stm
INTERNATIONAL ASSOCIATION OF SCIENTIFIC, TECHNICAL & MEDICAL PUBLISHERS
www.stm-assoc.org

28th November 2008

# STM Innovations Seminar

## New Streams, New Views, New Directions, Moving to a Future Beyond Text

Friday, 5 December 2008

Hilton London Kensington
179-199 Holland Park Avenue, London, UK

---

# Final Programme

**Program Chairs**
This event is co-chaired by Jonathan Clark (Elsevier, EVP Technology), David Martinsen (ACS Senior Scientist Strategic Planning and Analysis) and Eefke Smit (STM Director Standards and Technology).

8:45 - 9:30 Registration

9:30 - 10:30 **Is video the new killer app? Does STM have its own YouTube Generation?**
**Keynote: Mark Wood**, Chief Executive, ITN

This is a question often asked ever since YouTube launched in 2005, making it easy for anyone to post a video that millions could watch within a few minutes. What has been the impact of video since then? Is video changing the way people communicate? Has it changed the rules for media companies? And, most importantly, what future does it have in Research Publishing? Mark Wood is Chief Executive of ITN, one of the world's leading news and multimedia content companies. They create and supply content for a range of platforms from TV to radio, to mobile and online. Mark will explore the impact the world of the YouTube generation is having on traditional media businesses and will be looking for answers to these questions.
**Introduction: Jonathan Clark**, Vice President, Technology, Elsevier

10:30 - 11:00 Coffee Break

11:00 - 12:30 **Plenary Session: Visualisation**

Visualisation is not just 'cool' as a tool, it offers opportunities to understand and interact with data in ways that challenge the imagination. It enables what could not have been done in a paper world. This session will examine methods which provide end users with the ability to uncover trends and patterns within large collections of data, and also demonstrate tools for authors to communicate details about experiments which would be difficult with text alone.
**Chair**: **David Martinsen**, ACS Senior Scientist Strategic Planning and Analysis

11:00 - 11:30 **Shakespeare, God and Lonely Hearts: Transforming Data Access with Many Eyes**
**Jesse Kriss**, Research Developer, IBM Visual Communication Lab

Information visualization is no longer an experts-only domain. Like photos and video in earlier years, interactive visualization is starting to find its way on the web, becoming yet another medium for public discourse and personal expression. This is spurring all kinds of interesting side effects, creating changes not just in the scale of the audience, but in substance and purpose of the created artifacts.

11:30 - 12:00 **Visualisation for real - a publisher's experience**
**Toby Green**, Head of Publishing, OECD

It's all very well to play with pictures and data, the challenge is to build a publishing system capable of scaling and of connecting data objects with the scholarly publishing network via citations and library catalogues. Toby Green will present a series of case studies from OECD's first steps to include visual objects in its e-publications and online datasets and a proposal for bibliographic standards for data objects.

12:00 - 12:30 **Video Journal to Increase Efficiency and Standardization in Biological Research**
**Moshe Pritsker**, Editor-in-Chief and Co-founder of the Journal of Video Experiments

Biological science chronically suffers from the low reproducibility of experiments, as they are typically described in the Materials and Methods sections of scientific articles. The traditional text format of scientific journals cannot adequately describe complex experimental procedures, creating a critical bottleneck problem for biological research and drug discovery. Journal of Visualized Experiments (JoVE) addresses this problem using a novel video-based approach to scientific publishing. Visualization through video greatly facilitates the understanding of experimental procedures and increases the overall efficiency and standardization of biological research. So far, JoVE has published 21 monthly issues including over 250 video-articles on experimental approaches in neuroscience, immunology, developmental biology, microbiology and other fields. To facilitate integration of video into scientific publishing, JoVE has developed an organizational and technological structure to conduct production of scientific videos in research labs in the USA, Europe and Japan. For more information, visit the JoVE site www.jove.com .

12:30 - 14:00 Lunch

## 14:00 - 15:30 Parallel Sessions

**Parallel Session A: Raw Research Data**
Digital environments pose new opportunities for the access to large sets of original research data. By re-using existing data sets, research can be made more efficient and cost-effective and can be better validated and enhanced by means of re-analysis. There is increasing demand by authors in STM to add or refer to available datasets as mentioned in their publications. Science Infrastructure initiatives in various countries try to shape solutions for these opportunities. What is the publisher's role in the provision of and accessibility to raw research data ? And what kind of data mining tools can be offered to do more with content and data?
**Chair**: **Eefke Smit,** Director, Standards and Technology, International Association of STM Publishers

14:00 - 14:30 **Scientific articles enhanced with WEOs and other data-rich mechanisms.**
**Henry Rzepa**, Professor Computational Chemistry, Imperial College London

Professor Henry Rzepa has had research interests in both computational chemistry and chemical information for thirty years, and was a pioneer of implementing chemistry on the Web at its start in 1993. He is the author of around 250 scientific articles in the former and 50 in the latter category. The talk today will be a fusion of these interests, and in particular how modern informatics practices can substantially enhance the scientific perception of published articles, taking the "paper" well beyond its old cellulose based limits. The demonstration will comprise work carried out under the JISC-funded SPECTRa project: (Submission, Preservation and Exposure of Chemistry Teaching and Research Data). The project was based on the premise that chemistry and molecular sciences have two particular attributes; they are data rich, and that this data is finely grained within taxonomies that in some cases can be communally agreed upon. We selected three such areas; crystallography, NMR spectroscopy and computational chemistry, of which the latter will be demonstrated during the presentation.

14:30 - 15:00 **From Data to Knowledge: The Collection, Curation and Re-use of Research Data**
**Frank Allen**, CEO of CCDC

Since 1965, the CCDC has collected, validated and added value to primary research data from more than 465,000 published crystal structure determinations. The Cambridge Structural Database System, made available through an annual subscription model, also includes state of the art software for searching, visualising and analysing this data. The System is used at over 130 industrial companies and 2,000 academic sites worldwide as a basis for further research. Some 2,000 CSD-based publications describe applications in e.g. structural chemistry, drug design and development, crystal engineering and materials science.

15:00 - 15:30 **Raw Research Data in Germany: Building an Infrastructure**
**Jan Brase**, German National Library of Science and Technology, DOI Manager, TIB Hannover

Knowledge, as published through scientific literature, is the last step in a process originating from primary scientific data. These data are analysed, synthesised, interpreted, and the outcome of this process is published as a scientific article. Only a very small proportion of the original data are published in conventional scientific journals. Existing policies on data archiving notwithstanding, in today's practice data are primarily stored in private files, not in secure institutional repositories, and effectively are lost. This lack of access to scientific data is an obstacle, too much data remains underutilised.
To enable citations and better retrievability of data, the German Research Foundation (DFG) started the project Publication and Citation of Scientific Primary Data.
Starting with the field of earth science the German National Library of Science and Technology (TIB) has now established as a DOI-registration agency for scientific primary data. TIB provides a Digital Object Identifier (DOI) to registered data as a unique identifier. So far over 500,000 datasets have been registered by the TIB. C

**Parallel Session B: Digital Preservation**
Images, text, video, raw data sets, all of them cause data streams to grow in size. These need to be kept, stored and curated. Most of all, they pose new needs for digital preservation. What does it take to secure future availability of digital research sources, ranging from raw data to official publications? And whose role is it: institutes? libraries? (inter)national institutions? What can publishers add?
**Chair: Jonathan Clark**, Vice President, Technology, Elsevier

14:00 - 14:30 **Innovation and Preservation: Challenges in the E-Journal Space**
**Evan Owens**, CTO, Portico

Does Preservation always follow behind Innovation, perpetually doomed to play "catch-up"? Can we derive best practices for publishing innovation from experience in the preservation of existing electronic content? Portico, as part of its digital preservation program for published scholarly literature, has been intensively studying e-journal publication practices for the last six years. From that work has come a series of reports to the publishing community. This presentation, the fourth given at the STM Seminar Series, will explore the preservation challenges of innovations such as non-traditional units of content, article updating, and supplemental materials.

14:30 - 15:00 **How I Learned To Stop Worrying And Love New Media**
**Victoria Reich, David Rosenthal**, Board of Directors, Clockss

In January 1995 Jeff Rothenberg drew attention to the fragility of digital information by publishing "Ensuring the Longevity of Digital Documents". He identified media degradation, media obsolescence and format obsolescence as the three important threats. He prescribed regular media migration for the first two, and either format migration or emulation for the third. This vision has driven almost all subsequent work in digital preservation.
In May 1995 Stanford's HighWire press pioneered STM publishing's switch from paper to the world of digital media by putting the Journal of Biological Chemistry on the Web. The first journals used relatively few formats; every year since has bought new formats. STM journals are important and expensive. They have been a major focus of digital preservation efforts based on Rothenberg's vision, both format migration and emulation.
In October 1998 Stanford started work on the LOCKSS system, the first working technology for preserving electronic journals, and the basis for the CLOCKSS archive. Based on this experience we can report that, for reasons unrelated to preservation, the threats Rothenberg identified became insignificant. The actual problems of preserving STM journals, with their diversity of media, are quite different; scale, economics and intellectual property.

15:00 - 15:30 **Did I forget something? Roles and Responsibilities in Europe's Science landscape**
**Jeffrey van der Hoeven**, Project Manager Preservation, Royal Dutch Library, KB Netherlands

New trends in Europe's science landscape affect the roles and responsibilities of actors in this field. Sharing knowledge is a hot topic, but how do I know what is out there? Or was? Digital data sources need to be visible and sustainable at the same time. How do we achieve that? This presentation will give an overview of current changes in science in Europe and will identify topics that need to be solved to build Europe's science infrastructure.

15:30 - 16:00 **Break**

16:00 - 16:30 **Plenary Closing Session: Current Strategy and Future Directions of Scholarly Communication at Microsoft**
**Lee Dirks** -- Director, Education & Scholarly Communication, Microsoft External Research

New authoring tools will facilitate the recording of research results, dissemination of first results, submission of manuscripts and data sets, and much more. Authors nowadays use a variety of tools to create and submit their documents and data as part of the scholarly publishing process. There is usually some type of conversion required on the part of the publisher to incorporate these documents into the published collection. Although XML-based tools have been available for several years, these tools are not yet widely used by the scholarly authoring community. Microsoft's Scholarly Communications Division is seeking to address this void, but their vision goes beyond simply preparing documents for submission to publishers. This plenary talk will examine Microsoft's plans to empower scientists and other scholars to record and disseminate their manuscripts, experimental data, and other digital material to publishers and repositories, and to enable publishers to more easily process and publish submissions from authors.
**Chair**: **David Martinsen**, ACS Senior Scientist Strategic Planning and Analysis

STM and the STM Innovations Seminar Director, Eefke Smit, would like to thank the programme committee listed below for the hard work, commitment and time they have put in to helping organise the STM Innovations Seminar 2008 at Hilton London Kensington.

**Programme Committee**
David Martinsen - ACS
Jonathan Clark - Elsevier
Howard Ratner - Nature Publishing Group
Olaf Ernst - Springer
John Sack - HighWire Press
Tim Ingoldsby - AIP

**STM thanks the following organisations for helping sponsor the 2008 Innovations seminar:**

atypon systems

Scitation
A service of AIP