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

----

This document reports some discussions between the participants at the FAIR Heritage conference, which took place on June 17-18 2020 as a Web event. Some of the questions were discussed during the talks, therefore the answers do not appear in the document.

The document was used to facilitate the interaction between the conference participants. It did not undergo a professional editing process, nor a specific policy was adopted for commenting or replying to the asked questions.

For more information about the conference, including the slides of the presentations and the recordings, please visit: http://www.lestudium-ias.com/event/fair-heritage-digital-methods-scholarly-editing-and-tools-cultural-and-natural-heritage.

You can also contact the conference chairs:

Emilio Sanfilippo (namesurname AT gmail dot com)
Xavier Rodier (name dot surname AT univ-tours dot fr)

Day One (Wednesday June 17, 2020)

SESSION 1: 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

Cesar Gonzalez-Perez: There's a large body of literature in the software engineering field that shows that reusability is not a given that comes for free, something that can be added to data as an afterthought, but on the contrary it's something that must be engineering into the data from the start. However, most reusability efforts that I see around are concerned with making data reusable as an afterthought, rather than creating data that is reusable from the start. What's your take on this? How can we create datasets that are more reusable from their inception? Thank you!

Emilio Sanfilippo: What do you think that we can learn from the covid-19 pandemic with respect to digital culture equity and the FAIRness of digital objects?

Completely agree with the importance of community and the need of teaching materials about ontology engineering for knowledge representation/reasoning and data management.

Margriet: Thanks for the interesting and very informative presentation. My specialism is medieval history and culture. In my field our opinion about the nature of data can change over time, as attribution to an author, datation, localisation etc. My concern is that semantics and ontologies will not be able to keep pace with the progress of research and that it will become a repository of outdated data. Any ideas?

Beatrice Markhoff: @Margriet, in semantic Web there are ontologies such PROV-O to document the production and evolution of knowledge, CRMInf may be also a way to keep track of such evolutions…

Emilio: @Margriet some keywords to look for are - ontology evolution and knowledge evolution

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

Marion Lamé

Two topics :

(1)   EAGLE metadata
(2)   Ontology for describing ductus (epigraphical writings) (right part of the data model (« transcriptions »)

@alessandro I still do believe and think that your ontological model should be extended to the description fo writings ; this is cardinal when it comes to Text Bearing Object, whether they are epigraphical or not. PLUS compatibility at least with TEI and eventually with EpiDoc - but I would not encourage this last one. Once TEI compatibilty is done, EpiDoc follows, by nature.

Emmanuelle Morlock:

TEI/EpiDoc guidelines: http://www.stoa.org/epidoc/gl/latest/

TEI/EpiDoc Schema: https://sourceforge.net/p/epidoc/wiki/Schema/

Ale Mosca: Hi Marion! As far as I got, while working in the EPNet project, I cannot agree more with you. Honestly speaking, I pushed several times during the project lifespan the historians I was working with to properly study the connections with EpiDoc and TEI. When I left the project last here, this was an explicit task in the agenda. After that, I did not have any news on that from the Barcelona colleagues (i.e., the domain experts). One issue I was thinking about relates with finding some technique (from NLP technologies, maybe) to semi-automatic annotate their inscriptions/stamp in a way that is compatible with TEI (by, for instance, automatically identifying persons, places, edible items, activities, that are recognised as present in the inscriptions and reported in the inscription transcriptions). Technically speaking, I don't see any a-priori issue in moving towards the description of the writings you mention and, this way, extend the model we developed. Let me finally say that EPNet was probably a project with too many tasks, from agent-based simulation to complex networks analysis, passing through a bunch of non-trivial data management issues. Personally, I can talk about this last aspect only, and related to that I would definitely support a follow-up of the project 100% focused on data management: compatibility formal alignment with existing standards, integration of others available datasets, extensions of the domain

ontology. I left the project with the feeling that a lot of data management tasks have been left open, indeed.

Marion Lamé: Working expressions (such as named entities for instance,, i.e. how the community of experts express themselves with or without a computer) and writings are two different levels of abstraction of the digital scholarly editing process.

The first one regards at least the transcribe and edited text (critical and less diplomatic) and sometimes also its translation, and most often its comment. This deals with the textual system.

The second one regards the Digital Autoptic Processes and annotation + computer graphics. This deals with the writing and the contextual system of the inscribe object under study.

They are two different branches of computer sciences and also from the humanities standpoints. The first one regards a bit more linguistics (computational or not), the second one epigraphic and archaeological fields, computer graphics fields but also, in case of manuscripts, codicological and paleographical competences - paleographical works also fine for epigraphy anyway.

Ale Mosca: Marion, please consider that you are talking with a person working with mathematical logic and knowledge representation technologies in an artificial intelligence research group. What I learned about inscriptions and epigraphy is nothing compared to what an expert in this field knows. I've just interacted for three/four years with domain experts but does not make me feel comfortable jumping into deep discussions about these matters. I'm sorry about that! I would need more time and exchanges with you to get more about the issues you mention. Let's keep in contact in the next weeks: I'm more than happy to help, if I can!

Marion: Here you are : marion.lame@univ-tours.fr

Ale Mosca: here is mine alessandro.mosca@unibz.it

Marion: Your presentation was very clear and mathematical and logic fields are those who are able to unify the fields mentioned above with core editorial processes (imho).

I think the EAGLE data model requires a 2.0 version, the first one was attempting a first synthesis between verbal and non verbal information, not always easy to deal with. But it was a really great intellectual effort.

Loup BERNARD

You mentioned overlapping datasets that complete each other, we have the same issues in ArkeoGIS : How do You quote the end information ? i.e. do You quote any dataset used to produce the final document, and how ? Is the person who did the work mentioned ? the technology ? both ?

Ale Mosca: Hi Bernard! In principle, if you rely on OBDI technologies, the data identifiers are those of the original datasets: they are "simply" mapped to (possibly new) identifiers that are used to populate your domain ontology and support the query answering in a consistent way. That means that the so-called "landing pages" for the data that are integrated can be always retrieved from the system specification and consequently explored, together with their authorship, quoting, etc. Another option is the one of explicitly represent in your ontology the *provenance* information of the integrated data and we are currently working on this specific kind of information to properly deal with them in an OBDI setting.

Thank You Alessandro for the precisions, I hope we can exchange about those topics soon, Loup

Ale Mosca: With pleasure! Here is my email account you can freely use to reach me: alessandro.mosca@unibz.it

- 12:10 **Dr Cesar Gonzalez-Perez** - Ontologies of Cultural Heritage for Humans and Machines: The Cultural Heritage Abstract Reference Model

Emilio Sanfilippo:
Could you comment on the modeling of vague knowledge/data with ConML/CHARM?

Cesar: We distinguish ontological and epistemic vagueness. Ontological is about the world not having clear-cut boundaries, such as the area of a city or the extension of a mountain. Epistemic is about us not being certain of things, such as not being sure about for example how many children Alexander the Great had. ConML supports both to different degrees. Also, it supports the notion of *unknown*, which is different to "null" in databases or programming. "Null" means absence of data, whereas "unknown" means that there is data, but we don't know it. In a model having a Person class with a Job attribute, for example, p.Job = null means that person p doesn't have a job, whereas p.Job = unknown means that p does have a job, but we don't know what it is. This distinction is crucial.

ES: What are examples of ontological vagueness in the context of cultural heritage?

Cesar: The concept of *site* in archaeology is a good example. Sites are usually defined spatially, by having certain boundaries, and these boundaries are established interpretatively depending on archaeological evidences (artefacts, features, etc.) but also natural features of the landscape, human-made elements such as roads, etc. In this manner, the surface area, or the boundary, of any given site, is ontological vague, as it is not clear-cut.

A more mundane example is that of a hill. If I'm sitting at the top of a hill, it's clear that "I am on the hill"; if I'm sitting at the bottom of the valley, it's clear that "I'm not on the hill". But when I walk from the top downwards, there is no specific point at which this transition happens; it's gradual. The boundaries of a hill, therefore, are ontologically underdetermined, and we say they are ontologically vague.

ES: Could you also comment on what you mean by subjectivity?

Cesar: We mean the fact that different people may have different perspectives on the same thing. For example, neighbours in my village think that the old church needs restoration, but the local authorities say it doesn't. So, if we imagine a model having a Building class with a ConservationStatus attribute, this attribute should be marked as subjective,

so that datasets can store different values for it depending on who is speaking. This allows for multivocal datasets.

ES: Is CHARM formally represented in logic?

Cesar: No. It is expressed in ConML, which is a semi-formal conceptual language: www.conml.org You can download CHARM in Bundt format from www.charminfo.org, and the Bundt tools (to open and work with it) from www.conml.org/bundt.

ES: What kind of database technologies do you use in tandem with CHARM?

Cesar: If you mean how CHARM is stored, it is stored in Bundt format; this doesn't use a database but a file-based persister that we developed. If, on the contrary, you refer to what databases we use when implementing projects involving CHARM, then the answer is varied. We usually employ what's best suited for the problem at hand, depending on engineering and customer needs. We've used a bit of everything, from relational such as SQL Server to noSQL or even a file-based store for a mobile app once.

Beatrice:

Could you show an example of reinterpretation rule? What is the relationship between such rules and ontology mappings (i.e. mappings between ontologies)?

Cesar: Basically, an ontology mapping is constructed manually by an analyst by finding what concepts in an ontology "are the same as" concepts in the other. Reinterpretation rules work by assuming that there is an extended ontology that was constructed from a base ontology, and uses identity markers in model elements to establish a "mapping" automatically. The fact that both ontologies are related by an extension relationship is crucial for this to work. An example of a reinterpretation rule is this: "An object in the dataset having a non-root extended class in the extended model as type is reinterpreted to have the most immediate ancestor of said class that exists in the base model as type". There are many others. The modelling engine in Bundt applies these rules and generates a reinterpreted dataset automatically. If you are interested in reinterpretation rules, write me to cesar.gonzalez-perez@incipit.csic.es and I'll send you some material if you wish.

Djibril: In the conclusion of Cesar's presentation, it is said that: we have to focus on the domain, and remove the implementation noise. I don't know if i got it well or not, but sometimes, focusing on the domain means be as expressive as possible in the modelling. And we know that sometimes more expressivity deals with undecidability. And that last deals with big problems in the implementation. My question is how do you handle this issue ? Thank you

Cesar: OK, I think I have an answer. One of the crucial aspects of modelling is deciding what details of the world we keep and which we discard. Usually, this is determined by the model's goals and other contextual parameters, In

any case, expressivity must be understood not in terms of quantity *maximizing what the model contains(, but in terms of quality (expressing what we need to say in a manner that is as unambiguous and concise as possible). By eliminating detail, models allow us to fight complexity in the world. If we choose to capture so much detail that the model is still too complex (undecidable as you put it), then it's up to us to simplify a bit more. In any case, modularization and layering principles (which I mentioned in my talk) can help to achieve this.

Ale Mosca: Cesar, Djibril, If I can say something from a KR point of view, I would like to say that the main problem of any knowledge representation initiative is to find the right balance between expressivity of the languages in place and their (computational) complexity (in terms of reasoning with them). Actually, in KR, we always try to make explicit the connection between expressivity and complexity of the languages we devise in such a way that the users of them have a map to guide their decision (I need to reason about time intervals, fine, this is the complexity of this task I have to deal with; am I fine with reasoning with points in time, instead, then this is the complexity of reasoning with languages expressing timestamped info, and so on). Unfortunately, it is not always a matter of discarding (or, abstracting out) a few characteristics of the real word in order to get a (efficiently) computable knowledge representation model: we are plenty of super-simple example of logical theories (e.g., ontologies) that lead to undecidability (we nerdy like them a lot! ;-)). It is also important, I guess, to have a clear understanding of the ultimate goal of your modelling and representational activity: what the ontology/conceptual model will be used for? If it's for documentation purposes only, for instance, you shouldn't care at all about computational undecidability: different languages, that is different expressivity and computational complexity options, are used for different tasks.

My two cents!

Cesar: Thanks, Alessandro. I fully agree with you!

Djibril: Thank you for your answers, Cesar and Ale.

SESSION 2: DATA MANAGEMENT FOR NATURAL HERITAGE

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

Emilio Sanfilippo: from a data management perspective, it would be interesting to know what kind of information is represented in entomological labels.
Also, are there standards or common practices for creating the labels in a homogeneous way across multiple research organizations?

Do you use any reference model/biological taxonomic system? E.g., TAXREF

Loup BERNARD

Do You have a specific list of specimens linked to archaeological contexts ? and/or to archaeological/historical models linked to bugs ? In my experience, this is more important for us archaeologists that millions of bugs, which we are too often unable to use

Marianna: Dear Bernard, I am not aware of platforms where we can find archeological contexts linked to insects. I am aware however, of this manuscript, which seems to address this question in a better manner and more complete , than I could: :Elias, S. A. (2010). 7 The Use of Insect Fossils in Archeology. *Developments in Quaternary Sciences*, *12*, 89-121.

Vielen Dank !

Gerne, Bitte schön

Cécile Callou

Do you know DISSCO program, a network for digital unification of all European natural Science collection ? (https://www.dissco.eu/).

> **Marianna:** yes , thank you for mentioning this. We are aware of Dissco, and right now we have a data manager in Senckenberg with the goal of unifying all the databases we have in the institute (Aquila, Edaphobase, Paradox..) so we can unify and integrate into Dissco format in the future. Lets keep our fingers crossed to make the merge of all databases possible. That would be of great benefit for all of us. :)

And for the label transcription, you can consult "les Herbonautes", a collaborative platform crate specially for this question : http://lesherbonautes.mnhn.fr/

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

(ZOOM Name) / (Real Name) / Question

Emilio Sanfilippo: is there any specific ontology you use (or plan to use) for your projects?

Blandine Nouvel: about linking with PACTOLS thesaurus.

> **Cécile:** Dear Blandine , Yes we need to work together.

> **Blandine:** Shall we phone or email after the conference ?

> **Cécile:** of course Yes, my email is in abstract book. Got it!

SESSION 3: DIGITAL SCHOLARLY EDITING: THEORIES, METHODOLOGIES, AND TOOLS

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

(ZOOM Name) / (Real Name) / Question

- 16:00   Dr Carlo Meghini - Extending Digital Libraries with Digital Narratives

(ZOOM Name) / (Real Name) / Question

Emilio:

I'm wondering whether the formal representation of narratives can help in understanding whether two or more (literary) "works" (e.g., translations) have the same narrative, therefore they are the same work (or they are similar works, if their narratives are similar but not identical). This comment recalls @Elena Pierazzo's final discussion during her presentation.

Carlo:

In fact, we have been advised to use our model of narratives to detect fake news, by comparing the news with narratives of the same fact. I believe that all depends on the ability of getting the relevant details in the narratives. But we are already satisfied by the fact that thanks to our ontology we have a space for the problem, where at least some of the aspects can be formulated as technical questions on which a researcher can work.

Emilio:

From a practical perspective, how feasible is it to build narratives out of novels/poems in the form of RDF graphs? I imagine that building a RDF graph for the entire narrative of, e.g., the Divine Comedy can be quite expensive (time, energy, computational space..)

Carlo:

It is expensive to build narratives manually, but in my experience humanist scholars are used to do this kind of work, they spend a lot of times commenting a terzina or even a word in a terzina. The notes of the Divina Commedia are much more voluminous than the work itself, and complex in structure. And every time they have to restart from scratch. The problem is to protect the investment that they make by making their work available to others and re-usable by them in digital form.

Ludger Jansen:

Did you think about how to represent the series of events in time-travel fiction?

@Emilio: this is indeed a very interesting question! To me, in a more general perspective, one could ask: which kind of narrative can the ontology represent? From what I understand from Carlo's talk, the narrative has to be compliant with common sense reality

Carlo: As I said, we have included very only two axioms in our ontology about time, causes and events. One says that effects cannot happen before their causes.

The other says that sub-events occur within their super-events. The fabula of "Back to the future" clearly violates at least one of these axioms. So we can say that our ontology is limited in the capture of fictional stories. It's a trade-off: if we give up these axioms we would be able to model "Back to the future" but we would loose the ability of doing very reasonable inferences and enforcing a very reasonable form of consistency on all the reasonable narratives. I do not think it would be a great gain.

Ale Mosca: it would be funny to try with "https://en.wikipedia.org/wiki/The_Time_Traveler%27s_Wife" novel! Assuming that in this novel, there is *special* person who is able to move back and forth in time, all the rest of the narrative is perfectly reasonable and consistent ;-)

Carlo: In this case, my advise would be to omit the non-newtonian bit from the fabula, and keep it only in the narration where it does no harm :)

Ale Mosca : Do you think that the formal representation of narrative components (events) that you are proposing could be manipulated by a software agent to create/adapt the narratives on-the-fly to different users, with potentially different interests/requirements? I'm asking this 'cause I'm working at a project proposal where we would like to offer an automatic way to provide users with (educational) narratives about the cultural heritage of a given territory, according to their profile (whatever this means! e.g., difficulty in reading or others cognitive and physical disorders). The naive idea would be there to have a toolbox of "bricks" that can be used to create consistent narratives, according to the detected user context.

Carlo: We have this problem in Mingei. We must tell the same story to different users. We solved this problem by introducing another dimension of narratives, the presentation of the narrative, where a person is able to build a story of the fabula (another narration, if you like) borrowing different "pieces" of the narrative and assembling them the way they please.

Ale Mosca : Thank you very much for the suggestion Carlo! I'll bother you, maybe, in the near future to learn more from your experience! Thanks!

Carlo: You are welcome.

Elena Pierazzo

How feasible is a semi-automatic method applied to literature? [somewhat question connected to Ludger Jansen] I meant what happens when you have flashbacks/narrative gaps/dreams, not to mention irony. I understand how you can infer information from sort-of objective language like biography or similar, but in literature the linear sequence of the fabula is rarely found.

Carlo: This is no problem for the ontology: a narrator can arrange the events in the fabula the way they please, as long as they do not violate the basic laws of physics, which are captured as logical axioms in the ontology and help reasoning and keep the DL consistent. But I still believe that reading literature is a lot more fun than consuming it as a formal object! But if there are tasks that would benefit from a formal representation, then why not. In general, we do not see our narratives as tools for capturing novels at the formal level, we see them as tools for documenting, that is, we start from a process, we model it as a sequence of events, and then we build the narration by connecting the events to digital objects that document them. This is what we do in Mingei, perhaps in my presentation I did not emphasize it enough.

Beatrice Markhoff: thank you Carlo, seeing narratives as means of representing, thus keeping memory of, some creation processes is really a very nice idea: for me it also opens some perspectives regarding a (national) project, SESAMES, that some of the attendees (including me) here are working on.

Carlo: Thank you Beatrice, in fact I think narratives are a useful concept for digitization.

FLASH TALKS

- 16:30   **Marion Lamé** (speaking on behalf of Denise Ardesi) - Semantic coordination on Medieval cooking recipes

Blandine Nouvel: How do you manage the difference, if there is one, between the chaudron as it can be described in texts and the chaudron as an archaeological object form ?

Denise Ardesi: Thank you for the question. Yes there is a difference between chaudron nowaday and chauldron in medieval time. Thanks to Heterotoki we can provide a description in order to define the medieval chaudron. The target terminology should contain the concepts in modern languages and linked to their source in ancient languages + all the conepts in modern languages necessary for the description of any kind of medieval "chaudron". The modern terminology, partially linked with the ancient one, can then be used to annotate any digital artefact of archaeological "chaudron" according of what you need to describe. Aurélie Chantran is writing a phd on food history and archéology. Maybe if you are interested in this subject you can contact her: http://univ-paris1.academia.edu/AChantran

Would it fit to the PACTOLS definition for chaudron https://ark.frantiq.fr/ark:/26678/pcrtMchkJ5TnEz ?;)) if not, we'll have to improve our thesaurus with your help… I spotted you, Blandine! You are recognizable with such a comment!

The target terminology should contain the concepts in modern languages and linked to their source in ancient languages + all the concepts in modern languages necessary for the description of any kind of medieval "chaudron". The modern terminology, partially linked with the ancient one, can then be used to annotate any digital artefact of archaeological "chaudron" according to what you need to describe.

- 16:42   **Mériç Akdogan, Pauline Bellemere & Inoussa Kora Chabi -** DiVE Project

Emilio Sanfilippo: Can you comment on the idea of participative exhibition?

Pauline Bellemère: The idea of participative exhibition is the fact that someone else than you, for example one of your friends, can participate and put new content in your exhibition (which is, in our case, in a virtual 3D Museum).

Meriç AKDOĞAN: Actually, we would like to create a 3D platform which allows its users to create their own exhibitions based on DiVE's platform template or

participate the others people's exhibitions by adding theirs documents like photos, videos etc.

- 16:52 **Djibril Diarra -** Hybrid Artificial Intelligence for Medieval Illuminations

  Analysis

Emilio Sanfilippo: can you comment on the reason why a hybrid approach between machine learning and ontology engineering is required for your study? What do you gain with that? What are the advantages? How was the ontology fed into the deep learning algorithm?

Djibril : 1. The requirement of hybrid approach for the study is dual. First, the ontology is used for the expression and the preservation of the semantic of the objects contained in the illuminations. It allows us to describe the meaning of the explicit items in the illuminations and to formulate rules in order to infer implicit knowledge. This expression and description are done according to the medievist (medievalist) knowledge about the illuminations. Second, machine learning is used to facilitate the understanding of illuminations by allowing automatic detection of the illuminations' contains. It is useful for no specialist of the illuminations.

2. By this way, we gain the facility to extend the model in the case we have new illuminations. We also gain the facility to integrate our model with others.

3.The advantages are : data integration, the providing of tools for illuminations digital preservation and their understanding by no specialist of their medieval paintings.

4. The machine learning algorithm uses the ontology's concepts to propose automatic detection of the objects in the illumination uploaded in the system. The neural network evolves each time we add a new tag to the uploaded illumination.

- 17:04 **Dr Wieslawa Duzy -** Project "Historical Ontology of Urban Space"

  (HOUSe) - introduction

Damien Vurpillot: We are starting a project that could fit within the research topic (urban area in Tours city in France). I will see with my colleagues so that we can exchange with you.

Wieslawa Duzy: Brilliant! Some general info about the HOUSe project are available at our project website: https://urbanonto.ihpan.edu.pl/HOUSe_about.pdf And, please, use our urbanonto@ihpan.edu.pl e-mail equally with my wieslawa.duzy@gmail.com e-mail. Thank you!

- 17:12 **Prof. Ludger Jansen -** Why DH ontologies need standards too

Emilio Sanfilippo: Do you think that there are any relevant differences between ontology modeling for the life sciences and for the area of cultural heritage? For example, @Cesar claimed that modeling epistemological vagueness does play an important role for cultural heritage. Is it the case for the life sciences?

LJ: Epistemological vagueness is around everywhere. There are many known and unknown unknown around, both in science and humanities. So that should not be a difference. I try to argue for the similarities in http://ceur-ws.org/Vol-2518/paper-WODHSA5.pdf.

Emilio: I would be interested in understanding what are the differences (if there is any..) between data modeling / ontology modeling for the SHS, on the one hand, and for the life sciences (or engineering), on the other hand.

From ML to @ES: I am not sure that there is a huge difference in terms of technical processing. SHS might need more detailed editorial information.

ML: I do agree from the logical point of view, but not with the argument of "known" and "unknown" : in humanities part ong this vagueness will be due to subjective interpretations, not due to "known" and "unknown".

Emilio: We definitely need platforms for sharing cultural heritage ontologies. As @Geroge said, the platform Ontome is a start.

    LJ: I definitely need to take a look at this.

Loup Bernard

You might have problems with the multiningual ontologies, do the concepts have the same signification in different philosophy schools ? How to characterize a word in a foreign language in a database etc…

LJ: Ontologies are not about words, but about the things they refer to. You can have different labels from different languages for the same class term. On the other hand, even in one and the same language, a word might be ambiguous and you might need several classes to represent the thing(s) you refer to with this word.

Emilio: I think that a problem that Loup may have in mind is the fact that, e.g., universals are differently understood in different philosophical schools. So, how to deal with this kind of variety of meanings for the same notion? (Different approaches are actually feasible for this kind of problem.)

- 17:32 **Blandine Nouvel -** Deconstructing for reconstructing: the use of the BBT for reorganising the PACTOLS thesaurus

ML : In the future, how would you allow experts to, at the same time

- (1) suggest new terms and possible enrichment to OpenTheso ?
- BN : Make difference between opentheso, the managment system and PACTOLS, the vocabulary-thesaurus. See the Miled Rousset answer right down but any Contributor to PACTOLS is allowed to propose a new concept. Ask for an account. And a tool like Heterotoki could be of great help as it can be harvested by Opentheso (MR: yes, I am interested to discuss it)

(**MR**: with the new version of Opentheso, the candidate module will be completely redone and it is planned to offer this option with discussion)

- (2) And Mapping their terminologies with PACTOLS ?
- >> BN : while using the alignment functions of Opentheso. This needs that those vocabularies are standardized

(**Miled Rousset:** in Opentheso, it is possible to import a lot of alignments from another system (format SKOS))

@MR, could you harvest SKOS alignment, if exported and exposed on the web ? yes

Deal, then!

- 17:44 **Florian Hivert -** ROSER - Répertoire de l'Ornement Sculpté des Édifices de la Renaissance: An interoperable scholarly digital edition for the ornamentation in Renaissance edifice

Blandine Nouvel : 1- do you use any controlled vocabulary for your description and metadata and if yes, could it be mapped to PACTOLS ?;)

2- I missed something… could you precise the standard structure of your data ?

Florian Hivert : 1 - We will map the working terminologies like the Architectural Part and Iconography with "le Thésaurus de l'Inventaire Général" and "Iconclass". So the terms in itselves are local but they'll be linked with standards after the work of the scholar team. I need to ask my scholar team if they are interested in PACTOLS ahah, actually Jean Beuvier made a few tests in Heterotoki, matching his Iconography with PACTOLS. So it is possible !

2 - The standard structure relies on a granularity of classes, from the more precise to the wider structure : Motif -> Oeuvre/Partie d'Oeuvre (if the Oeuvre needs to be separated) -> Edifice. Then there are other properties that link the items with the

terminologies and other items like the Modele and the Personne. (I can give you the whole datamodel if you want, I couldn't put it in my slides).

- 17:54 **Dr Tomasz Panecki -** Digital editions of historical maps: between images and data. The example of "Gaul/Raczyński" topographic map (1807-1812)

http://atlasfontium.pl/index.php?article=gaul

Damien Vurpillot: The fact that you give access to a WMS service helps a lot with working with people from various backgrounds and that maybe not really into GIS. Maybe we should advertise more free web tools available to manipulate and create information (light online GIS tool we could say) over WMS/WFS services.

From experience with historians, it helped a lot with accessing complex spatial data.

Tomasz Panecki: Many people just use Google Earth for simple analyses (or ESRI online map tools). In fact I don't know any opensource/web tool. Still - we can use QGIS….

Loup Bernard : amazing tool, that you developed. If you are looking for some other french applications of the same kind, send a mail.

Tomasz Panecki: Yes, thanks. For now I know http://www.geohistoricaldata.org/

Damien Vurpillot : Custom interfaces for Cesium or Terriajs can work as light GIS for example

- 18:03 **Caroline Parfait & Marlène Arruga -** Development of digital platform about stained glass restoration data in France

   Miled Rousset: do you use thesaurus indexing with Omekas ?

Caroline Parfait :

We have tried to rely on thesaurus such as Thesaurus de l'inventaire du mobilier national, but without success because the stained-glass window is not described

Our work with omeka S is really recent

MR: for info, I developed a module for OmekaS which is included in ValueSuggest, it allows to connect to the thesaurus managed in Opentheso

Caroline: it's interesting, we could indeed need it and test it !

Emilio Sanfilippo: Have you already explored the state of the art about ontologies for glass artwork? I assume that you will need to adapt/extend CIDOC for your modeling purposes. Also, OBDA techniques may allow you to make connections between your relational database and a Semantic Web environment (if this is what you wish to do).

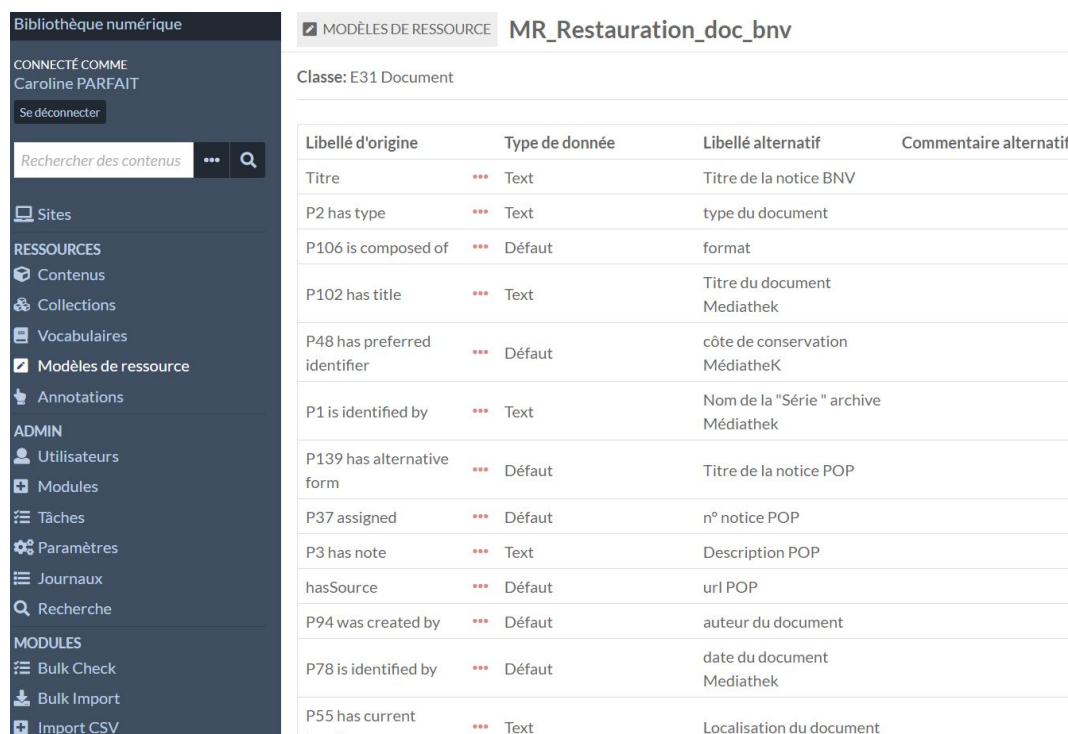**Caroline:** We use cidoc crm to describe documents such as condition reports and restoration "memoire"

I can send you a screenshot

Ok, i did not use odba , we going to find out with Marlène

Here is a screenshot of the template ressource I have started to develop to describe the documents, Class E31 document.

Now, I am creating a model to describe the E11 class modification that correspond to the restoration acts

Marlene Arruga : yes we looked at the state of the art about ontologies for glass artwork but we didn't find a specific web semantic vocabulary to describe restoration data. The best solution that we have found is Cidoc CRM.

Libellé d'origine		Type de donnée	Libellé alternatif	Commentaire alternatif
Titre	•••	Text	Titre de la notice BNV	
P2 has type	•••	Text	type du document	
P106 is composed of	•••	Défaut	format	
P102 has title	•••	Text	Titre du document Mediathek	
P48 has preferred identifier	•••	Défaut	côte de conservation MédiatheK	
P1 is identified by	•••	Text	Nom de la "Série " archive Médiathek	
P139 has alternative form	•••	Défaut	Titre de la notice POP	
P37 assigned	•••	Défaut	n° notice POP	
P3 has note	•••	Text	Description POP	
hasSource	•••	Défaut	url POP	
P94 was created by	•••	Défaut	auteur du document	
P78 is identified by	•••	Défaut	date du document Mediathek	
P55 has current location	•••	Text	Localisation du document	

- 18:13 **Dr Stéphanie Satre -** Epicherchell, an epigraphic dataset to share and publish

Blandine Nouvel: Great project ! it could be used as a model for epigraphic data

Stéphanie: Thank you ! As I said, it is and will always be a work in progress : we have still a lot to do *to* fit with the principles of accessibility, interoperability and reusability. EpiCherchell is a "core' as open as possible that we must align, link to other technologies.

How would you associate transcription to writings visible on the numerous images ?

I think that there is something to do with the IIIF specifications (annotations, manifest) in that way. Do you have ideas ?

We will indeed deepen the possibilities offered by the IIIF standard and its annotation system on these questions. It seems a very interesting way.

Virginie Fromageot: Congratulations Stéphanie ! It is a very rich project ! Did you know that Huma-num has implemented IIIF in Nakala ? I'm not using it. It's just an information.

Bruno Baudoin : Yes, indeed, but for the moment we are using the Image API from our own image server, itself hosted at TGIR with another project, and should stay in this system for EpiCherchell.

- ~~18:23 **Dr Cristina Vertan** - A fuzzy ontology for historical places people and events in historical texts - a case- study of historical texts of Dimitrie Cantemir~~

- 18:33 **Prof. Danielle Ziébelin -** Patrimalp, an integrated and interdisciplinary cultural heritage platform for the French Alps

Emilio Sanfilippo: I would be very interested in understanding how you use mereological/topological theories for your modeling purposes. I assume that you also need to represent material constitution (i.e., the fact that an object is made of a certain amount of matter). Foundational ontologies like DOLCE provide rich axiomatizations for these purposes.

> **Danielle Ziébelin :** Currently we have mainly modeled Part-Whole + Spatial relations. We have characterized some properties of the Part-Whole + Spatial relationships.
>
> Concerning the tangible object concepts, we have indeed MaterialEntity which is currently composed of eight classes. This work is in progress and the point of view taken is that "the matter has a volume", so we have focused again on spatial relationships. I know this is not enough, and when the modelisation of the volume part of the matter will be well defined, we will extend the model to other properties to modelize "amount of matter".

Day Two (Thursday June 18, 2020)

09:00   ARD Programme Intelligence des Patrimoines: Heritage(s) digital ecosystem

(ZOOM Name) / (Real Name) / Question

Damien Vurpillot: The first video we could not show in the first presentation: https://sharedocs.huma-num.fr/wl/?id=aNSJCI0xoXjVFTzPuL7XiWNFJGCXgOBw

Jorge Fins: Thank you for your presentation. It is great to see, notably on this video, the hard job done since your general assembly presentation last year.
    - Half a question/half a suggestion depending on what you did already : For DOI, Do you use Datacite, like DMP Opidor ? If yes, is any connection considered/possible with theirs, for your DMP project form ?
    - About Geonames data reuse, do you use Geonet Names Server or another way ?

Damien Vurpillot: Yes, we are using OPIDOR for the DMP and DOI management, we try as much as possible to use available tools. We are pushing forward the bilingual INRAE template in OPIDOR for the moment but we are thinking about a custom template of our own.
    The idea is to document clearly each part of the DMP for researchers to help them make the DMP even if it's something quite new for them. DMP can really be painful to make.

    For Geoname, we are using the official API. I was not aware of Geonet, I will look into it, thank you for the information.
    Jorge Fins: So it means, you have a Premium Data Subscription ?

Damien Vurpillot: Not yet but we will soon reach the request threshold, this is something we are considering when we move out of dev to production.**Jorge Fins**: Thanks for the answer. Maybe something can be done on a Huma-Num level, there ?

Damien Vurpillot: Yes, I think this service can be useful for a lot of people so that is something we should consider

Jorge Fins: I agree. For instance, I know that a few years ago, Telemeta tool (CNRS, now) used Geonames daily data dumps. Maybe something to discuss with them in order to be more people to negociate together.

09:40   OpenTheso and PACTOLS

(ZOOM Name) / (Real Name) / Question

Emilio Sanfilippo:
Did you reuse existing linguistic resources for the development of these projects? I'm thinking, e.g., about WordNet (see also the general documentation)
    (**M.R**. : no but I'm going to look forward to this project)

ES : Is there any formal specification behind OpenTheseo/PACTOLS? For example, the relation of subsumption between classes (terms) is commonly represented through IS-A semantic.
    (**MR. :** Can you tell me more about your question ?)

> **ES:** from a knowledge representation perspective, one wishes to give a precise meaning to terms within a vocabulary in formal/mathematical terms. For example, when you say in an ontology that the class of Person is subsumed by (IS-A) the class Mammal, this - formally - means that all instances of persons are instances of mammals, too. Differently, in linguistic resources like WordNet, for example, subsumption has a different meaning because terms do not stand for classes of entities, they only stand for words within a linguistic system. This is why I was asking about the specific meaning of taxonomic relations in your system (from @Blandine's comment, I understand that the semantic is given in terms of SKOS).
> A good reference for these discussions is: Hirst, Graeme. "Ontology and the lexicon." In Handbook on ontologies, pp. 269-292. Springer, Berlin, Heidelberg, 2009.

Blandine Nouvel: @Emilio: PACTOLS is a normalised thesaurus, with the ISO 25964 standard for thesaurus (multilingual and interoperability) and uses SKOS language. So, as I said yesterday, concepts are linked with hierarchic, associated and related relations. We use NT, sometimes NTi.
Vocabulary is controlled by researchers and chosen in bibliographic resources. We precise definition of each concept with expert communities and sometimes cnrtl or wikipedia. But we are very cautious with general resources. For example, WorldNet doesn't know flake as a lithic tool and its definition would need to be contextualised with archaeological domain.
    Translation stays a real issue that we try to solve with bibliographic resources and validation by archaeologists.

Emilio: Thanks, what is NT(i)?
Blandine: Narrow term, (i) is for instancial relation. It corresponds to a dependant term like cratère de Vix is a NTi of cratère

Emilio: ok, thanks, they are both very interesting projects indeed!

Loup BERNARD: amazing ! I'm looking forward to test it ! How do You decide what thesaurus can be aligned ? Does it need checking from your side ?

MR: each thesaurus can be aligned with the pre-programmed alignment sources in Opentheso

Blandine : aligned thesaurus have to be of a similar interest,of course. PACTOLS has chosen AAT for example for thematic coverage. Insee and Geonames for the micro-thesaurus LIEUX (among others) and we are working to align Anthroponyms with IdRef.
I don't understand your 2nd question ?...

Loup BERNARD: how do You solve disambiguation issues / control choices that are made ?

Emilio: This is a very interesting question!

Blandine : we solve it with different linking relations available with SKOS language, that is: closermatch, narrowermatch, broardermatch etc.. and if there is no match, we don't align concepts.

(**MR**: at the time of alignment, we also retrieve a definition and a URL that allows us to see the concept in its environment to be sure of our choice)

(**MR:** Opentheso will be translated into 4 languages (English, German, Italian and Spanish), a first stable version will be available at the end of june)

10:20   A Data Management Plan Template for Archaeology

(ZOOM Name) / (Real Name) / Question

Blandine Nouvel : Great work! Do you already know how many times the DMP has been downloaded ?

Nathalie Le Tellier-Becquart : The DMP OPIDoR platform enables us to know the usage statistics. Laurence Rageot, from the MASA consortium, has organized a training session in February 2020 at the house for social sciences in Tours, so these statistics are not only about the DMP template's use for research projects but can also give us information about its use in a training context (if 6 DMPs using the MASA template have been created at the same time, for instance). Between February and May 2020 the MASA DMP French template has been downloaded 15 times.

Loup BERNARD : Great work indeed ! You showed an english interface, is a french version available ?
Maybe I missed the information, but how is licensing ? Creative Commons or ETALAB are present ?
Perfect !

Emmanuelle Morlock : About the licencing, the template doesn't prescribe any licence because it's a choice that depends on the nature of the data, the requirements of the funders or the choices of the project's team. But it's important to document these choices in the DMP. You can do it globally for the whole set or at the dataset level. For example, you might want to use the field "Reuse conditions" to indicate the applicable licence. It's in the Dataset tab, section "Description of data". You might also want to say something about that in the section for describing the dissemination action items, let's say in the "conditions for reuse field". You will find some guidance there about licensing, but it's really up to the user. See a screenshot below:

Nathalie Le Tellier-Becquart : A French version has been available since February on the DMP OPIDoR platform (https://dmp.opidor.fr/). The English translation of the DMP template presented today is still ongoing and will probably be published online before September 2020. We thank the INIST translation service for it!

You can use creative commons or ETALAB licences if you wish (you have to fill in the information in the template).

10:50   OpenArcheo

(ZOOM Name) / (Real Name) / Question

Julian Richards: With regards to the question about whether OpenArchaeo is being promoted by ARIADNEplus

In ARIADNE we see our role as facilitating conversations and discussion between those working in cognate areas – so we have a working party (in which MASA participate) looking at various European efforts at developing ontologies with CIDOC-CRM for fieldwork recording – it currently includes representatives from partners in France, Italy, UK, Greece, Sweden and Norway.

Olivier Marlet : Indeed, this working group was launched on Tuesday 16 June. It is therefore currently in the process of being set up. This will be at a more detailed level than OpenArchaeo, whose vocation is to query simultaneously several archaeological data sets at different scales (regional inventories such as AERBA or field records such as ARSOL). The working group will try to see how to model more specifically field archaeological data but also how these data are mobilized to be published or to feed a scientific publication (in logicist format for example based on the reasoning modeling of the CIDOC CRMinf).

Concerning Sparnatural, the intuitive natural language query system to generate a SPARQL query, being also open source, it could also join the Ariadne+ infrastructure as such to be exploited by any SPARQL query platform.

SESSION 4: 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**

(ZOOM Name) / (Real Name) / Question

Jorge Fins / Jorge Fins / Many thanks for your presentation. Just as a reminder of the different questions I had in mind, and many were answered :
- Can this work lead to some help/recommendations for writing descriptions as a midlevel between fulltext and specific metadata ?

  SARA Tonelli: The issue with using machine learning and word embeddings (without additional features) is that it does not provide insights into what worked and what characteristics good descriptions have in common. Maybe we should add a further processing step that analyses good descriptions and highlights commonalities (for example presence of controlled vocabularies, structure, etc.)

- What about time spent that may be different based on different media/disciplines ?

  SARA Tonelli: If you mean the time spent by domain experts to manually label descriptions, there are surely differences, which however depend more on 1. Length of descriptions to be analysed and 2. Level of expertise of the human annotator.

- Have you in mind other domains for the future ?

  SARA Tonelli: The approach is rather language- and domain-agnostic, provided that enough training examples are available to train the classifier. We do not plan to add other domains because our results seem to be consistent across the three we tested, but if anyone would like to test the approach on a new domain we are happy to share the code and give help.

- Can different periods be distinguished (for instance, early modern period, with more exchange between domains) ?

  SARA Tonelli: It would be possible to apply any further distinction, provided that the training data are annotated also with this information. For example, if descriptions refer to artifacts from a certain period, and this information is explicitly associated with the

description, it is possible to further divide the data into subgroups and check which effect the time period may have.

- What about the modern challenge of interdisciplinarity (cross-domains) ?

    SARA Tonelli: Works like this are inherently inter-disciplinary. A domain expert is needed to collect and label the descriptions, while machine learning and NLP expertise are needed to run rigorous experiments. I must admit I have not seen many works in which such skills are effectively combined.

Beatrice Markhoff (beatrice.markhoff@univ-tours.fr): Thank you for this presentation. Regarding a part of the questions by Emilio and your answers, to whom it may interest, with my colleagues Arnaud Soulet and Arnaud Giacometti we are also investigating knowledge base (KB) completeness [1][2] using statistical methods, in particular in [2] we proposed an efficient algorithm for computing significant maximum cardinalities from the KB content. We conducted experiments on some very large Web Knowledge bases such as Dbpedia or Europeana, but also on a COINS knowledge base that Maria Theodoridou kindly provided us (see [3] only in French, sorry). We are currently building a website for providing a more easy-to-test version of this program, I could notify interested people when it will be online. We are also currently working on a solution for automatically extracting comparison tables from a given KB, this could be useful for KB profiling (helping exploring the content of a KB). The corresponding article is only submitted for now, but this site can give an idea (the Search allows to access to already generated comparison tables, you can try for instance « Paris »): https://lovelace-vs-turing.com/

Sara Tonelli: Very interesting link, thank you! I will check the references you have provided and let you know if we can come up with some ideas concerning the issue below.

BM: We are very interested to know if someone think that such automatic generation of comparison tables (for comparing a small set of entities) would be useful for applications in the Heritage field, if you have any idea about that please contact me, thank you.

[1] Arnaud Soulet, Arnaud Giacometti, Béatrice Markhoff, Fabian M. Sushanek : Representativeness of Knowledge Bases with the Generalized Benfod's Law, International Semantic Web Conference (ISWC 2018), October 2018, Monterey CA, USA, pp. 374-390. I can provide it on demand
[2] Arnaud Giacometti, Béatrice Markhoff, Arnaud Soulet : Mining Significant Maximum Cardinalities in Knowledge Bases, International Semantic Web Conference (ISWC 2019), October 2019, Auckland, New Zealand, pp. 182-199. I can provide it on demand
[3] http://ceur-ws.org/Vol-2133/cnia-paper6.pdf

Emilio: @Beatrice, thanks a lot for your comments and the references

SESSION 5: INFORMATION SYSTEMS FOR 3D DATA AND SEMANTIC ANNOTATIONS

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

    (ZOOM Name) / (Real Name) / Question

Ana Roxin : What is the language used for modelling the semantic data (e.g. RDFS / OWL 1 or OWL 2 what profile) ?

> **Roland**: Semantic data = data from the user, so it's not data in the sense of Semantic Web.

> **Roland:** Yes. But I am waiting news from one colleague to get a more comprehensive answer

> **Ana Roxin** : Thank you.

> **Roland**: (confirmation) semantic data are stored in a document database (MongoDB). No graph were used in this application

Emilio Sanfilippo: are there any relations between the semantic database and the spatial database? For example, are the 3D data semantically annotated?

> **Roland :** The link between the "semantic" info and the 3D point cloud is done through a tag stored in the system. From the user interface you have a "3D view" button. When you click on it you open the 3D viewer and you can see the tag on the point cloud. And on the other hand, you can tag the model or make a selection in the 3D point cloud and associate it to a "semantic" object.

ES: Have you experience with the use of machine learning methods to automatically classify/recognize entities in point-cloud data?

> **Roland:** Yes, Florent Poux is actively working in this field with very good results. I am going the put some references here : https://medium.com/@florentpoux; https://pointcloudproject.com/

Answer from **Florent Poux**: oui. on peut obtenir de bon résultats si les données d'entrainement sont de bonne qualité en supervisé. Mais je recommande la partie self-learning qui contribue bien   plus   à   l'évolution   du   domaine.   cf

https://towardsdatascience.com/fundamentals-to-clustering-high-dimensional-data-3d-point-clouds-3196ee56f5da

Florent Laroche: You used Geoverse. There are other technologies on the market; why this choice? Advantages and inconvenients?

> **Roland:** at that time, it was one of the best solutions (looking to the specification of the Tender) and we knew it. But it was in 2018.

Florent: Is software solution of Florent Poux opensource?

> **Roland:** Yes, you can have access to a lot of ressources: I am going to put some references here. https://medium.com/@florentpoux; https://pointcloudproject.com/

Answer from **Florent Poux**: les articles sont dispos en open access
https://www.mdpi.com/2220-9964/8/5/213
https://www.mdpi.com/2072-4292/10/9/1412
https://www.mdpi.com/2076-3263/7/4/96
une partie sera sous peu en open-source (nrmlt sept 2020 chez elsevier ISPRS)

Damien Vurpillot (comment): Point cloud selection seems really robust, in my opinion it's one of the most difficult challenges to tackle. Is it working on "semantic" subsets (like "Wall X" which would be identified afterwards by users) of the data which are byproducts of the "global" point cloud merging all acquisitions?

Roland: I would like to say yes, but could you refine your question?

Damien Vurpillot : I was wondering, do you select point cloud (like the stairs we saw in the video) based on a subset of the point cloud, either:
- because you can only select within the subset that was created by users
- or because the selection is created by selecting all points from a class from your smart point cloud (like a lasso tool from photoshop which can select the entire class even you roughly select the "semantic object" on the fly)

So far, the only robust tool I was able to find for point cloud, if you don't have a comprehensive classification of all objects within the point cloud, is inherited from research linked to astrophysic and how to select clusters of stars and such.

Roland: in the video, it is the coudenberg system with point cloud data managed with Geoverse. The selection (stairs) is made by the user and he(she) can link it to an object "stairs" in the "semantic" database.

It is different with the smart point cloud. In that case, if the concept "stairs" is accessible (through an ontology) you could classified a subset of points (aggregate points) as "stairs". Do not hesitate to see Florent Poux works (link above in the discussion).

>**Damien Vurpillot:** Thank you, I will read the document. I have been looking at Euclideon's work (as a company) for a while and that's the first time (to my knowledge) I see Geoverse used in this context (cultural heritage management). Really interesting.

Answer from F**lorent Poux**: cf https://www.mdpi.com/2220-9964/8/5/213
dans le cas présent (coud) indirectement, mais pour le futur possible. Je finalise un article la dessus avec Fabio, prévu pour fin Août normalement (multi-level unsupervised semantic segmentation), justement appliqué à des données du patrimoine très fines. c'est en lien avec les techniques de machine leearning sur des ensembles non euclidien (graph-learning).

- **14:50   Dr Florent Laroche - French National Project ReSeed - Semantic reverse-engineering of digital heritage objects**

(ZOOM Name) / (Real Name) / Question

Emilio Sanfilippo: can you comment on the use of semantic models in tandem with 3D models? You mentioned the use of Dublin Core, for annotating entities within 3D models?
What does "ontological freedom" mean? (I quote from slide n.20)

Florent: Wants to know more about research results
PhD defense of Loic JENSON - Monday 22 june 2h GMT Paris (14h)
https://ec-nantes.zoom.us/j/95406685635
"Apports à la construction d'outillage informatique pour les travaux Patrimoniaux. Analyse d'un carrefour disciplinaire et cahier des charges pour une méthodologie d'intégration de connaissances hétérogènes"

Jorge Fins : Do you study users feedbacks with the Museum ?

Florent: Yes you can read 2 scientific communications we have done:
- about Nantes1900 project: https://hal.archives-ouvertes.fr/hal-01243990v1
- Another one with a colleague in social sciences "Vers des applications numériques "durables" pour les institutions patrimoniales" https://hal.archives-ouvertes.fr/hal-01556585v1

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

    (ZOOM Name) / (Real Name) / Question

    Emilio Sanfilippo: I'm wondering whether there is any research work using IFC(OWL) for annotating point-cloud sets, recalling Roland's and Florent's talks

        => Damien Vurpillot: Maybe not so much in cultural heritage compared to urban/building management modeling and such?

    ES: I'm also wondering whether you can use ontologies/rule systems to check the consistency of a spatial model with respect to domain knowledge (e.g., inconsistencies with respect to geometric/topological knowledge)

    I remember that IFC reifies a lot of relations between classes, and this could pose problems from a Semantic Web perspective because you get quite large A-box models. Is this a problem for you? Has been this solved?

    CIDOC extension for BIM:
    Ronzino, Paola, Franco Niccolucci, Achille Felicetti, and Martin Doerr. "CRMba a CRM extension for the documentation of standing buildings." *International Journal on Digital Libraries* 17, no. 1 (2016): 71-78.
    http://www.cidoc-crm.org/

    Ronzino, P., Niccolucci, F., & D'Andrea, A. (2013). Built Heritage metadata schemas and the integration of architectural datasets using CIDOC-CRM. In *Online proceedings of the International Conference Built Heritage*.

    Roland : I wrote an ontology of space ;-), used recently for the new version of CityGML.
        Billen, R., Zaki, C. E., Servières, M., Moreau, G., & Hallot, P. (2012). Developing an ontology of space: application to 3D city modeling. In T., Leduc, G., Moreau, & R., Billen (Eds.), Usage, Usability, and Utility of 3D City models. Les Ulis, France: edp Sciences.

        http://smartcity.linkeddata.es/

        CityGML 3.0: New Functions Open Up New Applications

        https://link.springer.com/article/10.1007/s41064-020-00095-z

(Currently speaking) SESSION 6: DIGITAL DATA SERVICES FOR CULTURAL HERITAGE

- 16:15   Dr Livio De Luca

(ZOOM Name) / (Real Name) / Question

Florent Laroche: Can you imagine to use the Smart Semantic software recognition of Rolland Billen (University of Liège) to couple with your system?

Roland Billen: is the 3D viewer you developed open source?

Damien Vurpillot (comment) : Really happy and impressed to see how you make it possible to complement the comprehensive point cloud viewer (in Potree) with Aioli subsets and semantic annotations. They seem to work great together.
Something really interesting would be to extend the classification system inherited from Lidar las format, which is limited to a few specific categories, to take into account other kinds of classification. It seems that Markus (dev for Potree) is working on something in this regard.

Also, https://entwine.io/ is interesting for managing massive (billions and more) point clouds. It is currently used to display country lidar viewer (netherlands, danemark, etc.).

Florent Laroche suggestion for @Livio + @Roland: for those researches: using Potree as we all use it we can implement native Augmented Reality with next API WebXR (Standard W3C).
=> **Damien Vurpillot:** We also have developed a branch for VR but it's a bit tricky right now because of concurrent VR web model (mozilla and the other one). Markus is working on another branch to take advantage of GPU capacities on the web coming in next browser releases (currently in beta in chrome and mozilla).

Livio: thank you for all these inputs. I'm really interested in starting to discuss these perspectives concerning the multi-layered caracthérization of point-based representations. @Florent :yes, it should be very relevant; @Roland: we intend to publish the sources for the viewer and aioli (when finished..) ; @Damien: we're in touch with Markus about this point as well as for the managing of several segmentation layers … to visualise multi-layers 3D annotations/segmentaions coming from 'aioli'

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

(ZOOM Name) / (Real Name) / Question

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

(ZOOM Name) / (Real Name) / Question

Emilio Sanfilippo:
I would be interested in having a look at the mapping guidelines/approach you mentioned during the talk. Could you please add a link?

Olivier Marlet : I don't know if you can access the D4Science ARIADNE+ plateform : https://data.d4science.net/EyXn If not, I can share the PDF.

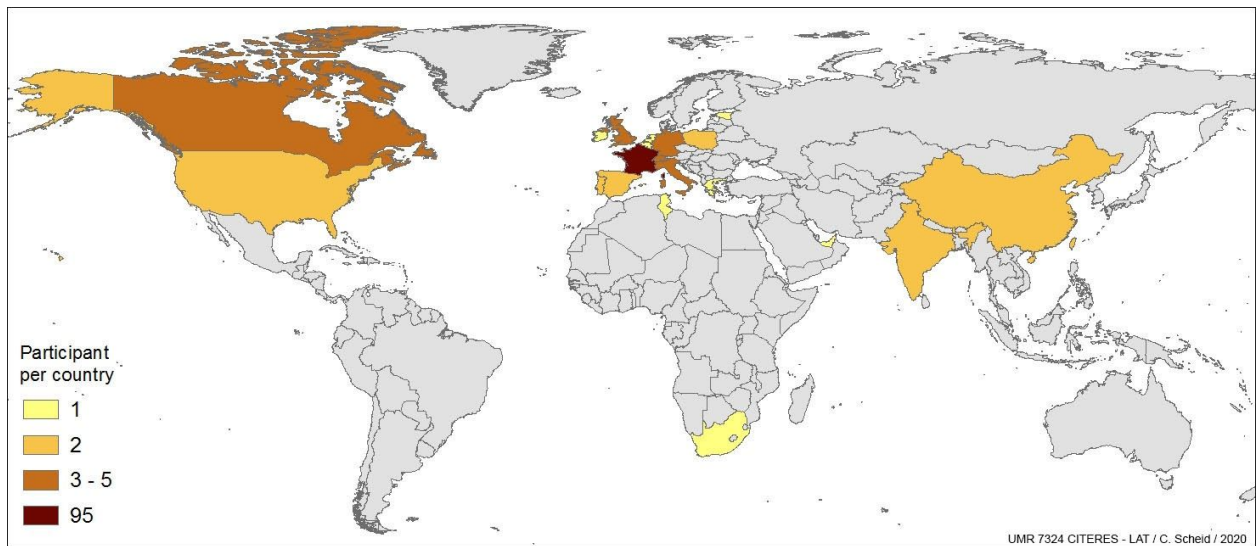**ES:** if you can share the pdf, that would be great

ES: Indeed, working with ontologies like CIDOC can be complex from a domain expert's perspective. Apart from ontologies, the community should also produce common guidelines for mapping datasets to CIDOC.
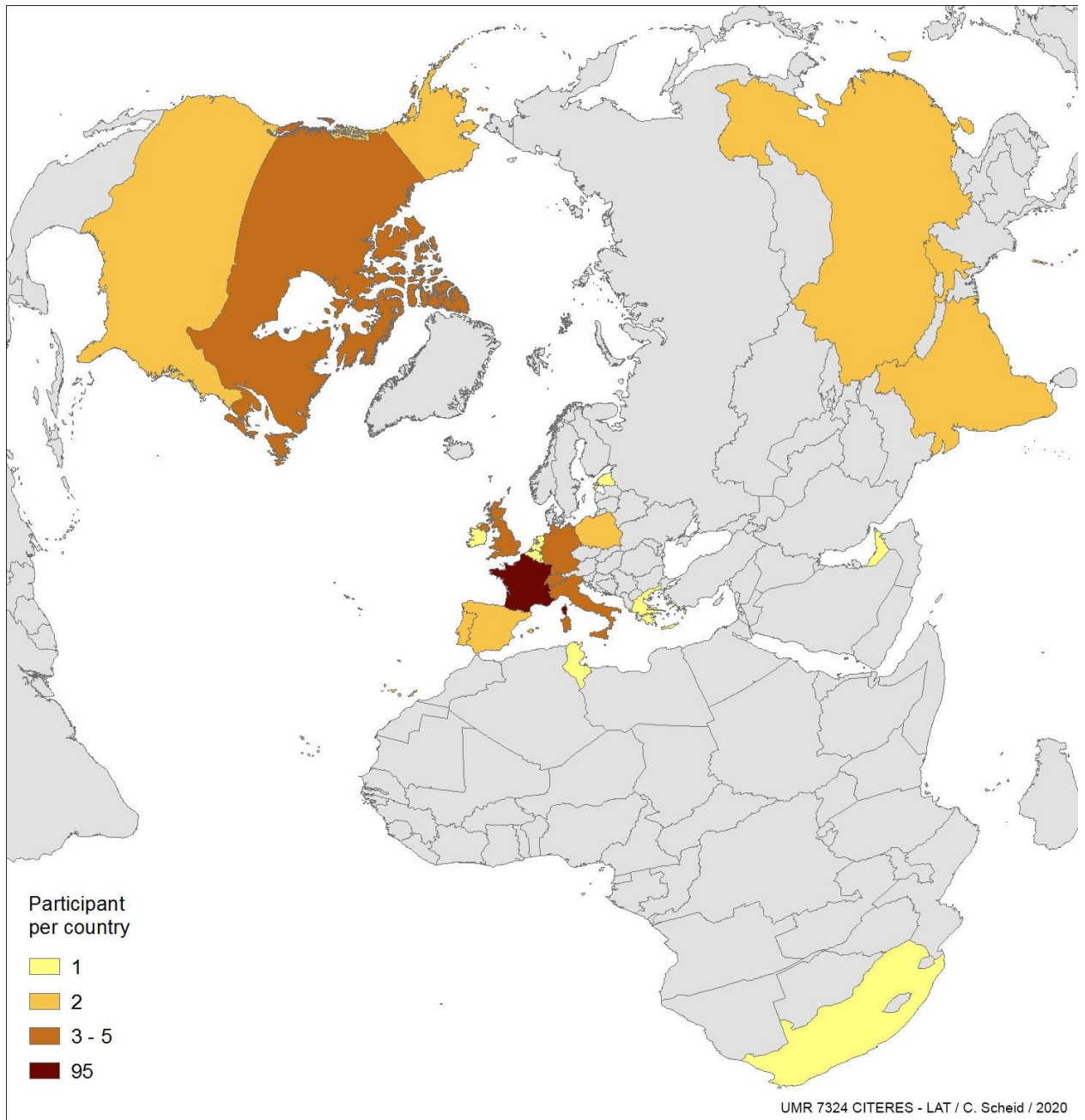
Loup Bernard : You rightly mentioned that there is a need for budget/resources for data cleaning : who pays and mostly who do You hire ?

Do You plan working on other languages as well ?

Beatrice Markhoff: Thank you very much for this presentation. I would like to know whether there is a public URL for a program to query the SPARQL Endpoint?

Provenance of the registered participants to the conference (classical and polar projection)

Participant per country
- 1
- 2
- 3 - 5
- 95

UMR 7324 CITERES - LAT / C. Scheid / 2020

Participant
per country

- 1
- 2
- 3 - 5
- 95

UMR 7324 CITERES - LAT / C. Scheid / 2020

Pictures of the day!

Day 1

Day2