

Stefan Gradmann & Jill Cousins



Humbolt University, DE

stefan.gradmann@uni-hamburg.de

Stefan Gradmann is Professor for Library and Information Science at Humboldt University in Berlin. Stefan started as a literary scholar (mainly working on Kafka and Joyce). He was a librarian in his 'second life', working in the library automation industry and was director of the GBV union catalogue. He directed Hamburg University's Computing Center in his 'third life' before now taking over his new position in Berlin. The focus of his work is on Digital Library architectures, DL interoperability and semantic web issues as well as on eScholarship and digital humanities issues.

Jill Cousins



The European Library

Jill.Cousins@KB.nl

Jill Cousins is Director of The European Library, Programme Director of *Europeana* and Executive Director of the EDL Foundation. Jill has many years experience in web publishing. Her past includes moving from the extremely commercial publishing world where she was European Business Development Director of VNU New Media to scholarly publishing. She managed the online journals section of Blackwell Publishing and set up the Oxford Publishing Society. Prior to publishing she had a variety of marketing and research careers in the information field. These ranged from being the Marketing and Event Director for Learned Information to managing her own research company, First Contact.

GRL2020 Position Paper

There are 2 major areas where our current work impacts on digital libraries. The first is largely the preserve of Jill Cousins and The European Library team. The second is a project *Europeana*, to produce a cross domain prototype, solving some of the issues and highlighting future necessary

research areas. This is run by the EDLnet Office with assistance from several external experts, such as Stefan Gradmann, as work package leaders.

1. The European Library is housed at the National Library of the Netherlands and is a service of the Conference of National Librarians. It is a joint portal bringing together the digital and non digital material of all the national libraries of Europe. The European Library provides a unique access point to the collections and catalogues of 33 national libraries. The amount of digital material is in the region of 6 million items. By the end of 2008, the number of participating libraries will be 45 of the 47 possible members of CENL. It is the only operational pan European service that accesses information at the item level for the cultural heritage sector.

2. *Europeana*: European digital library network, is a project to create a cross domain (museums, libraries, archives, audio visual collections) portal. It currently has 90 such organizations in its network, contributing to 5 work groups, aiming at solving the human and political; as well as, the technical and semantic issues of creating the joint portal. This project began in July 2007 and has so far produced a series of deliverables and an online demonstrations. The demonstrations, found at www.europeana.eu, shows what users have asked to be built. This demonstration is being subjected to user testing to determine what to develop for the prototype, due in November 2008. The prototype is a proof of concept. It will not put in place many of the back end processes that are integral to a service such as The European Library. *Europeana* has also uncovered several areas where it recommends further research and/or development.

There are a number of requirements that need to be addressed for further adoption of Digital Libraries.

The major requirement for both The European Library and Europeana is to be able to give access to the material held in the cultural heritage sector that will meet user requirements. This would be done by means of state of the art W3C architecture and would include awareness of semantic and multilingual interoperability. This means the ability to find things without determining first where they are housed. Searching should be seamless and results integrated and ranked. All 3 of these still have some way to go in the cross domain field.

A second requirement is that data needs to be truly interoperable regardless of its original format. This requires: a functional and architecture model built upon standards of the W3C architecture instead of proprietary library automation solutions; a model for digital information objects (f.a.k. as 'documents') that works across the cultural heritage divide; new solutions for data presentation and visualization; systematic automatic generation and improvement of metadata so that basic fields such as author, title, subject, description, time and place are standardized and shown.

Basic centralized tools that are accessible to all research is a further requirement. This includes multilingual searches with an inventory and mapping of all the relevant ontologies and tools for use with various software tools, requiring a knowledge of all the online dictionaries and currently non accounted for languages. In addition, a registry of all the authority files would be useful for both development and the filling of gaps. A registry of work on licensing and authorization role semantics and a registry of work on digital thumbprints and watermarking would also be examples of basic centralized tools that are needed.

Other requirements for digital libraries include: ranking capability to search results without the benefit of Google algorithm; Data weighting and results ranking; User influenced ranking using session cookies and logfile analysis; Scalable applications for multilingual searching; Scalable browsing that combines the vectors of space and time with who and what; persistent identifiers for cultural heritage digitized and born digital content.

It is important to identify the current state of the art. Recently completed European Library activities include the [EDLproject](#) which extends the multilingual capacity of the network, develops the European Library network's localisation and multilingual capabilities. This is aimed at improving access for end-users through multi language interfaces and advanced search mechanisms in a standardised way. The M2.1 Report on Unicode M2.2 Report on subject access tools are relevant to this subject.

At Present The European Library activities include: TELplus, an eContentPlus project that aims to do the following:

- To add and improve content of The European Library by: OCRing more than 20 million pages of important multi-lingual content now available only as images according to current best practice; Making the data that national libraries currently hold in proprietary databases (only searchable via the z39:50 protocol) OAI compliant and therefore harvestable for far greater usability; Adding the National Libraries of Bulgaria and Romania as full members of The European Library.
- To improve access and usability by: improving full text indexing and investigating automatic vocabulary mappings; learning from users through The European Library user group; creating a modular service infrastructure that enables users to integrate services with the European Library portal and adding a number of new services; work on user personalization to provide directions towards new services for users; providing translations of collection descriptions in 26 languages.
- Europeana / EDLnet (Europeana.eu) is responsible for the following deliverables of relevance:
 - The maquette showing what has been requested by users www.Europeana.eu
 - Deliverable 2.2 Initial Semantic and Technical Interoperability Requirements

Further deliverables from this project will be very important in informing both the state of the art and recommendations for research in digital libraries.

The vision is for a more coordinated approach on research, infrastructure and content supported by the European Commission. This is improving, but it seems that few people have an overview of all relevant projects and how they contribute to the whole picture. Piecemeal concertation events rarely have concrete, executable outcomes and forcing projects to conform to *Europeana's* goals has some short term benefits but could have a negative longer term effect. A network of networks coordinated at a European level by a new agency could enable the decentralization necessary but keep the duplication to a minimum and ensure that projects compliment each other. This has been given a kick start by the establishment of the EDL Foundation, but this Foundation needs a sustainable funding source to continue. Such an Agency could have strong links with the other continents work and research in the area of digital libraries, cross fertilizing and re using of research and applications.

Europeana.eu is a concrete step towards a common, coordinated access point to Europe's culture. The work of the Driver project is also very important in this goal. Furthermore, co-operation with initiatives such as OAI ORE for modeling complex digital information objects using standard web technology and architectures is currently building up. Other projects currently in negotiation: Arrow, Athena, EDLocal, EFG and EPA all contribute to creation and delivery of standards and content from the cultural heritage sectors. Next steps: Submissions under the current FP7/ICT and eContentPlus calls for research and development in the missing blocks and the creation of an operational Europeana.eu as a continuation of the work and the network of Europeana / EDLnet. EDL Foundation: to be put on a sounder financial footing.