David; Schulz, Hartmut; Waniek, Joanna J; Schulz-Bull, Detlef; Kucera, Michal (2009): *Seasonal and interannual variability of the planktic foraminiferal flux in the vicinity of the Azores Current.*

Deep-Sea Research Part I-Oceanographic Research Papers,
56(1), 107-124,

<http://dx.doi.org/10.1016/j.dsr.2008.08.009>



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

More Data example

- **Higgs particle**

ATLAS Collaboration (2013) HepData,

<http://doi.org/10.7484/INSPIREHEP.DATA.A78C.HK44>

- **ECOLI outbreak**

Li, D et al (2011):

Genomic data from Escherichia coli O104:H4 isolate TY-2482.

BGI Shenzhen.

<http://dx.doi.org/10.5524/100001>



Now what?

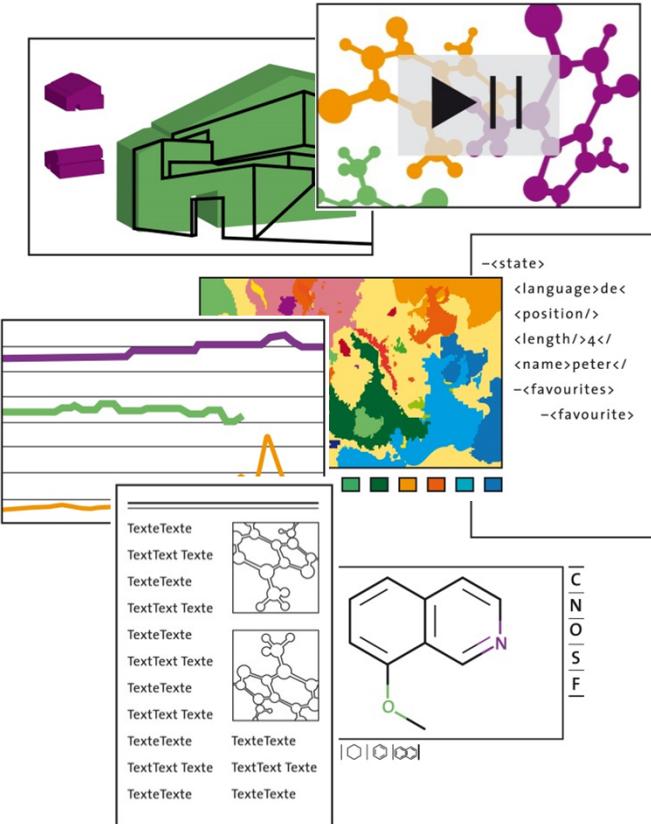
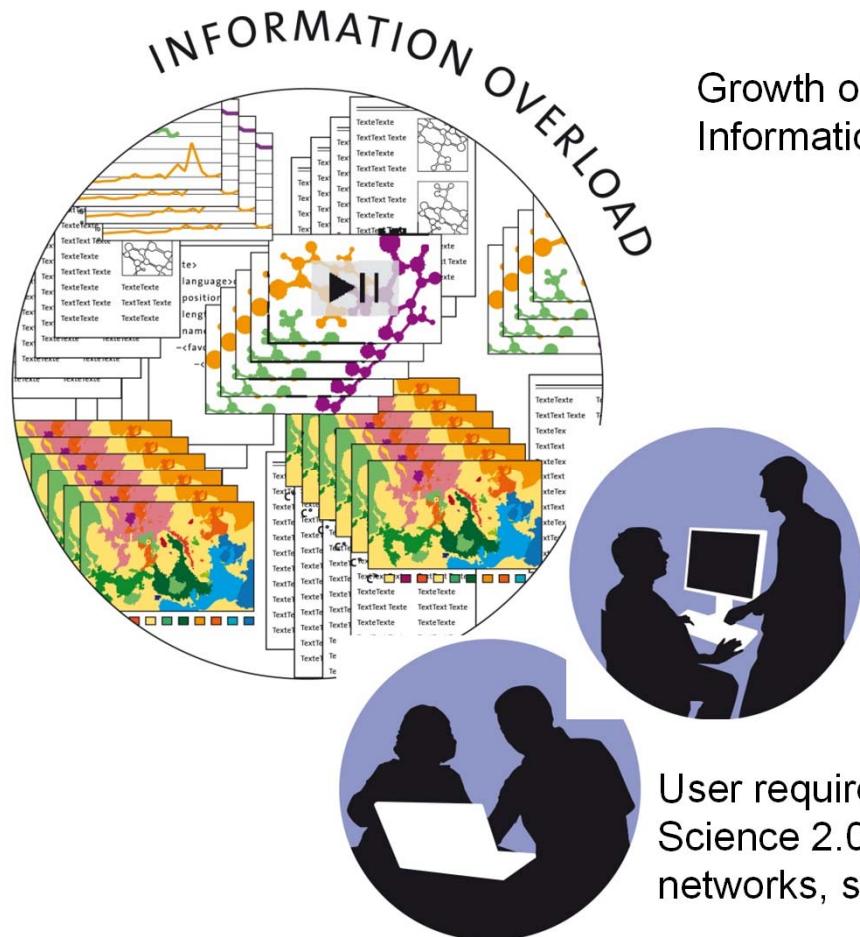


GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

The wave



Diversity of media types and formats

A threat?

- Information overload is only a problem for manual curation.
- Google is not complaining about data deluge—they're constantly trying to get *more* data.
- The more data you throw, the better the filter gets.
- **To develop and maintain these tools is a classical tasks for libraries!**
- **Don't turn off the taps, build boats.**

It is not only a challenge ...
... it is an opportunity

Libraries should ride the wave ...



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB

SUB

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

NIEDERSÄCHSISCHE STAATS- UND
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

SUB