

Riding the Wave: Move Beyond Text

Uwe Rosemann

German National Library of Science and Technology, Uwe.Rosemann@tib.uni-hannover.de

Irina Sens

German National Library of Science and Technology, Irina.Sens@tib.uni-hannover.de

Uwe Rosemann and Irina Sens, "Riding the Wave: Move Beyond Text." *Proceedings of the IATUL Conferences*. Paper 3.
<http://docs.lib.purdue.edu/iatul/2012/papers/3>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Riding the Wave: Move Beyond Text TIB's strategy in the context of non-textual materials

Uwe Rosemann, Irina Sens

Abstract Body

The European High Level Expert Group on Scientific data (2010) has formulated the challenges for a scientific infrastructure to be reached by 2030: "Our vision is a scientific e-infrastructure that supports seamless access, use, re-use, and trust of data. In a sense, the physical and technical infrastructure becomes invisible and the data themselves become the infrastructure – a valuable asset, on which science, technology, the economy and society can advance".

Here, "data" is not restricted to primary data but also includes all non-textual material (graphs, spectra, videos, 3D objects, etc.). The German National Library of Science and Technology (TIB) must now rise to these challenges: developing solutions together with the scientific community to make such data available, citable, sharable and usable, including visual search tools and enhanced content-based retrieval. With solutions such as DataCite and modular development for extraction, indexing and visual searching of new scientific metadata, TIB will ride this wave and will make all data accessible to its users fast, convenient and easy to use.

Fulltext

The Technische Informationsbibliothek Hannover (TIB) is the German National Library of Science and Technology. It is responsible for the literature and information provision for German industry and research in the specialised areas of technology, architecture, chemistry, computer science, mathematics and physics. Today, the TIB is the largest special library worldwide in its subject areas. It was founded in 1959 and this against the background of the early technological race to conquer space. The TIB is currently financed by the Federal Government and all the federal states and its total budget is 26 million Euros. The TIB is globally networked and has customers in 64 countries. Its main services are the provision of scientific content, the offer of an information portal for technology and natural sciences, a DOI service for the referencing of research data, as well as digital preservation of its scientific media and applied research and development of its services.

The main theme of this presentation is the consequences of the changes in the scientific process for scientific libraries. Today's scientists do not just require up-to-date literature, but also a comprehensive digital infrastructure in collaborative research which, alongside digital texts, also provides direct access to non-textual material such as images, statistics, software and other research data of all kinds.

There is generally a serious problem in research today; whilst scientific texts are, in principle, sufficiently well-documented and available, this does not apply to research data which is ascertained from experiments or other scientific studies and which forms the basis for the respective publications. Consequently, valuable (and expensive) scientific information is increasingly going missing!

This realisation has, of course, already been reflected in science policy and, as a result, there have been various experts' reports and recommendations which have emphasised the acute need for action - an example of this is given in a report from the European Commission entitled "Riding the wave. How Europe can gain access from the rising tide of scientific data."

In light of this, the TIB has adapted its strategy development, "Move beyond text". The TIB wants to position itself as an interface between scientists and scientific documents of all kinds, thus enabling access not only to scientific texts, but also scientific AV media, 3D objects, software, simulations and research data for its customers.

In order to be able to master this task operatively and naturally also financially, the TIB applied to the responsible ministries for additional funding. Thereupon, the TIB was evaluated twice by external experts who were provided with detailed information about its projects and services. A short time ago, political parties gave the go-ahead and the TIB is now going to establish a "Competence Centre for Non-Textual Materials."

What has happened thus far?

Back in 2005, the TIB became the first DOI registration agency for research data worldwide. The TIB had developed an infrastructure model which, in conjunction with scientific institutes, assumed the function of a data centre and allowed a complete workflow for the referencing of research data. This led - naturally in collaboration with other partners - to the founding of the DataCite consortium in London in 2009 which is now internationally established and offers a global service for the documentation of research data.

As part of the planned competence centre for non-textual materials, the TIB is intensively occupied with scientific AV media. It is already possible to search and view AV media in the TIB's GetInfo portal, but a special AV portal is now being developed which will offer a variety of specific visual search and indexing tools, such as object detection & clustering, genre analysis, Intelligent Character Recognition (ICR) or speaker detection and Automatic Speech Recognition (ASR).

Another big theme is the indexing and provision of 3D objects. The TIB, with various research institutes, carried out a project which has enabled the practical application of 3D objects of architectural models.

A further example of the application of non-textual methods is the handling of chemical structure data and its complete integration into the scientists' virtual research environment.

Another very research-intensive development area of the TIB is the methods for a visual search in research data. However, this presentation can only offer information on the initial steps, such as how graphical retrieval, a visual catalogue or a graphical results list of a search in numerical data could look in the future.

The TIB is already operating a large productive system for the digital preservation of digital texts with the national specialised libraries for medicine and economics. The problem of digital preservation is solvable for all the digital knowledge objects named in this presentation. The archiving of AV media or 3D objects does require specific technical and organisational solutions; in fact, the TIB is currently applying for an EU project with an international consortium in precisely this area.

With the establishment of its competence centre for non-textual materials and the provision of specific services, the TIB is planning to effectively support scientists' research processes in all the major aspects of information acquisition.

The TIB is pursuing the ultimate goal for its scientific users, "Interlinking and Search Across all Types of Digital Assets."

The TIB is getting ever closer to achieving this objective with its development projects and services.