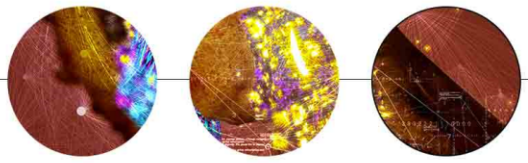
FAIR Heritage: Digital Methods, Scholarly Editing, and Tools for Cultural and Natural Heritage

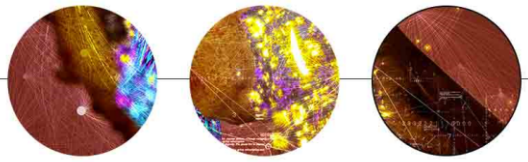
Wednesday 17th June & Thursday 18th June 2020

VIRTUAL MEETING



Scientific committee

Emmanuelle BRYAS French National Institute for Preventive Aecheological Research (Inrap) - France	**Pierre-Yves BUARD** House of Research in Social Sciences and Humanities of Caen (MRSH), University of Caen - France
Johann FORTE Centre for Advanced Studies in the Renaissance (CESR) / CNRS, University of Tours - France	**Roberta FERRARIO** Laboratory for Applied Ontology (LOA), ISTC-CNR - Italy
Ludger JANSEN Institute of Philosophy, Rostock University - Germany	**Corinne JOUYS-BARBELIN** National Archaeology Museum of Saint-Germain-en-Laye - France
Adeline JOFFRES TGIR Huma-Num / CNRS, University Aix-Marseille - France	**Marion LAMÉ** LE STUDIUM Research Fellow, ARD Intelligence des Patrimoines, Centre for Advanced Studies in the Renaissance (CESR) / CNRS, University of Tours - France
Béatrice MARKHOFF Laboratory of Fundamental and Applied Computer Science of Tours (LIFAT), University of Tours - France	**Alessandro MOSCA** Faculty of Computer Science, Free University of Bozen-Bolzano - Italy
Olivier MARLET CITERES / CNRS, MSH, University of Tours - France	**Aurélien MONTAGU** Le Studium Loire Valley Institute for Advanced Studies – Orléans & Tours - France
Marianna NICOLOSI ASMUNDO Department of Mathematics and Computer Science, University of Catania - Italy	**Blandine NOUVEL** Mediterranean House of Human Sciences & GdS Frantiq / CNRS, University Aix-Marseille - France
Elena PIERAZZO Centre for Advanced Studies in the Renaissance (CESR) / CNRS, University of Tours - France	**Nathalie LE TELLIER-BECQUART** House of Social Sciences (MSH) Mondes / CNRS, University Paris Nanterre & Paris 1 Panthéon-Sorbonne - France
Perrine THURINGER ARD Intelligence des Patrimoines, Centre for Advanced Studies in the Renaissance (CESR) / CNRS, University of Tours - France	**Miled ROUSSET** House of the Orient and the Mediterranean (MOM) / CNRS, University Lumière Lyon 2 - France
Damien VURPILLOT ARD Intelligence des Patrimoines, Centre for Advanced Studies in the Renaissance (CESR) / CNRS, University of Tours - France	



Introduction

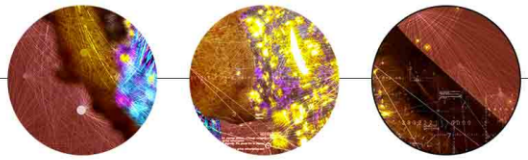
A plethora of data about cultural and natural heritage is nowadays available to both public and private institutions. These data are highly heterogeneous in terms of both formats and contents. In addition, the way in which they are organized and characterized depends on the socio-cultural contexts of different working communities, their research methodologies, languages, and ways of thinking. As a consequence of this heterogeneity, it is hard to find connections across multiple datasets or to agree on data publishing policies and shared vocabularies to describe the data in a common manner.

Researchers and stakeholders challenge this situation by relying on theories, methodologies, and technologies developed in Computational Linguistics, Conceptual modeling, Data Science, Artificial Intelligence, Knowledge Representation and Reasoning. These research efforts are characterized by a strong interdisciplinary nature and their methods are nowadays largely exploited in the Social and Human Sciences (SHS), including in the Digital Humanities.

The FAIR Heritage conference aims at bringing together scholars and stakeholders working with Open Science and the FAIR (*Findable*, *Accessible*, *Interoperable*, and *Reusable*) principles in the context of heritage studies. The latter principles play indeed a decisive role to define guidelines and valuable tools for managing data and digital resources in robust ways.

We are interested in research questions addressing both *methodological* and *application* challenges emerging from data management practices including data modeling, sharing, retrieval, integration, and mining. More specifically, topics of interest include (but are not limited to):

- Sustainability of the FAIR principles for cultural/natural heritage knowledge representation and data management;
- Current social challenges for the reuse of research datasets, including policies for data reuse;
- Presentation of digital data services, libraries, and software applications for the management of SHS data, as well as data about cultural/natural heritage. We are particularly interested in the use of technologies meeting the FAIR principles such as Semantic Web and Linked Open Data (LOD) technologies;
- Research and application challenges about data mining, conceptual modeling, knowledge representation and reasoning for the SHS and cultural/natural heritage. We are also interested in research efforts at the intersection between these fields, e.g., semantic data mining for information extraction, knowledge discovery or consistency checks;
- Conceptual and formal models for managing uncertain knowledge and data;
- Lessons learned from the use of metadata vocabularies and ontologies like CIDOC-CRM and FRBR, including discussions addressing their shortcomings and proposals to overcome them;
- Methodologies, languages, and tools for mapping Semantic Web ontologies to existing datasets and databases, including Ontology-Based Data Access (**OBDA**) approaches;
- Ontologies for geo-spatial or temporal knowledge representation and data management, including their exploitation in geographic information systems;



- Semantic characterization of visual data, e.g., 3D models, BIM models, 2D+ models, point-cloud sets, etc.

Domain of interests include: architecture, archeology, book studies, history, history of art, history of technology, literature, musicology, performing arts, philosophy, etc.

Organization: The conference is organized in the scope of the ARD 2020 Programme *Intelligence des Patrimoines* and the French MASA Consortium. One of the objectives of the FAIR Heritage conference is to connect the work done in the scope of these two initiatives with similar efforts at the international scale in the field of heritage studies.

Acknowledgments: We are grateful to Damien Vurpillot, Laurence Rageot, and colleagues at the CESR, LE STUDIUM Loire Valley Institute for Advanced Studies, the MASA Consortium, and the MSH Val de Loire for their kind support for the conference organization.

Convenors

Dr Emilio Maria Sanfilippo

LE STUDIUM Research Fellow at the Center for Advanced Renaissance Studies (CESR) / French National Center for Scientific Research (CNRS), University of Tours, Intelligence des Patrimoines Programme

I got both a Bachelor and Master degree in Philosophy from the University of Catania (Italy) with a focus on philosophy of science, philosophy of language, analytic metaphysics, and logic. I eventually got a Ph.D. in Computer Science from the University of Trento (Italy) discussing a thesis on the use of formal ontologies for product design and manufacturing. Since then my research focuses on conceptual modeling and ontology engineering for knowledge representation and data management. I like interdisciplinary research at the intersection between philosophy, computer science, cognitive sciences, and engineering. This attitude has brought me to explore the Digital Humanities, including data modeling and data integration for literary scholarly editing and music.

Contact: emiliosanfilippo@gmail.com

Dr Xavier RODIER

French National Center for Scientific Research (CNRS), Director of the Maison de Sciences de l'Homme Val de Loire (MSH VdL), MASA Consortium

Xavier RODIER is a research engineer at the CNRS, holder of an *Habilitation à Diriger des Recherches*. An archaeologist working on the processing of spatial information and computerisation of archaeological process (Archeomatic), his research forms part of an interface between archaeology, geography and computer science. Director of the *Maison des Sciences de l'Homme Val de Loire* and member of the *Laboratoire Archéologie et Territoires* (UMR CITERES), he heads the consortium Mémoires des archéologues et des sites archéologiques (MASA) of the French Very Large Research Infrastructure Huma-Num, which focuses on archiving, interoperability and publication in the Linked Open Data. He is partner of the ARDIANEplus project (H2020) and the COST action Saving European Archaeology from the Digital Dark Age (SEADDA).

Contact: Xavier.rodier@univ-tours.fr

Speakers

Dr Denise ARDESI

Center for Advanced Renaissance Studies (CESR) / CNRS, University of Tours - FR

Co-author: Dr Christian Steiner

After obtaining a master's degree in modern literature at the ENS- École Normale Supérieure de Lyon, Denise Ardesi continued her doctoral studies at the CESR- Centre d'Etudes Supérieures de la Renaissance in Tours where she specialized in French literature, especially on the dialogue that took place between Jewish and Catholic mysticism in the 15th - 16th centuries. In parallel, Denise Ardesi studied digital humanities, with a particular interest for XML-TEI encoding system. She is currently a research engineer for the CoReMA- Cooking Recipes of the Middle Ages program.

Christian Steiner is a PHD student on computer science at ZIM- Zentrum für Informationsmodellierung at the University of Graz.

Contact: denise.ardesi@univ-tours.fr

Semantic coordination on Medieval cooking recipes

The ANR-FWF CoReMA- Cooking Recipes of the Middle Ages program analyzes the medieval cooking recipes according to an international standard digital methodology and it will allow the creation of an ontology related to ingredients and culinary processes from the textual data directly observed in the primary sources, i.e. in the manuscripts. This multilingual analysis should reveal the migration of texts, competing or particular food behaviors, which have built the identities and food heritage of Europe. This paper aims to explain how complex is the semantic annotation of cooking recipes. In order to facilitate the editing of this complex semantic level and produce working expressions compatible with the Linked Open Data, we use Heterotoki and its export in SKOS of our Knowledge Organization System. Moreover, the CoReMA program wants to offer a working and data processing system that is compatible with the FAIR principles, especially with reusable XML code to allow new researches with other languages, LOD, controlled vocabularies and standard ontology about food. Finally, the CoReMA project also engages in activities for the public within social networking Programme CoReMA, blog *Hypotheses Cooking Recipes of the Middle Ages*, and an online channel CoReMA - *Corpus de recette du Moyen-Âge* with webinar and kitchen recipes.

Marlène ARRUGA

Center for Advanced Renaissance Studies (CESR) / CNRS, University of Tours - FR

Marlène Arruga and Caroline Parfait are 2nd year students in a M.A. in Digital Humanities at the Centre d'Études Supérieures de la Renaissance in Tours, France. They are specialized in Data Intelligence in Culture and Heritage.

Marlène has a degree in History and carried out research on the diffusion of writtenJapanese language in the French society from the mid-19th century to the early 20th century.

Contact: marlene.arruga@gmail.com

Caroline PARFAIT

Center for Advanced Renaissance Studies (CESR) / CNRS, University of Tours - FR

Caroline has a master's degree in visual arts gained at the École Nationale Supérieure Arts de Paris Cergy and a degree in philosophy at Sorbonne Université (Paris). Her current research focuses on the study of systems and standards for digital library development.

Contact: caroline.parfait@etu.univ-tours.fr

Developing a Digital Platform on Stained Glass Restoration Data in France

ENVi was first a graduate student project. The team originally included a third student, Romain Tigero. *ENVi, Espace Numérique du Vitrail*, aims at integrating all data collected throughout the whole restoration process of a stained-glass window in a database called *Base Numérique du Vitrail* (BNV). The *ENVi* platform is a unified system that enables professionals (researchers, administrative, and technical staff) to both work in a collaborative way and visualize the dataset on the *ENVi* web site. The data model was developed from the analysis of condition reports consulted at *Médiathèque de l'Architecture et du Patrimoine* (MAP).

This dataset is not homogeneous because of the variety of:

- The formats and media used for collecting and preserving data on stained glass and its restoration;
- The people involved in the field;
- The working process on paper documents (reports of past, current or future restorations).

We focused on establishing best practices for formatting data and metadata as well as linking the objects described to their related identifiers MAP and Joconde.

On a long-term basis, the BNV database aims at becoming a useful tool used for a scholarly data analysis by researchers who study the history and chronology of this material; in this respect, we plan to use Dublin Core and CIDOC-CRM to integrate our work within a Semantic Web approach.

In the future, our team wishes to work with governmental authorities in order to obtain agreement to make our data downloadable and reusable for the community of stained glass researchers and technicians. We are currently working on publishing the prototype of our website.

Prof. Roland BILLEN

Geomatics Unit, University of Liège - BL

Roland Billen holds a PhD degree in Geomatics from the University of Liege. He was a research fellow of the Belgian National Scientific Foundation and then became Lecturer at the department of Geographical and Earth Sciences of the University of Glasgow. He is now Full Professor at the department of Geography of the University of Liege.

He is active in several fields of spatial information sciences such as data acquisition techniques (Surveying, Laserscanning…), Geographical Information System (GIS) and qualitative spatial reasoning. He leads a research group specialised in 3D data acquisition and 3D GIS, including a track on 3D visualisation.

Contact: rbillen@uliege.be

The Coudenberg Heritage Building Information System

The Coudenberg archaeological site presents the remains of the palace of the Dukes of Brabant, destroyed by a fire in 1731 during the reign of Marie-Elisabeth of Austria, the Governess of the Netherlands. In 2018, the Geomatics Unit of the University of Liege has started the development of a 3D Heritage Building Information System based on the specifications established by the archaeologists, the site managers, the exhibition office, etc. It is composed of several modules (SQL and NoSQL database management systems / GIS / 3D point cloud viewer and editor / web Interfaces). The 3D data of the sites were acquired by a combination of laserscanning, photogrammetric and land surveying techniques. The system, currently in use, is going to be upgraded during the summer 2020.

The presentation will start by sketching the design and modeling approach (meta-model, general information system architecture). Then, the ad-hoc Coudenberg Heritage Building Information System will be detailed (conceptual data model / prototype / data acquisition procedures). Next, examples of uses of the system will be presented through demo videos (3D navigation / object creation / annotation and interaction / …). The presentation will end with some research and development perspectives.

Dr George BRUSEKER

Takin.solutions - BG

George Bruseker is Vice Chair of the CIDOC CRM Special Interest Group and an expert in the field of knowledge representation and semantics for cultural heritage. He has a PhD in Philosophy from the National and Kapodistrian University of Athens and has worked in diverse institutions within the field of cultural heritage and heritage research over the past fifteen years. He is deeply involved in the semantics community, participating in several major initiatives aiming to facilitate the adoption of semantic data including Data for History (d4h), Linked.Art, and the Arches Resource modeling Working Group (ARM WG). His research interests include the pedagogy of knowledge representation, semantic data workflow management and the modeling of art and architectural historical data.

Contact: george.bruseker@gmail.com

Realizing the Virtues (and Combating the Vices) of the Digital Shift in Heritage Scholarship: the Role of Knowledge Representation

The generalized adoption of digital tools and the ubiquitous creation of data in the heritage research community reflects an overall consensus trend within society. Digital methods have become the horizon of production and exchange of information. These methods have become a norm under the promise of gains in the quality of scholarly communication, equity, access, knowledge and simplicity. In a naive vision, these qualities are brought about simply by virtue of 'going digital'. In reality, such promises are not so easily paid out; they can only actually be realized through coordinated, community effort amongst scholars to achieve them. The movement towards FAIR data is a major step in this direction. Within such movements, the practice of knowledge representation (KR) has a key role to play. Community creation, maintenance and adoption of formal ontologies that can truly be used as a lingua franca both between scholars and computers but also amongst scholars lies at the core of supporting the mission of creating intelligible and sustainable data informing scholarly knowledge processes. This talk will discuss the role of KR in realizing FAIR heritage data. It will look at the basic achievements that are in place for the heritage semantics community, with a focus on the use of CIDOC CRM. It will also highlight the challenges, especially in KR, that lay ahead in bringing about the promised virtues of the digital shift and combating its incipient vices.

Dr Cécile CALLOU

National Museum of Natural History (MNHN) / CNRS, Sorbonne University - FR

Co-author : Isabelle BALY

Maître de conférences du Muséum national d'Histoire naturelle, archéozoologue spécialiste des périodes historiques dans l'unité de recherche « Archéozoologie, archéobotanique : sociétés, pratiques et environnements (AASPE) » (UMR 7209 CNRS/MNHN), Cécile Callou est également responsable de collection en charge des spécimens types de mammifères et de l'administration scientifique de la base de données des collections Zoologie Mammifères et Oiseaux et Anatomie comparée depuis 2004.

Depuis 2011, date de sa création, Cécile Callou est directrice de l'unité mixte de service à la recherche « Bases de données sur la Biodiversité, Écologie, Environnement et Sociétés (BBEES) » (UMS3468 CNRS/MNHN) et, à ce titre, impliquée dans de nombreux projets informatiques/numériques.

Contact: cecile.callou@mnhn.fr

Du projet d'inventaire archéologique de la faune et de la flore française vers un portail bioarchéologique global

Depuis sa création en 1993, la base de données nationale « Inventaires archéozoologiques et archéobotaniques de France (I2AF) » a connu de nombreuses évolutions au gré des projets et des programmes de recherches. D'abord « Atlas diachroniques des mammifères de France », puis « Inventaire des données archéozoologiques holocènes de la moitié nord de la France métropolitaine », la base contient aujourd'hui les données archéozoologiques et archéobotaniques françaises, de Métropole et d'Outre-Mer, pour une période pouvant débuter au Paléolithique supérieur.

Dans la perspective de mieux répondre aux exigences du plan national pour la science ouverte, un important travail de refonte a été entrepris. Ce travail s'attache à la fois à la technologie de la base (d'un format propriétaire vers un format open source), à l'amélioration de l'ergonomie de l'interface, avec pour objectif de faciliter la consultation et l'intégration de nouvelles données, aussi bien qu'à un élargissement possible de la couverture géographique et chronologique.

Dr Livio DE LUCA

Models and Simulations for Architecture and Heritage / CNRS - FR

Architect, PhD in Engineering (Arts et Métiers ParisTech), HDR in Computer Science, Livio De Luca is research director at CNRS and director of CNRS-MAP unit. General Co-chair of the UNESCO/IEEE/EG DigitalHeritage international congress (Marseille 2013, Grenade 2015) and coordinator and member of national and international actions, his research activities focus on surveying, geometric modeling and semantic enrichment of digital representations of heritage objects. Associate editor of the Journal of Cultural Heritage and the Journal on Computing and Cultural Heritage, since 2016 he is an appointed member of the CoNRS. His work was rewarded in 2007 by the Pierre Bézier Prize (Arts et Métiers Foundation), in 2016 by the Medal for Research and Technology (french Academy of Architecture) and in 2019 by the CNRS Medal of Innovation. He is today coordinator of the "digital data" working group of the CNRS/Ministry of Culture scientific site for the restoration of Notre-Dame de Paris.

Contact: **livio.deluca@map.cnrs.fr**

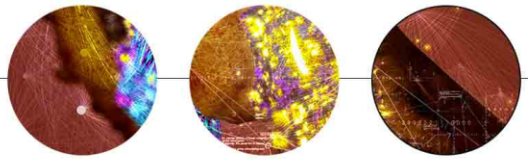
A digital ecosystem to document, in space and time, the restoration of Notre-Dame de Paris

The restoration of Notre-Dame de Paris is today an unprecedented opportunity to gather and analyze the many analytical and documentary resources produced by a large number of scientists and heritage professionals from different backgrounds on the same building.

In collaboration with the other working groups of the french CNRS (National Centre for Scientific Research) / MC (Ministry of Culture) task force on the restoration of the cathedral (stone, stained glass, wood, metals, structure, acoustics, heritage emotions, etc.), the "digital data" team will work on the introduction of an innovative ecosystem to bring together, analyze and correlate the multiple levels of reading that converge towards the production of new knowledge on the cathedral, its transformations over time, its materials, the alterations produced by the fire, as well as towards the restoration project.

By combining recent advances in knowledge engineering, computer vision and shape analysis, our approach will take into account two complementary aspects. On the one hand, the description of the different steps taken by scientists to move from the observation of raw data to their interpretation in relation to other contextual data from the analysis of typical operating chains. On the other hand, the analysis of the spatial, temporal and semantic overlap of regions of "multidisciplinary interest" based on the correlation of annotations, vocabulary terms, qualitative attributes and morphological features.

Beyond the introduction of new categorization and query functionalities within a unique data centralization and collaborative research platform, the analysis of data from such a plurality of perspectives involved (disciplinary fields, know-hows, interests, ...) opens up stimulating perspectives for the construction of a sort of "digital memory" of a collective adventure.



Djibril DIARRA

Distributed Knowledge and Artificial Intelligence (CIAD) / University of Bourgogne Franche-Comté - FR

PhD Student at the laboratory CIAD (Connaissances et Intelligence Articielle Distribuées, in french) at the University of Burgundy Franche Comté (UBFC) in France. Assistant researcher at the University of Social and Management Sciences of Bamako (USSGB) in Mali. My research domains of interest are: Knowledge Management, Artificial Intelligence, Big Data and Education's Technologies.

Contact: **dl4djibril@gmail.com**

Hybrid Artificial Intelligence for Medieval Illuminations Analysis

The preservation of cultural heritage requires a twofold objective. The first is the physical preservation of the cultural object. The second is the preservation of the meaning carried by that object. This meaning is often associated with a context which disappears over the time and then the exact meaning carried by the object becomes corrupted [2, 4]. This last can be reduced by a consensual and formal representation of the content's meaning. The goal is to allow an automatic detection of the content based on that formal representation in order to reinforce the popular understanding of the cultural collections [3, 4]. To achieve that, we constructed an hybrid Artificial Intelligence (AI) system combining both symbolic and connectionist approaches. The symbolic approach is a core ontology [1]. The question treated in this presentation is:

How Artificial Intelligence helps the medievist in the expression and the preservation of the meaning of cultural object (here the medieval illumination) according to its time's context?

The aforementioned ontology is dedicated to represent the context of influence in the medieval illumination. The specification of its terms is published online, at the URL *http://medievenl.ontology.checksem.fr*. Some mechanisms of extension are built in the system so that to improve the taxonomy during an annotation step of the illumination. The medievist combines selected graphical symbols and semantics (classes and properties) using the ontology. That annotation step trains a machine learning algorithm [5] to identify graphical items in the illumination. The users would be then able to understand better the cultural painting and to improve its annotation. This machine learning method constitutes the connectionist approach of our hybrid Artificial Intelligence system. This last is a Web application whose the URL is *http://app.illumination.checksem.fr/home*.

References:

[1] Diarra, D., Belkaroui, R., Clouzot, M., Nicolle, C.: Illumination3.0: A Semantic Annotation Platform Based on Ontology for Medieval Illuminations. In: Proceedings of 30ème Journées Francophones d'Ingénierie des Connaissances, IC @ PFIA 2019. pp. 231–234. AFIA, Toulouse, France (2019), https://hal.archives-ouvertes.fr/hal-02329636/file/djibrilD.pdf

[2] Doerr, M., Hunter, J., Lagoze, C.: Towards a core ontology for information integration. Journal of Digital information 4(1) (2006)

[3] Dörr, M.: The cidoc crm-an ontological approach to semantic interoperability of metadata, 2001. AI Magazine, Special Issue on Ontologies, Nov (2002)

[4] Hyvönen, E.: Publishing and using cultural heritage linked data on the semantic web. Synthesis Lectures on the Semantic Web: Theory and Technology 2(1), 1–159 (2012)

[5] Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. pp. 91–99 (2015)

DIVE_Team

Meriç AKDOGAN, Pauline BELLEMERE, Inoussa KORA CHABI

Center for Advanced Renaissance Studies (CESR) / CNRS, University of Tours - FR

The DiVE Team, founded in September 2019, is composed by three students in Digital Humanities of the Centre d'Études Supérieures de la Renaissance, Inoussa Kora Chabi, Pauline Bellemère and Meriç Akdogan. The DiVE project, under the supervision of Dr Marion Lamé, is our main student project during the first year of our M.A. *Sciences humaines et sociales MENTION Humanités numériques Parcours Médiation numérique de la culture et des patrimoines 2019-2021*.

Contact: **meric.akdogan@etu.univ-tours.fr**; **pauline.bellemere@etu-univ-tours.fr** ; **inoussa.korachabi@etu.univ-tours**

DiVE Project

The DiVE platform creates participative, virtual and multicultural exhibitions on 3D environment. DiVE started with the virtual exhibition Mon Sanitas created by our partner, the association Espace Passerelle, which seeks to share neighborhood perceptions through Tours's contemporary photographs. Usually, online exhibitions follow a top-down process: a team organizes an exhibition and provides this exhibition online. However, the DiVE platform encourages its users to create their own virtual exhibition, or to participate in other user's exhibitions in a secured environment where they meet offering a shared experience of their own photographs, videos, comments and even visits.

The upcoming exhibition of the MSH Centre Val de Loire, for the Covid-19 lockdown, inspired us for some logistic aspects and for DiVE's website structure. Our main work is the modeling of the exhibition process according to our criteria: participative, non-elitist, based on open access tools. For this reason, we would like to experiment with the prototype for modeling of our 3D platform by using Blender, a free and open-source tool. Unlike other 3D design software, Blender offers users a rich palette of functions for 3D environment construction. It can support several types of graphic and audiovisual formats. The files produced by Blender could also be readable by other 3D design software which facilitates the interoperability of the datasets and allows its reuse by others.

Dr Wieslawa DUZY

Institute of History, Polish Academy of Sciences - PL

Researcher (PhD) at the Institute of History, Polish Academy of Sciences. MA in history (2007), BA in sociology (2009), and PhD in history (2013). In 2010-2014 was working at Polish Statistical Office. In 2014-2015 was working at the Silesian University in Opava, Czech Republic. Currently working on Historical Atlas of Poland in 16th century and reconstruction of a settlement network in Warmia (Ermland). Fellow of the Herder Institute for Historical Research on East Central Europe - Institute of the Leibniz Association (2018).

Scientific interests: social history in 18th and 19th century – Spatio-temporal ontologies and databases – History of Ermland.

Contact: wieslawa.duzy@gmail.com

Historical Ontology of Urban Space

The Tadeusz Manteuffel Institute of History, Polish Academy of Sciences elaborates a project "Historical Ontology of Urban Space" (HOUSe) together with 9 international partners. Our goal is to elaborate an historical ontology suitable for European towns in the past. We are working on changes occurring in urban space from the 16th century to the beginning of the 20th century along with permanent elements that we think can constitute the town. We use various historical sources, including maps and city plans. The data collected in the project will refer to the historical space of Warsaw, but we will use administrative boundaries of present city to delineate this space. In other words: we are interested in all elements / objects that were located in the past in the area enclosed by the borders of Warsaw in the 21st century. We want the results of our work to be applied when working on a series of historic atlases of towns conducted under the auspices of the International Commission on Urban History. Our goal is to build an IT system, including a database and an ontology, allowing different teams dealing with town atlases to work in a uniform pattern, to enable future exchange of data sets and to improve data harmonization and comparative research.

NAWA
POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

More information about the ongoing project, including the data model, project team, and project partners: https://urbanonto.ihpan.edu.pl/ The project is financed by the Polish National Agency for Academic Exchange. Project social media: #urbanonto.

Dr Cesar GONZALEZ-PEREZ

Institute of Heritage Sciences, Spanish National Research Council (CSIC) - ES

Cesar has degrees in Biology, Electronics, Computing, and Geography and History. Currently, he is a Staff Scientist with Incipit CSIC, where he leads research in software engineering and cultural heritage aiming to develop theories, methodologies, and technologies to understand and assist knowledge generation in cultural heritage. Previously, he has worked in the areas of conceptual modeling, metamodeling, situational method engineering, and software development, with a special focus on heritage. He is also a member of UNE, the Spanish body represented at ISO, and has led 3 international standardization projects. Cesar is also an evaluator for public research agencies from 6 countries, has been a partner of 3 technology-based companies, and is the author of over 100 publications.

Contact: **cesar.gonzalez-perez@incipit.csic.es**

Ontologies of Cultural Heritage for Humans and Machines: The Cultural Heritage Abstract Reference Model

Since the mid-20th century, two research strategies have been struggling to represent the world. On the one hand, Artificial Intelligence developed ontologies, with a focus on allowing computers to carry out automated reasoning about facts of the world. On the other hand, Software Engineering came up with conceptual modeling, with a focus on capturing the nuances of the world to help communication between people. Both lines of work have been applied to cultural heritage with varying fortune.

Recent trends demonstrate that ontologies and conceptual models are not that different, and that combined approaches have advantages over either of them. Following this premise, we developed CHARM, the Cultural Heritage Abstract Reference Model (www.charminfo.org), an ontology of cultural heritage expressed in ConML (www.conml.org), a conceptual modeling language capable of describing "soft" issues such as subjectivity, temporality or vagueness. CHARM and ConML combine straightforward affordability to computing non-experts with a high level of formality for automated processing.

In this talk I will describe CHARM and the major areas covered by it. I will also explain how we can easily capture complex cultural heritage phenomena by using ConML built-in features, and how these two technologies together can allow you to standardise the documentation of cultural heritage while maintaining a high degree of specificity, something that is often impossible when adopting off-the-shelf standards.

Florian HIVERT

Center for Advanced Renaissance Studies (CESR) / CNRS, University of Tours - FR

Student in the M.A program *Intelligence des données de la culture du patrimoine at the Centre d'Études Supérieures de la Renaissances* during the years 2018-2020 and soon to be graduated, I worked during my degree on *ROSER*, a digital scholarly edition in order to learn the main principles of a Digital Humanities project.

ROSER initially started as a student project in 2018-19, led by Théo Roulet, Marie Petit and myself. This year, ROSER faced a deep restructuration on two specific aspects during my internship: (1) gateway with ARVIVA (2) Linked Open Data of its thesaurus. *ROSER* received the support of Intelligence des Patrimoines (Centre d'Étude de la Renaissance) and the *Projet Chartres (ARVIVA)*. The internship was under the supervision of Marion Boudon-Machuel, supervised by Marion Lamé and François Rosmann.

Contact: **florian.hivert@etu.univ-tours.fr**

ROSER - An interoperable scholarly digital edition for the ornamentation in Renaissance edifice

ROSER aims at offering a scholarly digital edition based on the corpus of two researchers in Art History, led by Jean Beuvier and Sarah Munoz. These corpora harvested a large amount of visual data on the Renaissance building ornamentation. The main part of the work focused on defining a meaningful data model for this specific type of cultural heritage object, to build up the interoperability with similar projects (*ARVIVA*) and publishing the dataset in a machine-friendly format to follow the FAIR principles.

The **CMS** Omeka-S was chosen because of its capacity to build up **RDF** triplets for the integrated data. Omeka-S builds his data models classes and properties only with ontologies. The ontologies used in *ROSER* can be standard ones, like *FOAF* or DCterms, and new ones, like the *ROSER* ontology created for the specific subject of this scholarly digital edition.

ROSER interoperability has already been tested with *ARVIVA* in the form of a gateway. The gateway matches the ontological properties of *ROSER* with the *ARVIVA* **SQL** database attributes and then it proceeds to transform ROSER items to the *ARVIVA* item structure.

To push further interoperability, *ROSER* uses *Heterotoki* to edit the working expressions of scholars. *Heterotoki* allows indeed researchers to define the use of a term (in an index or a glossary), map the term to controlled vocabularies like *PACTOLS* and *IconClass*, and construct a complex and enriched KOS based on *SKOS*.

Prof. Ludger JANSEN

University of Passau & University of Rostock - DE

Ludger Jansen holds a PhD in Philosophy, and teaches at the Universities of Passau and Rostock. His area of expertise is Metaphysics, including its history in Ancient Philosophy and contemporary analytic and applied approaches in ontology. His has written books on Aristotle's theory of dispositions (2nd ed. 2016), and on the ontology of groups and institutions (2017). Together with Barry Smith, he edited the first German-language introduction to Applied Ontology (2008). A complete list of publications can be found at http://purl.org/jansen/publ.

Contact: **ludger.jansen@uni-rostock.de**

Why DH ontologies need standards too

Data semantics and ontologies become more and more important in Digital Humanities. In this, DH are following other research area, where ontologies are much further developed. In this presentation, I will look in particular at the biomedical domain and argue that DH can learn a lot from the experience with ontology development in this area and thus avoid a lot of mistakes which have been made there some decades ago. DH communities should strive for shared standards and a network of collaboratively developed orthogonal reference ontologies.

Cf. L. Jansen, Ontologies for the Digital Humanities: Learning from the Life Sciences?, 1st International Workshop on Ontologies for Digital Humanities and their Social Analysis (WODHSA) within JOWO 2019, CEUR workshop proceedings, vol.2518, 2019.

Dr Florent LAROCHE

École Centrale of Nantes - FR

Florent Laroche is a "doctor-engineer" working as an assistant professor at Ecole Centrale de Nantes (France) and as a researcher in the laboratory IRCCyN (Research Institute for Communication and Cybernetics of Nantes, France - UMR CNRS 6597). He works on the translation of knowledge of the past in contemporary knowledge, readable and understandable in the present socio-technical system. One objective of this research is to reuse the technical heritage as a basis for innovation. His research topics are KM, PLM, information system modeling, interoperability, enterprise modeling, virtual engineering, reverse engineering. So as to, he used his engineering background for cultural heritage using digital tools to preserve and valorise old knowledge. He is expert for Museums and ICOMOS.

For more information: http://www.florentlaroche.net

Contact: **florent.laroche@ec-nantes.fr**

French National Project ReSeed - Semantic reverse-engineering of digital heritage objects

This project fits in the field of reverse engineering and meets the requirements and specifications of cultural heritage. The scientific goals are in line with the 7th "ANR Defi", aiming to build our future society based on information and communication. Nowadays, reverse engineering is widely used in manufacturing industries in order to record geometrical data from the products. However, digitizing physical data does not provide any knowledge about the products. Capitalizing this know-how and knowledge is critical to advance tomorrow's products. So we have identified two approaches:

- Classical knowledge management based on semantic sources
- Computer assisted design and scanning tools

Yet, there still does not exist any global process: technical issues born from interdisciplinary processes drive us to scientific issues: there is an important need to bring the different models into line and to carry out a case-by-case review of the usual industrial processes.

The ReSeed project aims for the development of a new technology: a tool and an interoperable format in order to digitize both historical semantic data and 3D physical objects. ReSeed will implement an ethical code assigned to guarantee the authenticity and uniqueness of the future semantically augmented numerical objects.

As the ANR project is interdisciplinary by nature, the composed consortium consists of:

- 4 research laboratories: 3 of science and 1 of humanities
- 2 companies: 1 of engineering and 1 of humanities
- 2 public bodies - providing use cases at a national scale, in scientific and technical fields.

The project is also supported by the Huma-Num TGIR (Large Scale French Research Infrastructure).

More information on the website of the project: www.reseed.fr

Dr Carlo MEGHINI

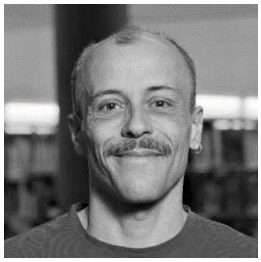
Institute of Science and Information Technology "A. Faedo", National Council of Research (CNR) - IT

Dr Carlo Meghini (M) graduated in Computer Science at the University of Pisa in 1979 and since 1984 works as a researcher at ISTI CNR. He is a member of the Networked Multimedia Information System laboratory, where he leads the Digital Library group (http://nemis.isti.cnr.it/groups/digital-libraries). His areas of interest are digital libraries and semantic technologies. He is involved in European projects since 1986, in the areas of Information Retrieval, Digital Libraries and Digital Preservation. From 2007 to 2014 he has been involved in the making of Europeana, the European digital library (www.europeana.eu), taking care of the scientific aspects of the project. He has coordinated the Coordination Action PRELIDA on the Preservation of Linked Data. He has been invited to give the Susan Hockey Lecture in 2018. He is member of the editorial board of the ACM Journal on Computing and Cultural Heritage.

Contact: **Carlo.Meghini@isti.cnr.it**

Extending Digital Libraries with Digital Narratives

Digital Libraries (DLs) abound with narratives, in the sense that every digital object in a DL tells some kind of story, regardless of the medium, the genre, or the type of the object. This is especially true for DLs in the Cultural Heritage domain. However, there is no track of narratives in the services offered by today's DLs. It is not possible, e.g., to discover a narrative, or to create one, or to compare two narratives. Of course, any DL offers a discovery service over its content; but this service addresses the objects that carry the narratives, whether books, audio-visual messages and the like; narratives per sé are not addressed. It maybe said, in short, that DLs ignore their contents. The talk will present an ontology for narratives, offering a two-level representation: a syntactical representation, constituted by the *narration* of the author, which may be a text or any other media object; and a *semantical* representation, constituted by a semantic network using the terms of the ontology and providing a formal complement to the narration. In narratological terms, the latter would be a formal representation of the *fabula* of the narrative. We discuss in some detail the ontology and present its application in the Mingei project, chartered at the representation and preservation of crafts.



Dr Alessandro MOSCA

Faculty of Computer Science, Free University of Bozen-Bolzano - IT

KRDB Research Centre of the Faculty of Computer Science, Free University of Bozen-Bolzano (Italy). His research activity focuses on knowledge representation (KR) and conceptual modeling for data management. His career has been always characterized by a strong degree of interdisciplinarity: he likes to build linguistic and conceptual bridges between different disciplines. This attitude gave him the opportunity to work with domain knowledge corpora by chemical engineers, historians, policy makers, artefact designers, among others. The idea of using the principles and KR technologies to allow the cross-domain interchange of meanings is surely one of the most relevant driving engine of his research activity.

Contact: **alessandro.mosca@unibz.it**

Ontology-based data integration in EPNet

Semantic technologies are rapidly changing the way historical research is digitally performed. Over the last decades, in social sciences and humanities, an immense amount of new quantifiable data have been accumulated, and made available in interchangeable formats, opening up new possibilities for solving old questions and posing new ones. This talk introduces a framework that eases the access of scholars to historical and cultural data about food production and commercial trade system during the Roman Empire, distributed across different data sources. The proposed approach relies on the Ontology-Based Data Access (OBDA) paradigm, where the different datasets are virtually integrated by a conceptual layer (also known as, an ontology) that provides to the user a clear point of access and a unified and unambiguous conceptual view. The framework has been implemented in the context of the EPNet project, "Production and distribution of food during the Roman Empire: Economics and political dynamics" (FP7-IDEAS-ERC) and it currently represent the backbone of the online "Roman open Data" (https://romanopendata.eu/) portal.

Blandine NOUVEL

Centre Camille Jullian (CCJ) & Mediterranean House of Human Sciences & GdS Frantiq / CNRS, University Aix-Marseille - FR

Blandine Nouvel learned archaeology and information sciences. After several experiences in libraries at technical and cultural organizations, she is now responsible for engineering and documentary resources at the Centre Camille Jullian, research laboratory in archeology in Aix-en-Provence (CNRS-Aix Marseille University). Convinced of the need to widen access to scientific information, she is particularly interested in the questions of standardization and structuring of the metadata that she implements in projects to publish research data. This led her to launch the reorganization of the PACTOLS thesaurus developed by the Federation and resources on Antiquity, whose technical characteristics and linguistic potential made it a pivotal vocabulary for archeology.

Contact: **nouvel@mmsh.univ-aix.fr**

Deconstructing for reconstructing: the use of the BBT for reorganising the PACTOLS thesaurus

The aim of the talk is to present the methodology used to reorganise the PACTOLS thesaurus of Frantiq, launched within the framework of the MASA consortium. PACTOLS is a multilingual and open repository about archaeology from Prehistory to the present and for Classics. It is organized into six micro-thesaurus at the root of its name (Peuples, Anthroponymes, Chronologie, Toponymes, Oeuvres, Lieux, Sujets). The goal is to turn it into a tool interoperable with information systems beyond its original documentary purpose, and usable by archaeologists as a repository for managing scientific data. During the talk, we will describe the choice of tools, the organisation of work within the steering group and the collaborations with specialists for the upgrading and development of the vocabulary while showing the strengths and limitations of some experiments. Above all it will show how the introduction of the conceptual categories of the BackBone Thesaurus of DARIAH, modelled on the CIDOC-CRM ontology, through a progressive deconstruction/reconstruction process, eventually had an impact on all micro thesauri and questioned the organisation of knowledge so far proposed.

Dr Tomasz PANECKI

Institute of History, Polish Academy of Sciences - PL

Tomasz Panecki, PhD – assistant professor at the Department of Historical Atlas (Institute of History, Polish Academy of Sciences). Obtained master degree in history and geography and defended doctoral thesis at the Faculty of Geography and Regional Studies (University of Warsaw) entitled "The Concept of Historical Topographic Objects' Database". Scientific interests include: old maps digital editions & historical mapping including history of cartography, spatiotemporal ontologies and databases, conceptualization of place and its categories in diachronic context, cartographic representation of real world entities.

Contact: **tpanecki@uw.edu.pl**

Digital editions of historical maps: between images and data. The example of "Gaul/Raczyński" topographic map (1807-1812)

The aim of the author is to present and discuss methodological problems related to the development of old maps' digital editions on the example of the "Gaul/Raczyński" map. The map consists of 8 sheets covering 8 districts of Poznań Department of the Great Duchy of Warsaw and its scale is approx. 1:125 000 with elaborations dates 1807-1812 and its scope of content is rather typical for maps of this period. This manuscript map was colorful, although what remains today are only its black-and-white photocopies as the original was lost during the Second World War.

Edition is here understood both as a process of providing access to the map and data derived from it as well as the final result of the work. Editing is a "critical representation of a historical document", so all actions should be scientifically justified and the edition should include the representation of the map. It also requires detailed studies in terms of map's scope of content and circumstances of its elaboration as well as digitization process including georeferencing and vectorisation. Finally, editing also means visualizing data (e.g. in webGIS application) and organizing spatial knowledge as well as maintaining the data.

The "Gaul/Raczyński" map edition is comprised of six components: map images, georeferenced and mosaicked sheets, map with reconstructed colors, spatial database, webGIS application and documentation. It is available on the website.

Prof. Elena PIERAZZO

Center for Advanced Renaissance Studies (CESR) / CNRS, University of Tours – FR

Elena Pierazzo is Professor of Digital Humanities at Centre d'Études Superieures de la Renaissance, University of Tours where she directs the masters in Digital Humanities. She has a PhD in Italian Philology: her specialisms are Italian Renaissance, digital edition of manuscripts, digital editing, and text encoding. Her most recent publication is Digital Scholarly Editing: Theories, Models and Methods (2015). She has been the Chair of the Text Encoding Initiative (TEI) and involved in the TEI user community, with a special interest in modern and medieval manuscripts. Co-chair of the Programme Committee of the DH2019 conference, she was the co-chairs the working group on digital editions of the European Network NeDiMAH and was one of the scientists-in-chief for the ITN DiXiT.

Contact: **elena.pierazzo@univ-tours.fr**

The edition of draft manuscripts: theories, (data) models and methods

Editing drafts and authorial manuscripts more generally is a difficult task, whether for publication on paper or on screen, and this is because manuscripts themselves are complex, data-rich objects. In fact, draft manuscripts represent in a bi-dimensional space the multi-dimensional activity that is authoring, which in turn can be broken down into many complex processes conducted over time which are only partially recorded on the space of the page. Digital editions seem to be the only way to capture at least some of these aspects, but in order to tackle the editorial problems connected to drafts, digital editors need to consider these manifold issues, separate them, and then classify and model them for both the computer and the end user. The digital editions produced so far are all somewhat unsatisfactory in these respects, as they are either too difficult to use, or they oversimplify the research questions. Centred on the concept of "representation", this paper will approach the open problem of the edition of authorial manuscripts from three points of view: 1. The physical document and its virtual representation; 2. The representation of the underlying creative process; 3. The representation of the evolution of the concept of work. For all these approaches, existing data models and their standardisation will be examined, as well as the ways these data could and should be exploited in order to provide much deeper insights into the authorial process and its cultural relevance.

Prof. Julian D. Richards

Department of Archaeology, University of York - UK

Julian Richards is Professor of Archaeology at the University of York. He is Director of the Centre for Digital Heritage, and founding Director of the Archaeology Data Service, and the e-journal Internet Archaeology (both established in 1996). His direct involvement in archaeological computing began in 1980 when he started his PhD research studying pre-Christian Anglo-Saxon burial ritual using the computing power of an ICL mainframe and an early Z80 micro-computer. In 1985 he co-authored the first textbook in archaeological computing for Cambridge University Press, and has subsequently written numerous papers and edited a number of books on the applications of information technology in archaeology. He is currently Deputy Coordinator for the European ARIADNE e-infrastructure for archaeological research, and Action Chair for the COST Action SEADDA: Saving European Archaeology from the Digital Dark Age.

Contact: **julian.richards@york.ac.uk**

Making Archaeology FAIR: in the Archaeology Data Service and ARIADNE

The Archaeology Data Service has been promoting FAIR data in the UK heritage sector since 1996, three years before Google was founded. We are the only accredited digital repository for heritage data in the UK, and one of only five UK repositories which has been awarded the Core Trust Seal. We preserve and disseminate over 1.3 million records for the archaeology of the UK, over 60,000 unpublished fieldwork reports, and over 1800 project archives, all available under an open licence and with permanent digital object identifiers. This presentation will discuss our approach to making Archaeology FAIR, and some of the solutions, and difficulties we face. In particular, the R of FAIR is the most challenging, as how do we best encourage and facilitate data Re-use? Providing full metadata, following data standards, and applying controlled vocabularies is only a partial answer, as the lack of adequate paradata is frequently a further obstacle. I will also introduce some of the solutions to European interoperability being implemented within the ARIADNE e-infrastructure, and discuss the roles played by the CIDOC-CRM, ARIADNE's AO-Cat ontology, the Getty Art and Architecture Thesaurus, and Periodo, all of which provide building blocks for FAIR European cultural heritage.

Miled Rousset

House of the Orient and the Mediterranean (MOM) / CNRS, University Lumière Lyon 2 - FR

Responsible for the WST (Semantic Web & Thesauri) technological platform of the national network of Maisons des Sciences de l'Homme (RnMSH).

With 20 years of experience in the field of semantic indexing, management, data fusion and dissemination of document collections, I am the "IT Director of GDS3378 (Groupement de Services) of the FRANTIQ network (Fédération et Ressources sur l'ANTIQuité)".

Contact: **miled.rousset@mom.fr**

OpenTheso and PACTOLS

After 2 years of preparation and development on the new version of Opentheso, I present you the result of this work.
We called upon a professional designer for the redesign of the graphical interface. The design of the software was also revised, several partners were involved in the project (Frantiq, MASA, Huma-Num).
A stable version of Opentheso will soon be released in open-source on GitHub.

Dr Ana Roxin

University of Burgundy (UBFC), Computer Science Laboratory of Burgundy (LIB) - FR

Ana Roxin is Associate Professor in Computer Science, habilitated to conduct research, at the University of Burgundy. Her research interests address specifying descriptions that allow computers to simulate human understanding, reasoning and problem solving. She has proven experience across Semantic Web, Linked Data, rule-based reasoning, semantic interoperability, BIM, GIS and explainable AI. She sits on national, European and international standardisation bodies e.g. AFNOR, CEN, ISO. She is a member of buildingSMART International, and of the W3C Linked Building Data Community Group. Involved in the program committees of several international conferences (ISWC, ESWC, IEEE SITIS, EC3), she is regularly solicited as an expert evaluator for H2020 proposal evaluations and project monitoring.

Contact: **ana-maria.roxin@ubfc.fr**

Semantic-based BIM Modeling: Approaches and related Standardization Actions

The talk will address how ontologies rule-based reasoning approaches can help BIM-related activities and projects. The first part of the talk will present a brief overview of the concepts needed to best understand the approaches depicted in the rest of the talk:

a) what is BIM and main related standards that apply e.g. Information Delivery Manuals (IDM), Model-View Definition (MVD) and Industrial Foundation Classes (IFC).
b) what is ontology design and rule-based reasoning, and how it compares to (semi-)automatic learning approaches e.g. Machine Learning, Deep Learning.

Following this short introduction, the second part of the talk will present several approaches, as conceived and implemented through research and industrial projects: semantic checking of compliance rules, IFC view extraction, MVD adaptation through semantic rules, etc.

The third part of the talk will discuss existing, ongoing and future standardization activities making use of Semantic Web technologies, in the BIM domain. Namely the Semantic and Modeling Linking guide discussed at CEN along with the ISO Information Container for Data Drop will be presented.

Dr Stéphanie SATRE

Centre Camille Jullian - FR

Stéphanie Satre is an engineer for the CNRS working on sources' analysis and database conception. Member of the Centre Camille Jullian since 2010. She is co-responsible of the transverse axis: Research data, corpuses and databases and member of the Digital Humanities team. She is in charge of data curation (especially databases) and web dissemination. Her research activities focus on roman Sculpture. Online projects:

- SculptuRo (beta version): database on Roman sculpture
- Epicherchell: collaborative tool for the management and study of antic inscriptions of Cherchell.
- CoReA: digital library of the laboratory archives. Project conducted in partnership with INIST and Frantiq.
- AreAr: Archival treatment of archaeological documentation.

Contact: **satre@mmsh.univ-aix.fr**

Epicherchell, an epigraphic dataset to share and publish

From the outset, the EpiCherchell programm (funded by *MASA*) was set up within the landscape of international open community for digital epigraphy using the same technologies for data exchange (XML-TEI, DC metadata), their citability, their archiving.

The database is a collaborative tool for studying the corpus of the inscriptions of *Caesarea Mauretaniae* (Algeria) based on the rich photographic documentation of the Centre Camille Jullian. EpiCherchell gathers the information or offers a link with the related documents though creating a network: some records are linked to archival papers available in AreAr (Archives and Archeology platform at CCJ).

Three main goals: create an epigraphic corpus from a digital tool according to the FAIR principles in order to efficiently use and share epigraphic information among researchers, students and epigraphists; make a scientific publication of the corpus online; contribute to the training of young Algerian and French researchers in epigraphy and digital matters.

EpiCherchell is designed as a publication tool for the epigraphic corpus using the XML-TEI model. As a standard file, it allows the exchange of data enriched by links to repositories, gazeteers (Geonames, Pleiades) or vocabularies (EAGLE, PACTOLS). It constitutes a pivot file from which various forms of publications can be produced.

EpiCherchell is based on the FAIR principles to facilitate the acquisition and study of data, to allow the collaboration of geographically distant researchers and thus to enhance the sharing of knowledge between the two shores of the Mediterranean sea.

Dr Marianna SIMOES

Senckenberg Deutsches Entomologisches Institut (SDEI) / University of Hamburg - DE

I'm an evolutionary biologist with research foci lying at the intersection of phylogenetics, spatial ecology and ecological biogeography, with a particular interest in probing macroevolutionary patterns and processes. Currently, I am researcher and curator of the Coleoptera collection at the Senckenberg German Entomological Institute (SDEI), with research largely collection-based,focused on the evolution of leaf beetles (Chrysomeloidea), a very large and diverse group of phytophagous beetles. With the majority of my projects being collection-based, I am highly interested in the field of biodiversity informatics, promoting and generating methods to solve problems of organizing, accessing, visualizing and analyzing primary biodiversity data.

Contact: **mariannavpsimoes@gmail.com**

Primary data and Entomological collections: utility, best practices and challenges

Museum collections provide a rich source of historical, specimen-based, occurrence data that represent known occurrences of species in time and space, also known as primary biodiversity data. The increase of digitally available primary biodiversity data has been a positive result of sharing initiatives in the natural history museum community and among citizen scientists. Utility of this data is vastly applied on assessment and validation of biological patterns and processes, being essential to the integrity of biological knowledge. Owing to the heterogeneity of sources, however, limitations related to data quality control emerge, as incomplete and/or erroneous information at different stages of input must be overcome. Therefore, use of such data must be preceded by detailed data cleaning, a process that involves detecting, flagging, documenting and either correcting or eliminating suspect records from specific analyses. Due to the size and complexity of entomological collections, the magnitude of the task in digitizing and making primary data available are unrivaled in natural history collections. Thus, in my talk I intend to demonstrate the utility of entomological primary data on biodiversity assessments, and demonstrate best practices applied while seeking to reduce/eliminate mistakes associated during the digitization process, focusing on the approach used on the digitization of the coleoptera collection hosted by the Senckenberg German Entomological Institute (SDEI).

Dr Sara TONELLI

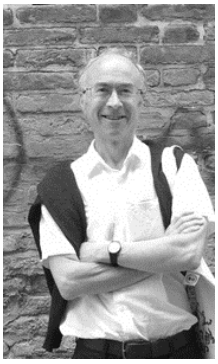
Bruno Kessler Foundation (FBK), Trento - IT

Sara Tonelli is the head of the Digital Humanities research group at Fondazione Bruno Kessler in Trento, and adjunct professor of Language Interfaces at the Dept. of Psychology and Cognitive Science, University of Trento. She got a PhD in Language Sciences from Università Ca' Foscari in 2010, after which she joined FBK first as a post-doc in the Natural Language Processing group and then as a tenured researcher leading the newly founded DH group. She has been involved in several national and European projects dealing with information extraction and processing from historical archives and digital cultural heritage collections. Her research interests are highly interdisciplinary, trying to apply and adapt advanced text analysis approaches to historical and cultural investigation.

Contact: **satonelli@fbk.eu**

What NLP can do for Metadata Quality: The Case of Descriptions in Cultural Heritage Records

Metadata allow the access to a wide variety of cultural heritage resources made available through repositories, digital libraries and catalogues. Usually taking the form of a structured set of descriptive elements, metadata assist in the identification, location, processing, tracking, preserving, sharing and retrieval of information, while facilitating content and access management. However, low metadata quality, such as incorrect information provided by textual record elements or inconsistency is still an open issue in many repositories and the manual evaluation of each record is not affordable, even for middle-sized collections. Recent advances in machine learning and natural language processing, however, can greatly support curators in checking metadata quality when text is present, since they enable to automatically evaluate several records in a short time, highlighting lacking information in records and providing aggregate statistics on entire collections. In this talk, I will present ongoing work aiming at automatizing metadata quality analysis using a machine learning approach. The preliminary results obtained on descriptions of visual artworks, archaeology and architecture extracted from the Italian digital library 'Cultura Italia' are promising, and show that supervised machine learning can be effectively used to assist curators in assessing the content of descriptions, significantly reducing the time needed to check metadata quality.



Prof. Douglas TUDHOPE

University of South Wales - UK

Co-author: Ceri BINDING

AHRC funded STAR, STELLAR and SENESCHAL projects (Semantic Tools for Archaeological Resources), in collaboration with English Heritage and the 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 an Editor of the journal, New Review of Hypermedia and Multimedia. He is a member of the Networked Knowledge Organisation Systems/Services (NKOS) network. He was a member ISO TC46/SC9/SC8 (and NISO) working group developing the new thesaurus standard (ISO 25964). http://hypermedia.research.southwales.ac.uk/kos/

Contact: **douglas.tudhope@southwales.ac.uk**

Reflections on the semantic integration of archaeological datasets and grey literature reports

The availability of different forms of open data within the field of cultural heritage and the increasing influence of the FAIR principles offer many opportunities for data reuse but also considerable challenges. This presentation reflects on experience from semantic integration exercises involving archaeological datasets and grey literature reports, primarily from a case study for the ARIADNE FP7 infrastructure project on item level data integration. This involved archaeological data and reports in different languages following a broad theme of wooden material, objects and samples dated via dendrochronological analysis. The semantic framework combined the Getty Art and Architecture Thesaurus (AAT) with a set of classes from the CIDOC CRM. The work required the combination of data cleaning, natural language processing and data integration methods with a web application offering semantic search. The same template based data conversion method was used for extracts from the archaeological datasets and also the information extracted from the archaeological reports. With finite resources it is not possible to develop parallel implementations in order to compare major design alternatives and thus not easy to know the consequences of one implementation choice over another. We reflect on some practical implementation choices and issues encountered during the data integration projects and discuss approaches to producing appropriate metadata for reuse.

Dr Cristina VERTAN

University of Hamburg - DE

Cristina Vertan is computer scientist working as senior researcher at the University of Hamburg. She holds a PhD in Computer Science from the University of Bucharest and a Master in Statistics from the Free University Brussels Her main research interests are data modeling in digital humanities, with particular focus on historical documents and natural language processing for less resourced rich morphology languages. She is author of more than 100 scientific publications and coordinated the IT-Unit of several digital humanities projects related to cultural heritage in various domain (classical Ethiopic, classical Maya, classical philology, ottoman studies).

Contact: **fshao60@uni-hamburg.de**

A fuzzy ontology for historical places people and events in historical texts - a case- study of historical texts of Dimitrie Cantemir

The massive digitization activities in recent years made available many historical works, as scans with corresponding metadata, or transcripted and annotated in TEI .For the analysis of such texts, the annotation of people and places of great importance. For modern texts, text mining and language technology processes are used to identify, annotate and link so-called "named entities" and to link them with registers and/or geo-coordinates. For historical texts, the problem is much more complicated for several reasons:

- Some geographic positions are not or just partially known, linking with geo coordinates is sometimes impossible;
- The names have been changed throughout history
- Metaphorical expressions related to place names according to several traditions;
- The historical maps are unreliable;
- People often do not know birth and death information, or are not sure of this information.

In this contribution we will explain how to build a semantic network of places and people for the European regions of the Ottoman Empire. The work is performed within the project HerCoRe (https://www.inf.uni-hamburg.de/inst/dmp/hercore/projects.html) in which we investigate among others, the reliability of his historical assertions by:highlighting the linguistic expression of uncertainty and vagueness The core of the investigation and annotation process is a is a Fuzzy OWL Ontology, trying to model he ottoman empire world in its administrative, social, geographical and religious facets.

Prof. Danielle ZIÉBELIN

Department of Computer Science, Grenoble Alpes University – FR

Co-authors: Matthew Sreeves, Fatima Danash, Karine Aubry, Emilie Chalmin

Danielle Ziébelin is a Professor at the Department of Computer Science at Université Grenoble Alpes, since 2010. From 2013 to the present she is responsible for International relations in the Department and from 2017 to the present she is also in charge of the Master track Artificial Intelligence and Web. Her research interests include knowledge representation, Semantic Web, spatio-temporal reasoning. She uses Semantic Web technologies to provide semantic interoperability to heterogeneous data resources and Linked Open Data facilities in many application domains, especially that of cultural heritage. Her focus is on extending knowledge representation languages to provide adequate modelization languages for different users in their respective fields.

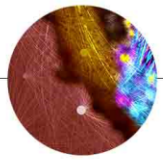
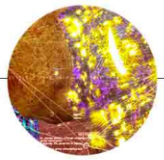
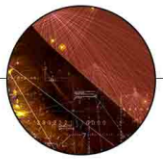
Contact: **Danielle.Ziebelin@imag.fr**

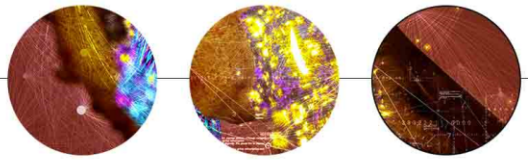
Patrimalp, an integrated and interdisciplinary cultural heritage platform for the French Alps

Understanding the creation, evolution and transformations of a heritage artefact requires data from different themes such as art history, archaeology, material science, geosciences, restoration and conservation as well as knowledge on the sustainable management of this artefact [1].
The data collected on the cultural heritage artefact is analysed to form a conceptualization of its representation: information on the structure, design, restoration and evolution are stored on the platform. For example pigments are analysed through the ages, stored in a pigment library [2], connected to the artefacts in which they appear, and tested on their evolution and degradation over time.
The platform is based on a semantic web architecture, aggregating data from different thematic information systems according to the multiple ontologies approach [3]. This induces a graph-based knowledge representation for the artefacts and provides the necessary elements for building the life cycle of it with elements from the different domains previously mentioned. The knowledge graph includes the contribution of each discipline by providing a mapping for the different discipline ontologies and standards (OpenTheso, Getty-AAT). In addition to the graph, the architecture includes a query module allowing the use of external resources (open linked data). For future work, the architecture will follow a global structuring ontology approach [3] using a parthood model including space and time [4].

References:

[1] https://www.researchgate.net/project/Pigmentotheque
[2] From archeological layers to schematic rock art? Integrated study of the Neolithic pigments and pigmented rocks at the Rocher du Château (Western Alps, Savoie, France) Claudia Defrasne, Emilie Chalmin, Ludovic Bellot-Gurlet, Eric Thirault, Guy André Archaeological and Anthropological Sciences, Springer, 2019, 11 (11), pp.6065-6091. ⟨10.1007/s12520-019-00882-9⟩
[3] Ontology-based data integration in multi-disciplinary engineering environments: a review - FJ Ekaputra, M Sabou, E Serral, E Kiesling... - Open Journal of Information Systems (OJIS) Volume 4, Issue 1, 2017
[4] A Parthood Approach for the conceptual modelling of Tangible Objects Composition (TOC) - an application on Cultural Heritage (CH) Fatima Danash, Danielle Ziebelin and Emilie Chalmin eswc2020- posters (submitted)



Scientific programme

Wednesday June 17th 2020

09:00 | **Technical introduction & institutional presentation**
Prof. Benoist Pierre, Sophie Gabillet, Dr Xavier Rodier, Dr Emilio M. Sanfilippo & Dr Damien Vurpillot

09:30 | **Official Opening**
Dr Xavier Rodier & Dr Emilio M. Sanfilippo

First session: Ontologies and semantic web technologies for cultural heritage

10:30 | **Dr George Bruseker**
Realizing the Virtues (and Combating the Vices) of the Digital Shift in Heritage Scholarship: the Role of Knowledge Representation

11:15 | *Coffee break*

11:30 | **Dr Alessandro Mosca**
Ontology-based data integration in EPNet

12:00 | **Dr Cesar Gonzalez-Perez**
Ontologies of Cultural Heritage for Humans and Machines: The Cultural Heritage Abstract Reference Model

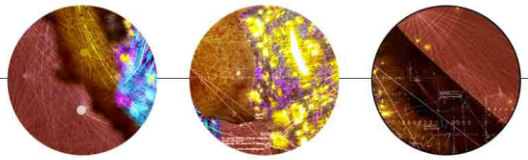
12:30 | *Lunch break*

Second session: Data management for natural heritage

14:00 | **Dr Marianna Simoes**
Primary data and Entomological collections: utility, best practices and challenges

14:45 | **Dr Cécile Callou**
Du projet d'inventaire archéologique de la faune et de la flore française vers un portail bioarchéologique global

15:15 | *Break*



Third session: Digital scholarly editing: theories, methodologies, and tools

15:30 | **Prof. Elena Pierazzo**
The edition of draft manuscripts: theories, (data) models and methods

16:00 | **Dr Carlo Meghini**
Extending Digital Libraries with Digital Narratives

FLASH TALKS

16:30 | **Denise Ardesi & Christian Steiner**
Semantic coordination on Medieval cooking recipes

16:40 | **Mériç Akdogan, Pauline Bellemere & Inoussa Kora Chabi**
DiVE Project

16:50 | **Djibril Diarra**
Hybrid Artificial Intelligence for Medieval Illuminations Analysis

17:00 | **Dr Wieslawa Duzy**
Project "Historical Ontology of Urban Space" (HOUSe) - introduction

17:10 | **Florian Hivert**
ROSER - Répertoire de l'Ornement Sculpté des Édifices de la Renaissance: An interoperable scholarly digital edition for the ornamentation in Renaissance edifice

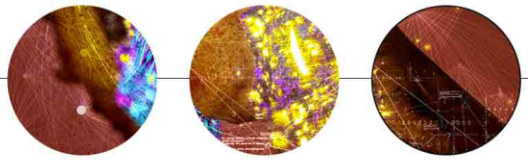
17:20 | **Blandine Nouvel**
Deconstructing for reconstructing: the use of the BBT for reorganising the PACTOLS thesaurus

17:30 | **Prof. Ludger Jansen**
Why DH ontologies need standards too

17:40 | **Dr Tomasz Panecki**
Digital editions of historical maps: between images and data. The example of "Gaul/Raczyński" topographic map (1807-1812)

17:50 | **Caroline Parfait & Marlène, Arruga**
Development of digital platform about stained glass restoration data in France

18:00 | **Dr Stéphanie Satre**
Epicherchell, an epigraphic dataset to share and publish



18:20 | **Prof. Danielle Ziébelin**
Patrimalp, an integrated and interdisciplinary cultural heritage platform for the French Alps

18:30 | **Final discussions**

Thursday 18th June 2020

09:00 | **Damien Vurpillot & Perrine Thuringer**
ARD Programme Intelligence des Patrimoines: Heritage(s) digital ecosystem

09:30 | **Miled Rousset**
OpenTheso and PACTOLS

10:00 | *Coffee break*

10:15 | **Emmanuelle Bryas, Nathalie Le Tellier-Becquart & Emmanuelle Morlock**
A Data Management Plan Template for Archaeology

10:45 | **Olivier Marlet**
OpenArcheo

Fourth session: Data mining and nlp for cultural heritage

11:30 | **Dr Sara Tonelli**
What NLP can do for Metadata Quality: The Case of Descriptions in Cultural Heritage Records

12:30 | *Lunch*

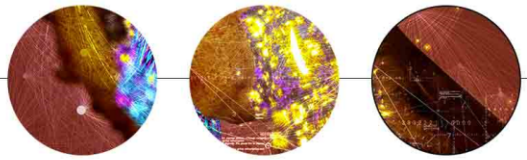
Fifth session: Information systems for 3d data and semantic annotations

14:00 | **Prof. Roland Billen**
The Coudenberg Heritage Building Information System

14:45 | **Dr Florent Laroche**
French National Project ReSeed - Semantic reverse-engineering of digital heritage objects

15:15 | **Dr Ana Roxin**
Semantic-based BIM Modeling: Approaches and related Standardization Actions

15:45 | *Break*

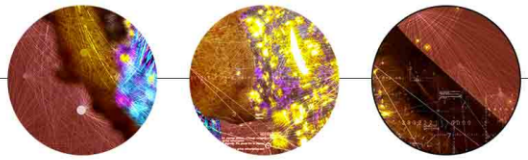


Sixth session: Digital data services for cultural heritage

16:15 | **Dr Livio De Luca**
A digital ecosystem to document, in space and time, the restoration of Notre-Dame de Paris

17:00 | **Prof. Julian D. Richards**
Making Archaeology FAIR: in the Archaeology Data Service and ARIADNE

17:30 | **Prof. Douglas Tudhope**
Reflections on the semantic integration of archaeological datasets and grey literature reports



Partners

ARD 2020 PROGRAMME INTELLIGENCE DES PATRIMOINES

An interdisciplinary research programme dedicated to innovation, training and scientific development, Intelligence des Patrimoines offers a new understanding of cultural and natural heritage.

Intelligence des Patrimoines is an interdisciplinary scientific research and innovation programme serving the tourism heritage economy in the Centre-Val de Loire region and offering new education and employment perspectives.
It consists in an unprecedented approach to the promotion of the territory combining scientific research with the socio-economic world to design new innovative services and products.
As part of the ARD programme (Ambition Research Development), it is supported by the Centre-Val de Loire Region and hosted within the Centre d'études supérieures de la Renaissance (CESR) in Tours.
This programme brings together thirty-three different laboratories and three hundred and sixty researchers from various teaching and research institutions in Tours and Orléans such as: Tours University, Orleans University, CNRS, INRAE, INSA, BRGM, ESCEM and LE STUDIUM.

The ARD programme Intelligence des Patrimoines is articulated around five major interdisciplinary research topics towards the development of interdisciplinary scientific activities:

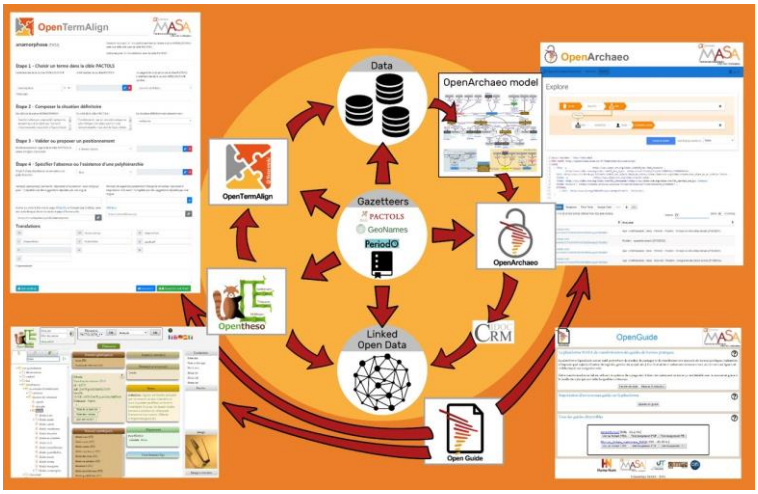
- Chambord-Chateaux
- Vine and Wine
- Gastronomy, Health and Wellbeing
- Loire and Rivers
- Monuments, Parks and Urban Gardens

The programme is developing a transversal digital ecosystem named Heritage(S) composed of three platforms dedicated to the collect and dissemination of cultural and natural heritage data.
It also offers new multidisciplinary and professional training courses at Master and Research Doctoral levels (École supérieure en Intelligence des Patrimoines) and has created a thematic academic incubator – the Smart Tourism Lab - for the development of startups on the regional territory and to support entrepreneurial projects dedicated to promote and renew the tourist experience around heritage.

THE MASA CONSORTIUM

The consortium Mémoires des Archéologues et des Sites Archéologiques, from the very large research infrastructure Huma-Num, is rooted at the heart of the Open Science movement with the aim of transmitting and implementing the FAIR principles in the French archaeological community in association with international infrastructures. MASA mission is to produce guidelines to help researchers better choose digital formats and find long-term and interoperable solutions. Faced to the archaeological data accumulation and the heterogeneity of the systems and structures of the datasets, MASA focused the objective of sharing archaeological data and archives in the Linked Open Data in order to promote the emergence of archaeological data on the Semantic Web.

To meet this objective, MASA proposes a process of data manipulation from acquisition to publication according to a systemic approach that respects FAIR principles. The MASA ecosystem is composed of bricks for publishing and sharing archaeological datasets. Once processed, documented and standardized, the archaeological datasets are put online according to the standards in force (XML, TEI, EAD...). Standardized repositories are used for spatial (GeoNames) and temporal (PeriodO) information. For descriptive information, MASA uses the PACTOLS thesaurus via the OpenTheso thesaurus manager. The matching of dataset vocabulary with standardized thesauri can be achieved through the online application OpenTermAlign, which generates a SKOS file of the aligned vocabulary. The OpenArchaeo platform ensures the interoperability of several datasets, in a MASA triplestore or in external triplestores, offering an intuitive query interface that translates requests into SPARQL according to the OpenArchaeo model for mapping archaeological data with the CIDOC CRM ontology. A good practice guide in the OpenGuide platform developed for this purpose documents each step of the process. With this digital ecosystem, the MASA consortium relies on the data culture of archaeologists and their long experience in computerization to get the community to respect the FAIR principles and to open these corpuses in the Linked Open Data.



LE STUDIUM Loire Valley Institute for Advanced Studies

Established in 1996 and inspired by the historical, geographical and human cultures of the Loire Valley, LE STUDIUM has evolved to become a multidisciplinary Loire Valley Institute for Advanced Studies (IAS), operating in the Centre-Val de Loire region of France. LE STUDIUM has its headquarters in the city centre of Orleans in the amazing facilities of a newly renovated 17th century building. LE STUDIUM IAS offers programmes that include public and private research stakeholders, to increase research capacity, promote quality international collaborative research projects and support innovation and social and economic development.

LE STUDIUM IAS offers to internationally competitive senior research scientists the opportunity to discover and work in one of the IAS's affiliate laboratories from the University of Tours, the University of Orleans, National Institute of Applied Sciences (INSA) Centre Val de Loire and ESAD Orléans, as well as of nationally accredited research institutions located in the region Centre-Val de Loire (BRGM, CEA, CNRS, INSERM, INRAE). Our goal is to develop and nurture interdisciplinary approaches as innovative tools for addressing some of the key scientific, socio-economic and cultural questions of the 21st century. For the period 2015-2021, LE STUDIUM operates with an additional award from the European Commission - Marie Skłodowska-Curie Actions (MSCA) - for the mobility of experienced researchers. LE STUDIUM is also the official partner of the Ambition Research and Development 2020 (ARD 2020) programmes initiated by the Region Centre-Val de Loire to support the regional specialisation strategy around five main axes: biopharmaceuticals, renewable energies, cosmetics, environmental metrology and natural and cultural heritage.

LE STUDIUM has attracted over two hundred LE STUDIUM RESEARCH FELLOWS and LE STUDIUM RESEARCH PROFESSORS for long term residencies coming from 47 countries of Europe, America, Africa, Asia and Oceania. International experienced researchers and their research projects are selected through an annual call for application open to all scientific disciplines or goal-oriented calls for application and with the support of an international scientific council.

In addition to the contribution in their host laboratories, researchers participate in the scientific life of the IAS through attendance at monthly interdisciplinary meetings called LE STUDIUM THURSDAYS gathering members of the regional scientific community and industries.

Researchers are also invited to organize a two-day LE STUDIUM CONFERENCE. It provides them with the opportunity to invite internationally renowned researchers to a cross-disciplinary conference, on a topical issue, to review progress, discuss future studies and strategies and stimulate advances and practical applications in the chosen field. The present LE STUDIUM CONFERENCE named "FAIR HERITAGE: Digital Methods, Scholarly Editing and Tools for Cultural and Natural Heritage " is the 108th in a series started at the end of 2010.

We thank you for your participation and wish you an interesting and intellectually stimulating conference.