

# Session 2: Aggregation & (Meta) Data Quality

Europeana Tech Conference  
4th of October 2011



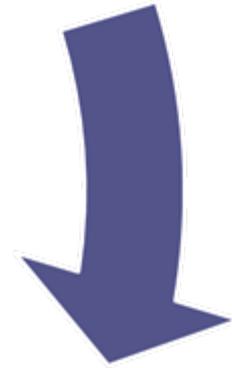
## Evaluation

- Navigating the troubled waters of W3C standards
- What have we achieved that Google hasn't?



## Content

- A look at (meta)data
- What opportunities lie ahead with the EDM



## Content aggregation

- Establishing sustainable aggregation infrastructures
- Content enrichment



## Use and Re-use

- Exploring rich content
- EDM field-trial within the domain of Archaeology
- Meet the Semantic Web



# Franziska de Jong

«*There is no data like Metadata*»

- Professor of language technology at the University of Twente
- Director of the Erasmus Studio at the Erasmus University in Rotterdam
- Main research interest is in the field of multimedia indexing, text mining, semantic access, cross-language retrieval and the disclosure of cultural heritage collections
- Frequently involved in international programme committees, expert groups and review panels.
- Was a member of the EU/NSF 'spoken word archives' working group.
- Principal investigator of the NWO-CATCH project CHoral (2006-2010) and coordinator of IST project PuppyIR (2009-2012).
- Since 2008 a member of the Governing Board of the Netherlands Organization for Scientific Research (NWO).



# Stefan Gradmann

## *The opportunities of the Europeana Data Model*

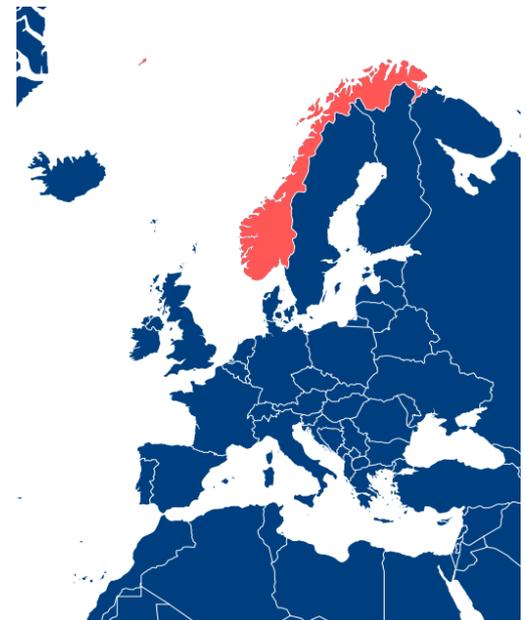
- Professor teaching knowledge management and semantic knowledge architectures at the School of Library and Information Science of Humboldt-Universität since 2008
- Other focal areas include digital libraries, library automation and use of information technology in the realm of sense and interpretation in the 'Digital Humanities'
- Has been heavily involved in the creation of Europeana, where he is responsible for semantic interoperability.
- Since 2008, Stefan Gradmann is president of the German Society of Information Science and Information Practice (DGI).



# Gunnar Urtegaard

*National Infrastructure and Content Aggregation in Norway*

- At the helm of the unit of e-culture at Arts Council Norway.
- For 20 years he was director of the County Archive of Sogn og Fjordane.
- He has been project manager for a number of projects involving new technology and cultural institutions including both in Norway and abroad
- He is an experienced “practitioner” in establishing regional and national harvesting infrastructures.
- Recent involvement in Europeana include EuropeanaLocal



# Nuno Freire

## *Information Extraction in Metadata*

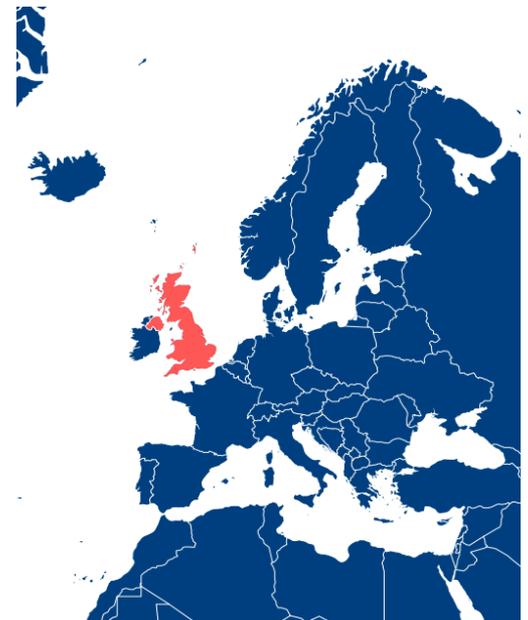
- Interoperability Architect at The European Library with an MSc. in Informatics and Computer Engineering from Instituto Superior Tecnico of the Technical University of Lisbon.
- His career so far has been focused on international projects in the area of digital libraries.
- Areas of interest include interoperability, information retrieval and data quality, particularly in its application to digital libraries and catalogues.
- He participated in several European projects including NEDLIB, TELproject, LEAF, MALVINE, DIGMAP, TELplus, Europeana Connect, ARROW, ARROW Plus and EuropeanaLibraries.



# Mark Hall

*Aggregating cultural heritage collections using automatically generated topic hierarchies*

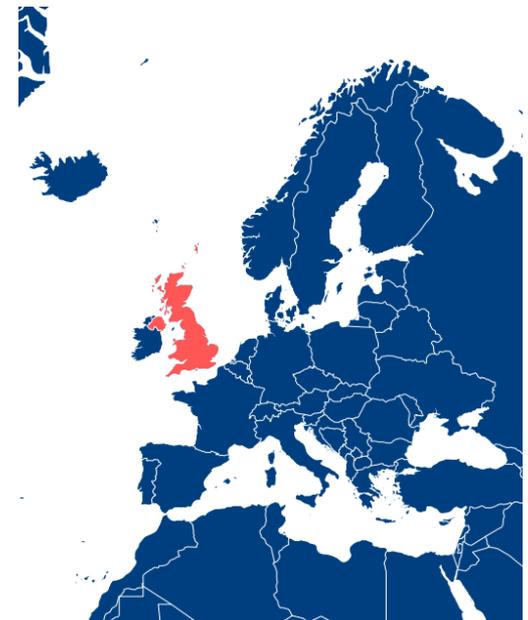
- Research assistant with the Information School at The University of Sheffield, UK
- His research interests focus on computationally modeling human behaviour to improve man-machine communication. He also maintains an interest in information retrieval, spatial data, computational linguistics, interface design / usability, and automatic data-integration.
- Did his PhD at Cardiff University, Wales on the automatic interpretation and generation of spatial language. He then spent ten months working at The UK National Archives, developing systems to automatically geo-locate their historic data-sets
- Current work focuses on developing user interfaces for navigating large cultural heritage collections in the PATHS project



# Kate Fernie

## *The suitability of EDM for domain specific content, CARARE*

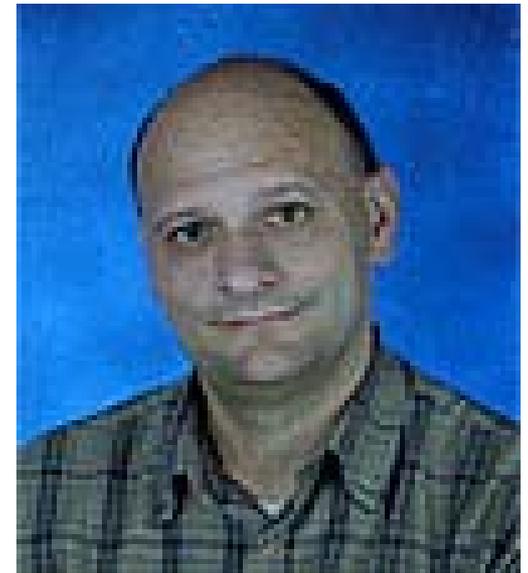
- Technical coordinator for the CARARE project on behalf of MDR Partners.
- From 2004-2006 ICT Advisor at the Museums, Libraries and Archives Council managing the CALIMERA, MINERVA, BRICKS and MICHAEL projects.
- She joined MLA from English Heritage where she was responsible for promoting standards and systems for heritage information management to over 100 local authorities in the UK.
- During a 3-year secondment to the Archaeology Data Service Kate produced on-line learning resources, was editor to a series of best practice guidelines and was coordinator to the Historic Environment Information Resources Network.
- Since 2008 she has worked as a project manager with MDR Partners, as a specialist adviser and editor and as Digital Preservation Specialist at the UK's Parliamentary Archives.
- Kate is currently project manager for CARARE, PATHS and DigCurV



# Paul Hermans

## *Metadata and the Semantic Web*

- Lead architect of [Erfgoedplus.be](http://Erfgoedplus.be) ('heritage-plus'), a Belgian project aimed at disseminating all types of heritage from the provinces of Limburg and Flemish Brabant and the city of Leuven to the public by applying semantic web technology.
- [Erfgoedplus.be](http://Erfgoedplus.be) uses RDF/XML, OWL and SKOS to describe relationships to heritage types, concepts, objects, people, place and time.
- He is the technical architect and developer of the Dublin Core based metadata standard of the Dutch Government, OWMS, and of the Linked Data publication of the related vocabularies.
- Hermans has been chairman of the Belgian-Luxembourgian SGML/XML User's Group for years and he is a fellow of the Hogenhevelcollege KULeuven, research group Informatics (LIRIS).



# Phil Archer

## *Choosing the Right W3C Standard*

- Joined the W3C staff after many years representing one its member organisations.
- Chair of the first Incubator Group, a lightweight mechanism for developing new ideas for the Web, he went on to chair the POWDER Working Group whose standards form part of the Semantic Web technology suite.
- Was an original member of the Mobile Web Best practices Working Group.
- As well as work at the W3C Phil's career has encompassed broadcasting, teaching, linked data publishing, copy writing, and, perhaps incongruously, countryside conservation.

