



ONE-DAY SEMINAR ORGANIZED BY AXE 4 GERIICO,  
UNIVERSITY OF LILLE

Date : 18th of March / 18 mars

Time : 10:30 am - 4h45 pm local time / 10h30 - 16h45

Lien de connexion : <https://univ-lille-fr.zoom.us/j/93156342957>

## WEB OF DATA AND SEMANTIC ENRICHMENT

## WEB DE DONNÉES ET ENRICHISSEMENT SÉMANTIQUE





# PROGRAM

## MORNING SESSION

### **INTRODUCTION (10H30 - 10H45)**

Widad Mustafa El Hadi

### **TRANSFORMING THE INTEGRATIVE LEVELS CLASSIFICATION SCHEME TO SKOS RDF (10H45 - 11H15)**

Douglas Tudhope, Faculty of  
Computing, Engineering and  
Science, University of South Wales,  
UK

### **CROWDSOURCING AND CLASSIFICATION: GUIDELINES FOR TECHNICAL AND SOCIOLOGICAL ASSESSMENT (11H20 - 11H40)**

Daniel Martinez Avila, iSchool,  
Universidad Carlos III de Madrid,  
España, Professeur invité UDL,  
département SID-Geriico

Discussion

# PROGRAM

## AFTERNOON SESSION

Moderateur/Chair Amel Fraisse

### **WEB-SCALE DATA INTEGRATION FOR CREATING CYBER-PHYSICAL SOCIAL SYSTEMS (14H - 14H30)**

Marta Sabou, Vienna University of  
Technology, Austria

Moderateur/Chair Laurence Favier

### **LIBRARY LINKED DATA SCHEMAS IN THE SEMANTIC WEB (14H40- 15H10)**

Hyoungjoo Park & Margaret E. I. Kipp,  
School of Information Studies at the  
University of Wisconsin – Milwaukee, USA

### **STRUCTURER ET ÉDITORIALISER DES DONNÉES DU SECTEUR CULTUREL : PROJETS ET EXPÉRIMENTATIONS AU SEIN DE L'INSTITUT NATIONAL D'HISTOIRE DE L'ART (15H10- 15H40)**

Antoine Courtin, maître de conférences  
associé à l'Université Paris Nanterre, INHA

Discussion

**Concluding remarks & closing session**  
Chair Joana Casenave



# ABSTRACT /RÉSUMÉ

## TRANSFORMING THE INTEGRATIVE LEVELS CLASSIFICATION SCHEME TO SKOS RDF

The Integrative Levels Classification (ILC) is a comprehensive 'freely faceted' knowledge organization system not previously expressed as SKOS. This presentation reflects on recent work, in collaboration with Ceri Binding and Claudio Gnoli, transforming the ILC to SKOS RDF representation. The ILC is a complex classification scheme with various features that are challenging to represent in SKOS. Previous work expressing other complex classification schemes in SKOS is briefly reviewed. The ILC SKOS representation is described and significant aspects of the implementation of the transformation using the STELETO transformation tool are described. Various issues arising from the SKOS transformation, both technical and theoretical, are discussed. These include the challenges posed by compound faceted descriptors with coordination of concepts, a limited extension of SKOS to incorporate RDF properties in order to model the ILC's facet indicators together with specialized sub-properties of `skos:related`. Some possible routes towards addressing the complexity of complex classification schemes in indexing and search applications are suggested.



# ABSTRACT /RÉSUMÉ

## **CROWDSOURCING AND CLASSIFICATION: GUIDELINES FOR TECHNICAL AND SOCIOLOGICAL ASSESSMENT**

Crowdsourcing is a broad term referring to the collaboration of large numbers of distributed people on a common product, usually organized through information. Crowdsourcing is a concept originated within the open source/free software movement that has been applied in multiple contexts. In Library and Information Science and knowledge organization, it has been used for the development of classifications. This presentation discusses the assessment of a crowdsourcing classification using the guidelines for the open-source software development model (f Eric Raymond's "The Cathedral and the Bazaar" and the sociological analysis of "Homesteading the Noosphere"). As a case study, it uses the guidelines for the assessment of the Open Shelves Classification developed by the LibraryThing community



# ABSTRACT /RÉSUMÉ

## WEB-SCALE DATA INTEGRATION FOR CREATING CYBER-PHYSICAL SOCIAL SYSTEM

This study examined linked data models proposed or in use by libraries around the world. The researchers examined different library types (e.g., national libraries, academic libraries, government libraries, public libraries) and cultural heritage centres around the world (e.g., North America, European Union and Asia). Library linked data models world wide were constructed using a wide variety of data models and term choices, though the final records were designed to describe the same types of materials using the same concepts. The choices of linked data models in each library can create interoperability concerns for linked data services due to the substantial differences in linked data models developed via local decisions. The researchers further examined user perceptions and understanding of metadata elements used in library linked data services.



# ABSTRACT /RÉSUMÉ

## STRUCTURER ET ÉDITORIALISER DES DONNÉES DU SECTEUR CULTUREL : PROJETS ET EXPÉRIMENTATIONS AU SEIN DE L'INSTITUT NATIONAL D'HISTOIRE DE L'ART

Cette intervention aura pour but de fournir un panorama des principaux projets en cours développés au service numérique de la recherche de l'INHA. À partir de ces cas d'études émergeront certains principes ou stratégies mis en place à l'intersection entre l'exigence des bonnes pratiques afin d'intégrer ces données dans l'éco-système du web données (alignement de référentiels, stratégie d'exposition de données, projet de réinformatisation, etc.) et la proposition de formats ou object numérique hybrides permettant d'offrir un espace d'éditorialisation personnalisé à chaque projet de recherche (data storytelling, blogging scientifique, etc.).



# ABSTRACT /RÉSUMÉ

## LIBRARY LINKED DATA SCHEMAS IN THE SEMANTIC WEB

Cyber-physical Social System (CPSS) are complex systems that span the boundaries of the cyber, physical and social spheres. They play an important role in a variety of domains ranging from industry to smart city applications. As such, these systems necessarily need to take into account, combine and make sense of heterogeneous data sources from legacy systems, from the physical layer and also the social groups that are part of/use the system. The collection, cleansing and integration of these data sources represents a major effort not only during the operation of the system, but also during its engineering and design. To address this need, we develop methods and techniques based on Semantic Web and Linked Data technologies for the acquisition and integration of heterogeneous data from disparate structured, semi-structured and unstructured sources, including open data and social data. In this talk, we present the overall system architecture with a core focus on data acquisition and integration and its prototypical implementation within an adaptive planning use case for public transportation scheduling.





# SHORT BIO

## DANIEL MARTÍNEZ ÁVILA

Daniel Martínez Ávila received his PhD degree in Documentation from Universidad Carlos III de Madrid in 2012. He worked as a researcher and adjunct instructor at the School of Information Studies at the University of Wisconsin-Milwaukee, United States, in 2013-2014. In 2013 he also taught at Bryant & Stratton College Milwaukee in the United States. In 2014 he started to work at the Department of Information Science of São Paulo State University (UNESP), in Brazil.

During his time at UNESP (2014-2019), he co-chaired the doctoral program in Information Science. He has been a guest professor in the graduate programs in Information Science of the University of Salamanca (2015) and the University of Coimbra (2017-2018), Portugal.

Today, he still collaborates with the Graduate Program in Information Science at São Paulo State University (UNESP), Brazil, and supervises several master's students and doctoral students at UNESP and at the University of Coimbra, Portugal. Since 2014 he serves as International Coordinator of the Satija Research Foundation for Library and Information Science (SRFLIS), India. He is a member of the ISKO (International Society for Knowledge Organization) Scientific Advisory Council, a member of the Institute for Knowledge Organization and Structure, United States, a member of the Committee 50 of AENOR (the Spanish Association for Standardization and Certification), and a member of several editorial boards of JCR and Scopus indexed journals from Germany, South Korea, and the Netherlands. He has been invited to give talks at the University of Urbana-Champaign, United States, the University of Wisconsin-Milwaukee, United States, and the University of Copenhagen, Denmark, and he has authored more than 100 publications including journal articles, books, book chapters, and conference papers published in the United States, United Kingdom, Spain, France, Germany, Netherlands, India, Brazil, Portugal, South Korea, Sri Lanka, Mexico, and Argentina.



# SHORT BIO

## ANTOINE COURTIN

Formé initialement en histoire de l'art puis à l'École Nationale des Chartes en master en nouvelles technologies, diplômé en 2010, Antoine Courtin a été successivement product manager dans le secteur privé, ingénieur d'études au sein du Labex Les passés dans le présent avant d'intégrer l'Institut National d'Histoire de l'Art en 2015. Il y développe un service numérique de la recherche dont il en devient le responsable en 2019. En parallèle, Antoine Courtin devient en 2017, maître de conférences associé à l'Université Paris Nanterre. Cv en ligne : [antoinecourtin.github.io/markdown-cv/](https://antoinecourtin.github.io/markdown-cv/) | Twitter : [@seeksanuersname](https://twitter.com/seeksanuersname) |



# SHORT BIO

**MARGARET E. I. KIPP**

Margaret E. I. Kipp is an Associate Professor and coordinator of the Knowledge Organization Research Group, School of Information Studies, University of Wisconsin-Milwaukee. Her research interests include social tagging, Linked Data and the Semantic Web, data science, and information retrieval. She is currently working on several other projects in Linked Data including a project modelling Chinese Name and Title authorities. Previously, she worked as a programmer/analyst in Geographic Information Systems (GIS), including on the design of data models and user interfaces in multiple languages for GIS databases.



# SHORT BIO

## HYOUNGJOO PARK

Hyounghoo Park is a visiting assistant professor in the School of Information Studies at the University of Wisconsin - Milwaukee. Her research focuses on data/software curation, sharing, reuse, citation and management in scholarly communication as well as linked open data. She was awarded the Eugene Garfield Doctoral Dissertation Fellowship by the Beta Phi Mu international honor society. She received a Data Share Fellowship with the Whole Tale - Research Data Alliance. Dr. Park earlier worked as a research and development analyst with the LG Corporation.



# SHORT BIO

## MARTA SABOU

Dr. Marta Sabou is an FWF Elise-Richter Fellow at the Vienna University of Technology, where she leads the Semantic Systems Research Lab which performs foundational and applied research in the area of information systems enabled by semantic (web) technologies. Prior to this she held positions as Research Fellow at the Knowledge Media Institute (Open University, UK), Assistant Professor at the Department of New Media Technology (MODUL University, AT) and Key Expert in Semantic Technologies (Siemens). Her work is situated at the confluence of Semantic Web and Human Computation research areas. She is an accomplished academic (over 100 peer-reviewed papers, h-index 45) and takes an active role in the Semantic Web research community. She acts as an editorial board member for three journals that publish Semantic Web research (SWJ, IJSWIS, JoDS) and has been engaged in several senior conference organization activities, including: Knowledge Graph Track co-chair at ESWC'2021; WebScience Track co-chair at WWW'16; Resources Track co-chair at ISWC'16 and program co-chair for ESWC'15 and iSemantics'2013.



# SHORT BIO

## DOUGLAS TUDHOPE

Douglas Tudhope is Professor in the Faculty of Computing, Engineering and Science, University of South Wales and leads the Hypermedia Research Group. He was PI on the AHRC funded STAR, STELLAR and SENESCHAL projects (Semantic Tools for Archaeological Resources), in collaboration with English Heritage and the preceding EPSRC funded FACET project, in collaboration with the Science Museum, investigating thesaurus-based query expansion. He led the Linking Archaeological Data Work Package for the ARIADNE FP7 Infrastructures Project. Since 1977, he has been Editor of the journal, *New Review of Hypermedia and Multimedia*. He is a member of the Networked Knowledge Organisation Systems/Services (NKOS) network and has organised or co-organised NKOS workshops at ECDL/TPDL, DCMI. He was a member ISO TC46/SC9/SC8 (and NISO) working group developing a new thesaurus standard (ISO 25964).