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

Emilio M. Sanfilippo[1] and Xavier Rodier[2]
[1]Le Studium Research Fellow at CESR University of Tours, France
[2]UMR 7324 CITERES-LAT CNRS/University of Tours

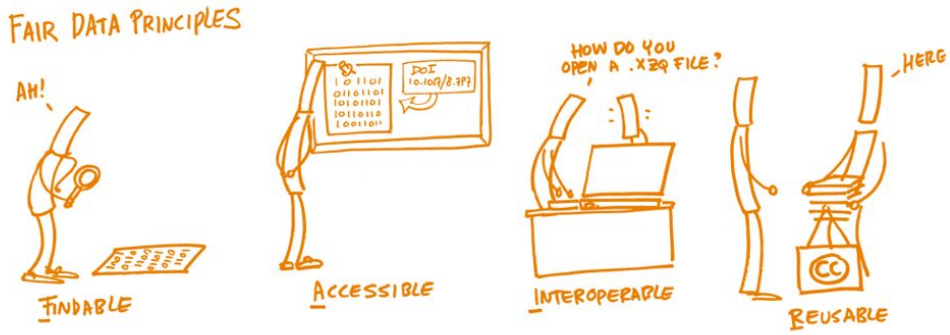
Data management for cultural and natural heritage

Scholars in both the natural sciences and the social and human sciences working about natural/cultural heritage currently assist to the **proliferation of**:

- Data produced by research efforts
- Controlled vocabularies (ontologies included)
- Tools and digital services
- Policies and practices

Proliferation but also fragmentation …


European Open Science Cloud
open science
Metadata
Sustainability
social and human sciences
natural heritage
Literature
semantic technologies
cultural heritage
Music
Semantic Web
History
Accessible
life sciences
Reusable
Data-intensive science
Digital Humanities
Findable
Geospatial systems
Interoperability
FAIR principles
Archeology
History of Art
Scholarly Editing
Digital Culture
WordItOut



Credit: OpenAire

The FAIR principles

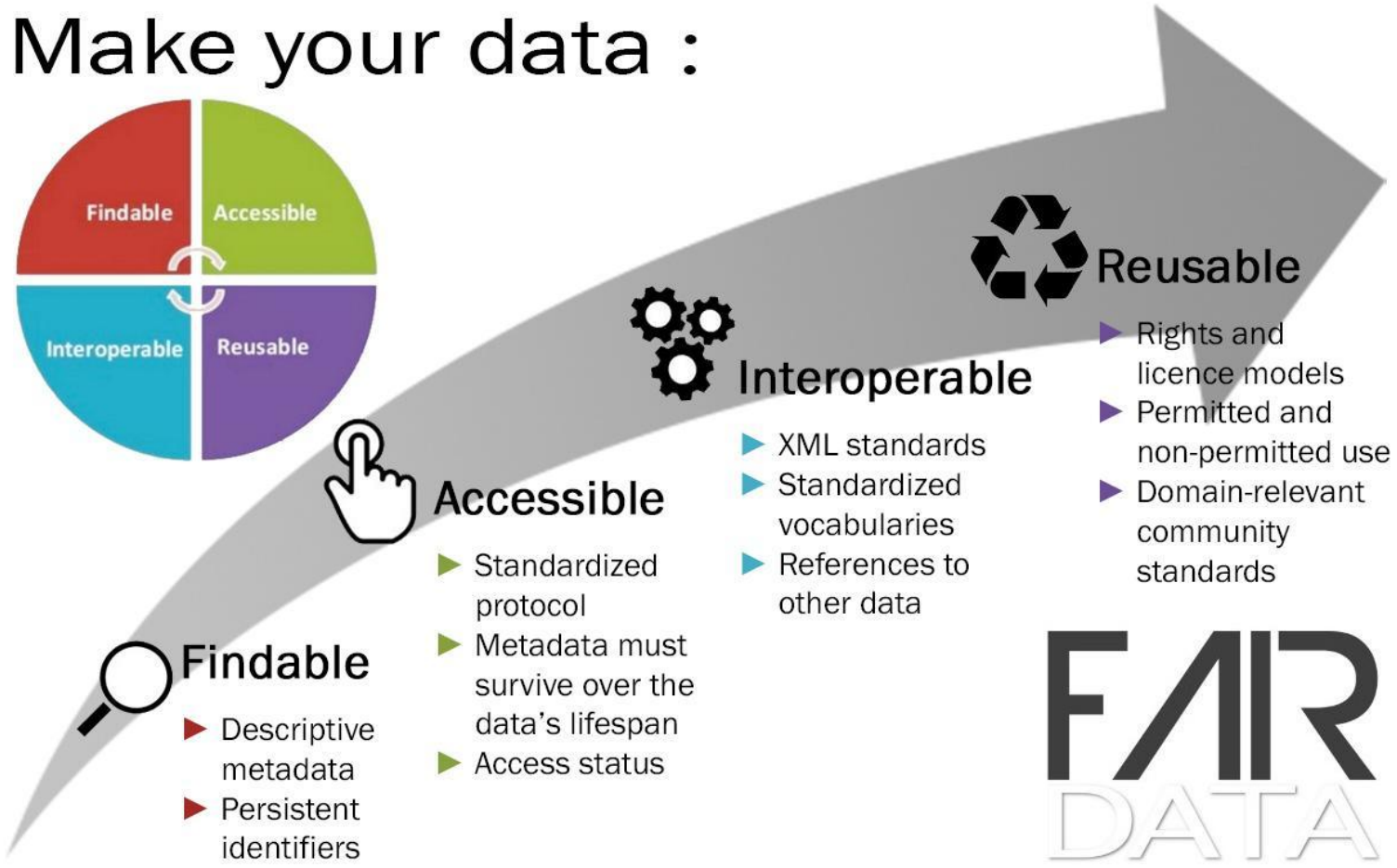
FAIR principles: The origins

Workshop on ***Jointly Designing a Data Fairport***, Leiden, Netherlands, 2014, definition of:

> "Minimal set of community-agreed guiding principles and practices [by which] stakeholders and scholars could more easily discover, access, appropriately integrate and reuse [...] the vast quantities of information being generated by contemporary data-intensive science."

Wilkinson, Mark D., et al. "**The FAIR Guiding Principles for scientific data management and stewardship**." *Scientific data* 3 (**2016**) https://doi.org/10.1038/sdata.2016.18

The 4 foundational FAIR principles



FAIR principles

The FAIR guiding principles: https://doi.org/10.1038/sdata.2016.18

To be Findable:

F1. (meta)data are assigned a globally unique and persistent identifier

F2. data are described with rich metadata (defined by R1 below)

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

A1. (meta)data are retrievable by their identifier using a standardized communications protocol

A1.1. the protocol is free, open and universally implementable

A1.2. the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation

I2. (meta)data uses vocabularies that follow FAIR principles

I3. (meta)data include qualified references to other (meta)data

To be reusable:

R1. (meta)data are richly described with a plurality of accurate and relevant attributes

R1.1. (meta)data are released with a clear and accessible data usage license

R1.2. (meta)data are associated with data provenance

R1.3. (meta)data meet domain relevant community standards

Originally from:
Wilkinson, Mark D., et al. "**The FAIR Guiding Principles for scientific data management and stewardship**." Scientific data 3 (2016).

Remarks

1. High-level, domain-independent principles; they can be applied to a wide range of scholarly outputs;

2. Precede implementation choices, i.e., the FAIR principles do not suggest any specific implementation setting;

3. Provide guidance without being a standard, e.g., in the sense of an ISO, EN or W3C standard.

Implementation considerations (F)

Principle F - Findable

F1. Use of persistent and globally unique identifiers (e.g., URIs) behind single projects and communities.

Example: use of Digital Object Identifier (DOI) - DOIs enables also **access**; use of URIs for metadata vocabularies and ontologies (e.g., PURL)

F2. Use of rich metadata vocabularies to describe and index digital resources

Example: Dublin Core; FAIRsharing [repository]

Implementation considerations (A)

Principle A - Accessible

A1. Digital resources are retrievable by their identifier using a standardized, free, and universally implementable communication protocol

Example: use of URIs with the Hypertext Transfer Protocol (HTTP)

Lannom et al. (2020): a study across **115** European **natural science museums** shows more than **100 commercial** and **in-house solutions** are in use:

"This makes it difficult to create a seamless virtual collection and an approach that aggregates multiple heterogeneous collection management system inputs is needed"

Implementation considerations (I)

Principle I - Interoperable

I1. (Meta-)data use a formal, accessible, shared, and broadly applicable language for knowledge representation

Example: use of **W3C**'s recommendations for machine-processable knowledge representation, i.e., the Resource Description Framework (RDF), the RDF Schema (RDFS), and the Web Ontology Language (OWL)

I2. (Meta-)data use vocabularies that meet FAIR principles

Example: from **thesauri** and **metadata vocabularies** (e.g., DublinCore) to **axiomatized ontologies** (ArCo, etc.)

Implementation considerations (R)

Principle R - Reusable

R1.1 (Meta-)data are released with a clear and accessible data usage license

Example: CC0 1.0 Universal (CC0 1.0) Public Domain Dedication

R1.2 (Meta-)data are associated with detailed provenance (e.g., how/why the resource was generated, by whom, using what fundings, who should be given credits, etc.)

Example: PROV-template [PROV-enance]

The FAIR principles at the European Commission

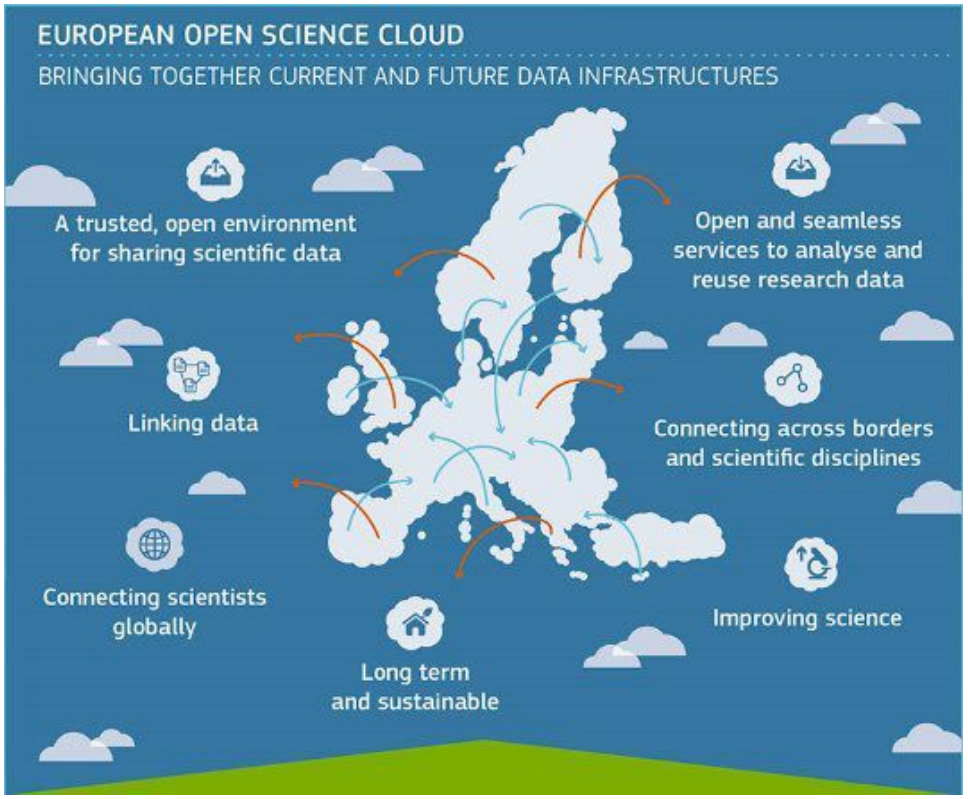
European Open Science Cloud

European Open Science Cloud (portal) [2015-2020]:

- Developing an infrastructure providing its users with services promoting open science practices

See also:

- FAIRsFAIR project



European Open Science Cloud Declaration

- **Data culture** - European science must be grounded in a **common** culture of data stewardship.
- **Open access by-default** - All researchers in Europe must enjoy access to an open-by-default, efficient and cross-disciplinary research data environment supported by FAIR data principles.
- **Research Data Repositories** - Scientist must be able to find, re-use, deposit and share data via trusted data repositories that implement FAIR data principles and that ensure long-term sustainability of research data across all discipline.
- **Skills** - The necessary skills and education in research data management, data stewardship and data science should be provided throughout the EU as part of higher education, the training system and on-the-job best practice in the industry.
- **Semantic Layer** - Research data must be both syntactically and semantically understandable, allowing **meaningful** data exchange and reuse among scientific disciplines and countries.
- **Fair Data Governance** - To make FAIR data a reality, it is imperative to engage stakeholders and relevant multipliers, based on a solid stakeholder engagement strategy, on inter-institutional arrangements, well-established frameworks and decision making flows.

Turning FAIR into reality

In implementing the FAIR principles, high priority to:

- The development, refinement and adoption of **shared** vocabularies, ontologies, metadata specifications and standards which are central to interoperability and reuse at scale

From: Turning FAIR into reality, Report of the Commission FAIR Data Expert Group (FAIR Data EG)

The FAIR principles for cultural/natural heritage

The FAIR principles for cultural/natural heritage

International initiatives and consortia supporting digitally-enabled research and training modules covering the adoption of the FAIR principles

Examples:

- **Natural heritage and life sciences, e.g.,** BBMRI-ERIC, EPOS, ICOS, IS-ENES3

- **Cultural heritage, e.g,** DARIAH, Huma-Num, Consortium of European Social Science Data Archives (CESSDA) [see, e.g., CESSDA's data management guide]

The FAIR principles for cultural/natural heritage

Web services and repositories for sharing vocabularies and/or data

Some examples:

- **Natural heritage and life sciences**, e.g., BioPortal, AgroPortal, OLS, OBOFoundry
- **Cultural heritage** (including **Archeology**, **Social and Human Sciences**): BARTOC, Ariadne, ArCo, Biblissima, data.bnf, OpenArcheo, Europeana [see how Europeana meets the FAIR principles, link]

The FAIR principles for cultural/natural heritage

Use of Semantic Web ontologies

Examples:

- **Natural heritage and life sciences**: Foundational Model of Anatomy, Gene Ontology, Chemical Entities of Biological Interest Ontology, etc.

- **Cultural heritage**: CIDOC Conceptual Reference Model, FRBR, ArCo, Europeana Data Model, CESSDA Catalog, CHARM, etc.

Some references

1. Garijo, Daniel, and María Poveda-Villalón. "Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web." arXiv preprint arXiv:2003.13084 (2020).
2. Guizzardi, G.. "Ontology, ontologies and the “I” of FAIR." Data Int. (2020): 181-191.
3. Jacobsen, A., et al. "A generic workflow for the data FAIRification process." Data Int. (2020): 56-65.
4. Jacobsen, A., et al. "FAIR principles: interpretations and implementation considerations." Data Int. (2020): 10-29.
5. Joffres, A., et al. "The Impact of FAIR Principles on Scientific Communities in (Digital) Humanities. An Example of French Research Consortia in Archaeology, Ethnology, Literature and Linguistics." 2018.
6. Lannom, L., Koureas, D., & Hardisty, A. R. (2020). FAIR data and services in biodiversity science and geoscience. Data Intelligence, 122-130.
7. Wilkinson, M. D., et al. "The FAIR Guiding Principles for scientific data management and stewardship." Scientific data 3 (2016).

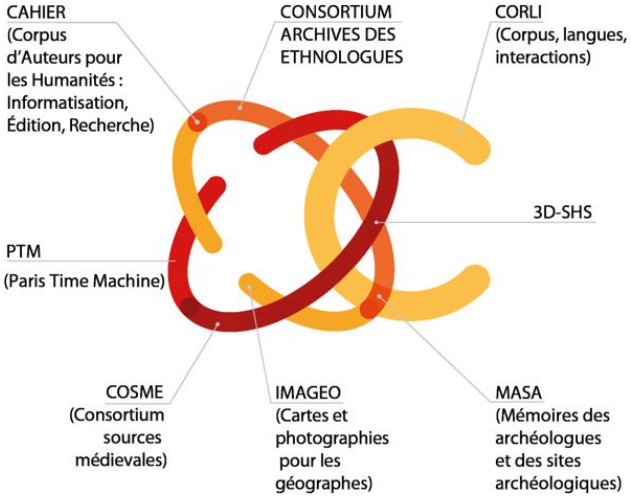
Some references

See also:

- Data Intelligence, vol. 2, no. 1-2. Special Issue on *Emergent FAIR Practices*. Issue Editors: Barend Mons, Erik Schultes & Annika Jacobsen
- Harrower, Natalie, et al. Sustainable and FAIR Data Sharing in the Humanities. ALLEA-All European Academies, 2020.

MASA CONSORTIUM

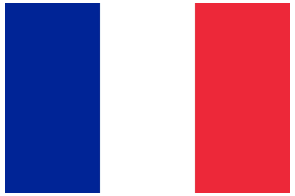
Mémoire des archéologues et des sites archéologiques



Huma-Num

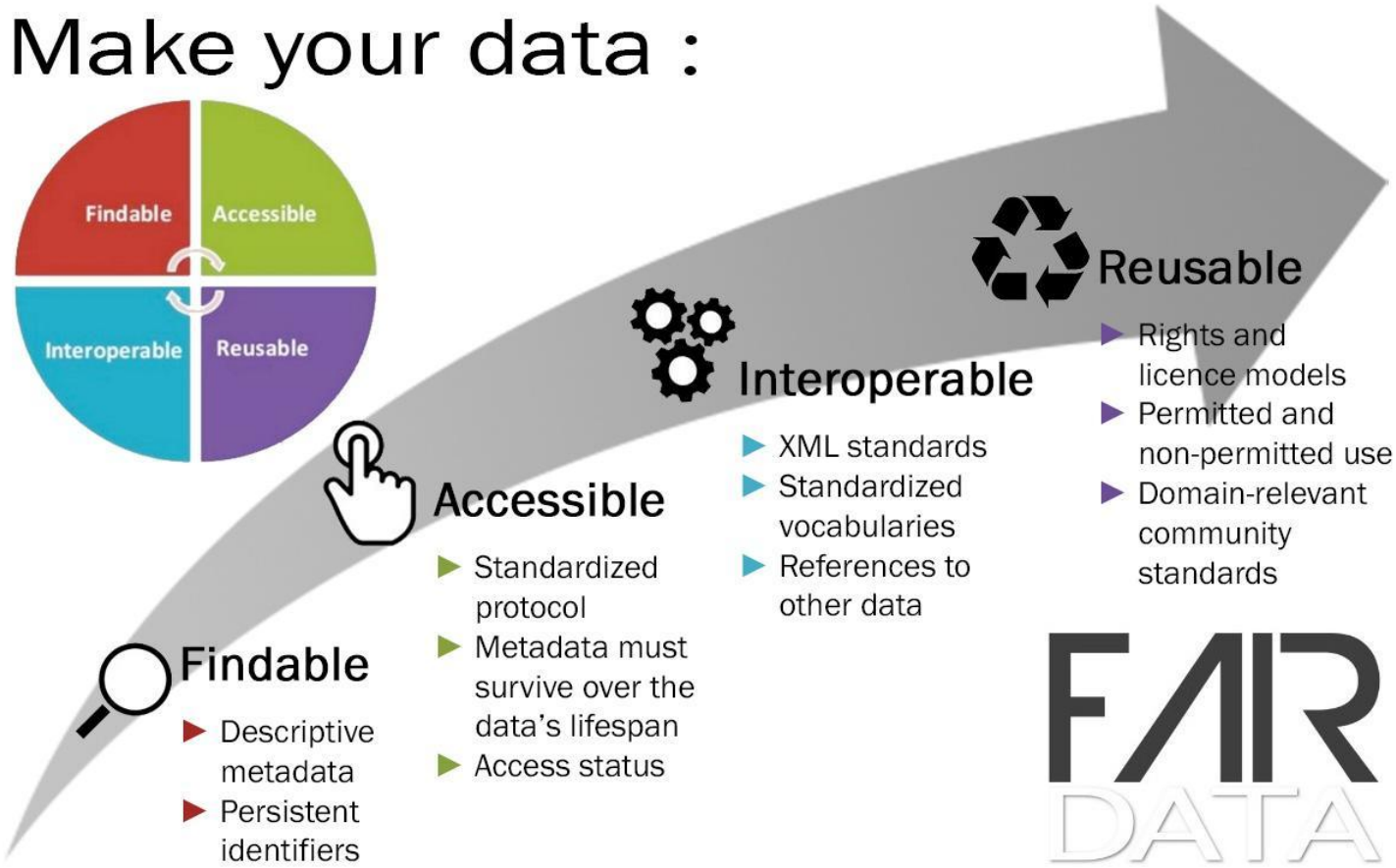
- la Maison des Sciences de l'Homme du Val de Loire, Orléans-Tours
- la Maison de la Recherche en Sciences Humaines, Caen
- la Maison de l'Orient et de la Méditerranée, Lyon
- la Maison Méditerranéenne des Sciences de l'Homme, d'Aix-en-Provence
- la Maison de Sciences de l'Homme Mondes, Nanterre
- la Maison Interuniversitaire des Sciences de l'Homme Alsace, Strasbourg
- le Musée d'archéologie nationale, Saint Germain-en-Laye
- le Groupement de Services Frantiq
- l'INRAP (depuis le 1er janvier 2017)
- le Réseau des écoles françaises à l'étranger

INTERFACING WITH NATIONAL AND INTERNATIONAL INFRASTRUCTURES



FAIR DATA PRINCIPLES

MASA disseminates the FAIR principles to the French archaeological community.



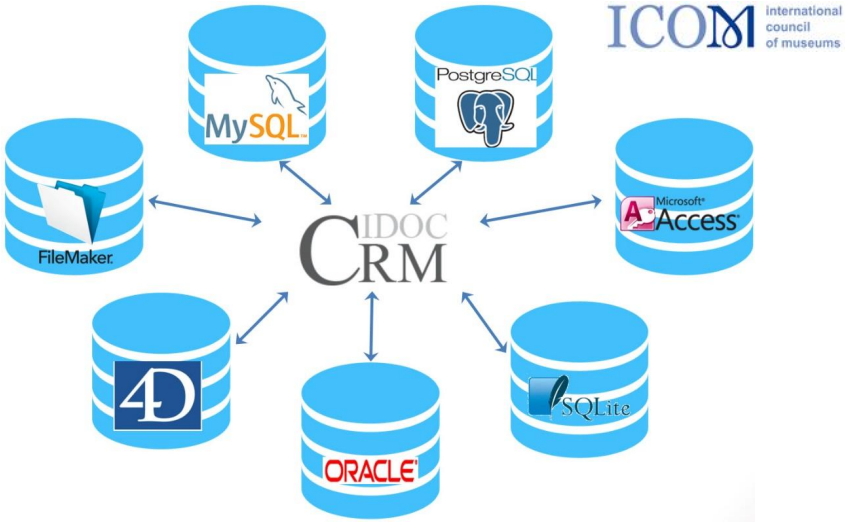
GAZETTEERS AND ONTOLOGY

- Uniform description (data and metadata)
- Interoperability

GeoNames

PeriodO

PACTOLS

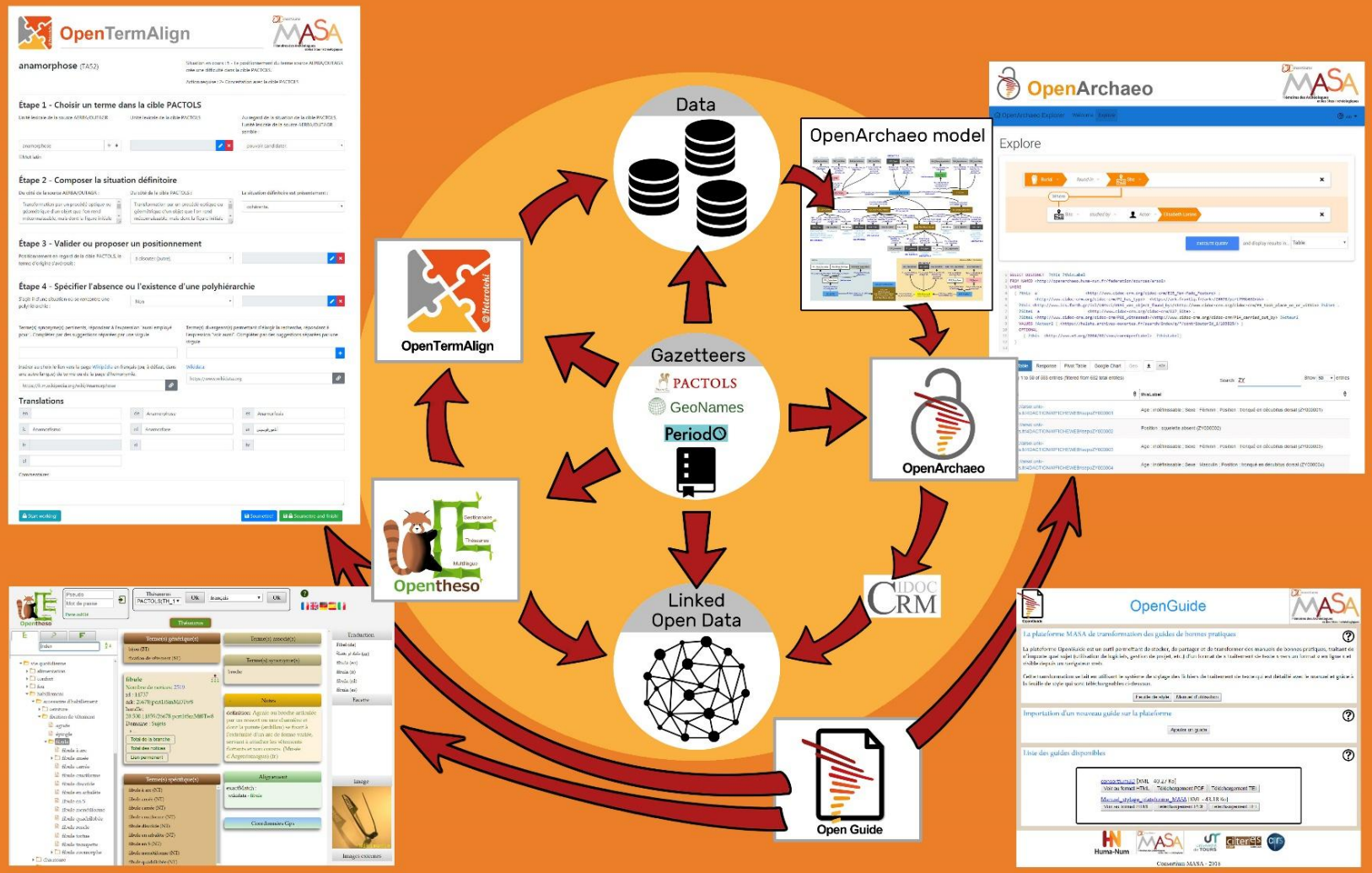


MASA TOOLBOX

- **OpenTermAlign** : Alignment to standardized vocabularies
- **PACTOLS** : archaeological vocabulary, compatibility with DARIAH's BackBone Thesaurus
- **Opentheso** : thesaurus manager
- **OpenArchaeo** : user friendly interface for querying archaeological datasets via CIDOC-CRM
- **LogicistWriter** : help for logicist writing from a diagram
- **OpenGuide** : best practices editing platform

Consortiums MASA
Mémoires des Archéologues et des Sites Archéologiques

MASA ECOSYSTEM



PACTOLS

- Vocabulary controlled by scientific literature and experts in the fields
- Open Access (ODbL v1.0)
- Standardization of the structure (ISO and SKOS)
- Multilingual
- Perennial identification (Ark, Handle)
- Stronger SKOS structuring
- Compatibility with DARIAH BackBone Thesaurus
- Exploiting conceptual categories

- ***activities* / activités**
  - disciplines / *disciplines*
  - *human interactions* / interactions humaines
  - *intentional destructions* / destructions volontaires
  - *functions* / fonctions
  - other activities / autres activités
- ***natural processes* / Processus naturel**
  - *natural disasters* / catastrophes naturelles
  - *Geneses* / Genèses
- ***materials* / matériaux**
- ***material things* / entités matérielles**
  - *mobile objects* / objets mobiles
  - *built environment* / environnement bâti
  - *physical features* / caractéristiques physiques
  - *structural parts of material things* / parties d'entités matérielles
- ***types of epochs* / types d'époques**
- ***conceptual objects* / objets conceptuels**
  - *symbolic objects* / objets symboliques
  - *propostionnal objects* / objets propositionnels
  - *methods* / méthodes
  - *concepts* / concepts
- ***groups and collectivities* / groupes et collectivités**
- ***roles* / rôles**
  - *offices* / rôles officiels
  - *roles of interpersonnal relations* / rôles dans les relations interpersonnelles
- **geopolitical units / unités géopolitiques**

OPENTHESO

Multilingual, multi-hierarchical thesaurus manager

- Open access (CeCILL_C license)
- Standardized
  - ISO 25964-1 2011 and ISO 25964-2:2012
  - SKOS - Unicode
- Interoperable
  - Citability: ID Ark and Handle
  - Import/export : SKOS, Turtle, JsonLD

Diffusion

- TGIR Huma-Num, all consortia and users
- Github

STANDARD VOCABULARY OPENTERMALIGN

Alignement et positionnement entre terminologies hétérogènes

OpenTermAlign

Tous 603/603 | Candidats 209 | Non-candidats 28 | Concertation cible PACTOLS 132 | Concertation source AERBA/OUTAGR 109 | Dépôt 387

Rechercher :

	Identifiants	U. L. source AERBA/OUTAGR	U. L. cible PACTOLS	W	Situation	Action		
1	RSL2us152	abandon	abandon de lieu (35358)		8 - Aucun problème.	4- Dépôt	Alignons !	
2	AD1L1	abside	abside (21832)	W	8 - Aucun problème.	4- Dépôt	Alignons !	
3	AD2L2	accès	entrée (14628)	W	8 - Aucun problème.	4- Dépôt	Alignons !	
4	RSL2mou72	accessoire vestimentaire	accessoire d'habillement (155895) (155895)	W	8 - Aucun problème.	4- Dépôt	Alignons !	
5	O2L48	accessoires animaliers	CANDIDAT		8 - Aucun problème.	4- Dépôt	Alignons !	
6	RSL2mou73	activité agricole	agriculture (13130) (13130)	W	8 - Aucun problème.	4- Dépôt	Alignons !	
7	RSL2mof14	activité artisanale	artisanat (13408) (13408)	W	8 - Aucun problème.	4- Dépôt	Alignons !	
8	RSL2mou74	activité métallurgique	métallurgie (15846) (15846)	W	8 - Aucun problème.	4- Dépôt	Alignons !	

anamorphose (TA52)

Situation en cours : 5 - Le positionnement du terme source AERBA/OUTAGR crée une difficulté dans la cible PACTOLS.

Action requise : 2- Concertation avec la cible PACTOLS

Étape 1 - Choisir un terme dans la cible PACTOLS

Unité lexicale de la source AERBA/OUTAGR

anamorphose fr

Mot latin

Unité lexicale de la cible PACTOLS

Au regard de la situation de la cible PACTOLS, l'unité lexicale de la source AERBA/OUTAGR semble :

pouvoir candidater.

Étape 2 - Composer la situation définitoire

Du côté de la source AERBA/OUTAGR :

Transformation par un procédé optique ou géométrique d'un objet que l'on rend méconnaissable, mais dont la figure initiale

Du côté de la cible PACTOLS :

Transformation par un procédé optique ou géométrique d'un objet que l'on rend méconnaissable, mais dont la figure initiale

La situation définitoire est présentement :

cohérente.

Étape 3 - Valider ou proposer un positionnement

Positionnement en regard de la cible PACTOLS, le terme d'origine s'avérerait :

à discuter (autre).

Étape 4 - Spécifier l'absence ou l'existence d'une polyhiérarchie

S'agit-il d'une situation où se rencontre une polyhiérarchie :

Non

Terme(s) synonyme(s) pertinents, répondant à l'expression "aussi employé pour". Compléter par des suggestions séparées par une virgule

Terme(s) divergeant(s) permettant d'élargir la recherche, répondant à l'expression "voir aussi". Compléter par des suggestions séparées par une virgule

Insérer au choix le lien vers la page Wikipédia en français (ou, à défaut, dans une autre langue) du terme ou de la page d'homonymie.

https://fr.m.wikipedia.org/wiki/Anamorphose

Wikidata:

https://www.wikidata.org

Translations

en		de	Anamorphose	es	Anamorfosis
it	Anamorfismo	nl	Anamorfose	ar	أنامورفوسيس
fr		sl		hr	
el					

Commentaires

Start working! Soumettre! Soumettre and finish!

RE-USE OF DATA

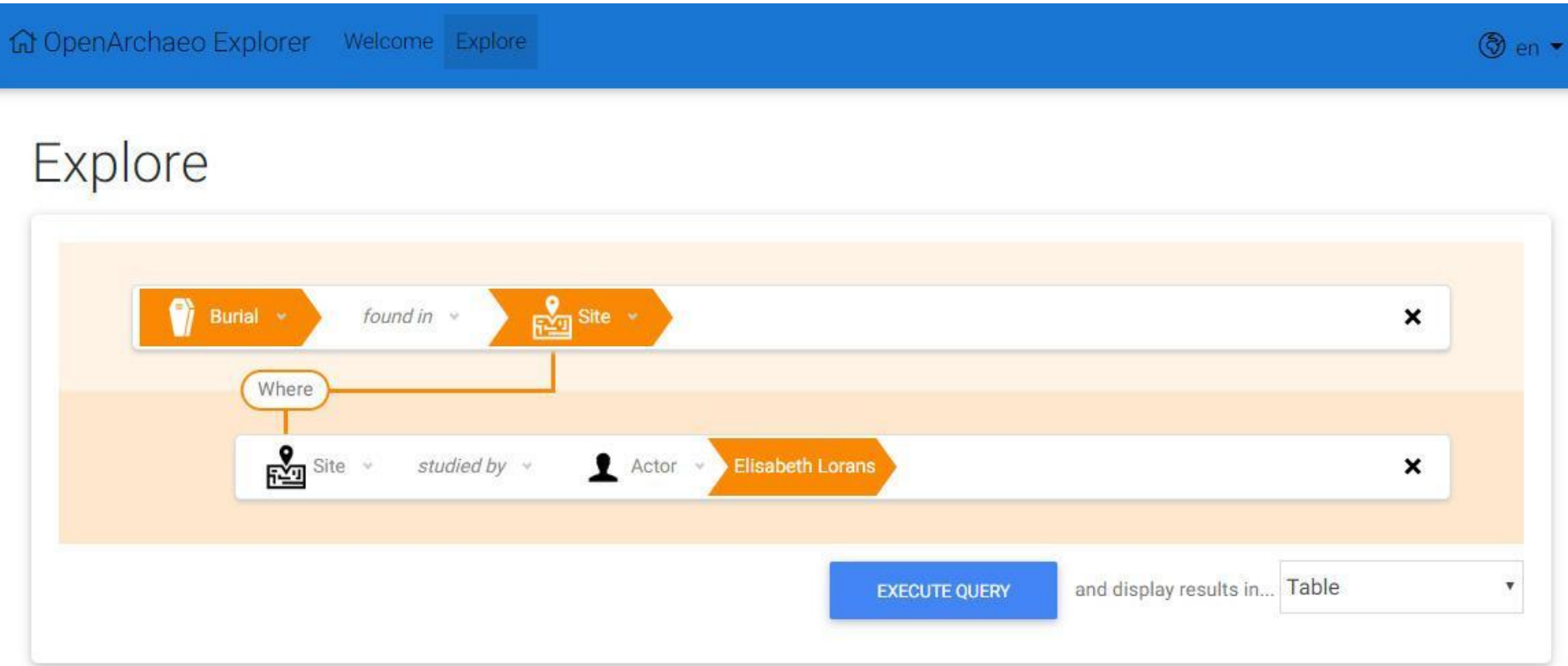
OpenArchaeo

OPENARCHAEO

SPARNA

- Platform for querying archaeological data sets via CIDOC CRM
- Ongoing development by par SPARNA (Thomas Francart)
- Internal or external Triplestore
- SPARQL query generator with CIDOC CRM
- Generic model for matching archaeo data sets to CIDOC CRM

OPENARCHAEO



Consortiums MASA Mémoires des Archéologues et des Sites Archéologiques

SPARNA

OPENARCHAEO

OpenArchaeo

OpenArchaeo Explorer | Welcome | Explore | en

```
SELECT DISTINCT  ?this ?thisLabel
FROM NAMED <http://openarchaeo.huma-num.fr/federation/sources/arsol>
WHERE
  { ?this  a                    <http://www.cidoc-crm.org/cidoc-crm/E25_Man-Made_Feature> ;
           <http://www.cidoc-crm.org/cidoc-crm/P2_has_type>  <https://ark.frantiq.fr/ark:/26678/pcrt795b632nWw> .
    ?this <http://www.ics.forth.gr/isl/CRMsci/O19i_was_object_found_by>/<http://www.cidoc-crm.org/cidoc-crm/P8_took_place_on_or_within> ?Site1 .
    ?Site1  a                     <http://www.cidoc-crm.org/cidoc-crm/E27_Site> .
    ?Site1 <http://www.cidoc-crm.org/cidoc-crm/P8i_witnessed>/<http://www.cidoc-crm.org/cidoc-crm/P14_carried_out_by> ?Acteur2
    VALUES ?Acteur2 { <https://halshs.archives-ouvertes.fr/search/index/q/*/contributorId_i/103825/> }
    OPTIONAL
      { ?this  <http://www.w3.org/2004/02/skos/core#prefLabel>  ?thisLabel}
  }
```

Table | Response | Pivot Table | Google Chart | Geo | </>

Showing 1 to 50 of 94 entries

Search:

Show 50 entries

	this	thisLabel
1	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isepuAJ000001	Age : indéfinissable ; Sexe : Indéterminé ; Position : squelette absent (AJ000001)
2	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isepuAJ000002	Age : adulte (>17 ans) ; Sexe : Masculin ; Position : en décubitus dorsal (AJ000002)
3	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isepuAJ000003	Age : adulte (>17 ans) ; Sexe : Masculin ; Position : en décubitus dorsal (AJ000003)
4	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isepuAJ000004	Age : indéfinissable ; Sexe : Indéterminé ; Position : en décubitus dorsal (AJ000004)

Consortiums

MASA

Mémoires des Archéologues et des Sites Archéologiques

OpenArchaeo Explorer Welcom

```
1  SELECT DISTINCT  ?this ?thisLabel
2  FROM NAMED <http://openarchaeo.huma
3  WHERE
4    { ?this  a                    <h
5             <http://www.cidoc-crm.or
6      ?this <http://www.ics.forth.gr/
7      ?Site1  a                     <
8      ?Site1 <http://www.cidoc-crm.or
9      VALUES ?Acteur2 { <https://hals
10     OPTIONAL
11       { ?this  <http://www.w3.org/2
12   }
13
```

Table | Response | Pivot Table

Showing 1 to 50 of 94 entries

	this
1	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isep
2	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isep
3	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isep
4	http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/isep

ArSol

Archives du Sol - © 1996 - 2018 UMR 7324 CITERES - Laboratoire Archéologie et Territoires

RECHERCHE | CONSULTATION STRATIGRAPHIQUE | ARCHIVE

+Recherche

+Résultats

Tours_site 17 - Marmoutier

Sépulture fouillée par Charles Lelong. Tombeau du cardinal Charles de Bourbon, bâti en moyen appareil, voûté plein cintre, muré à l'est par une maçonnerie grossière comportant des remplois romans. Nombreux graffitis sur les murs latéraux. Corps reposait dans un cercueil de plomb de 4mm d'épaisseur, anthropomorphe (dont le couvercle était décoré d'un écu analogue à celui du cardinal) logé à son tour dans un cercueil de bois peint d'un blason fleurdelysé surmonté d'une croix tréflée et d'une croix de Malte, posé directement sur le sol. Le cercueil en plomb présentait l'inscription : "EN 1610 IUIN 15". Le crâne du squelette a été scié perpendiculairement au front. Importante arthrose polyarticulaire et anomalie congénitale de la jonction crâne-rachis par soudure de l'atlas à l'occipital. La dalle formant le seuil du caveau était probablement d'origine antique (trou de louve) portant l'épitaphe d'un sous diacre, Dodenus, mort en 835. NI=5 Actuel secteur 2

Sépulture 2

Retour à la Fiche du Départ

Anthropologie

Classe d'âge :	adulte (>17 ans)	Position MSO:	
Sexe :	Masculin	Position MSR:	
Age :			

Architecture de la tombe

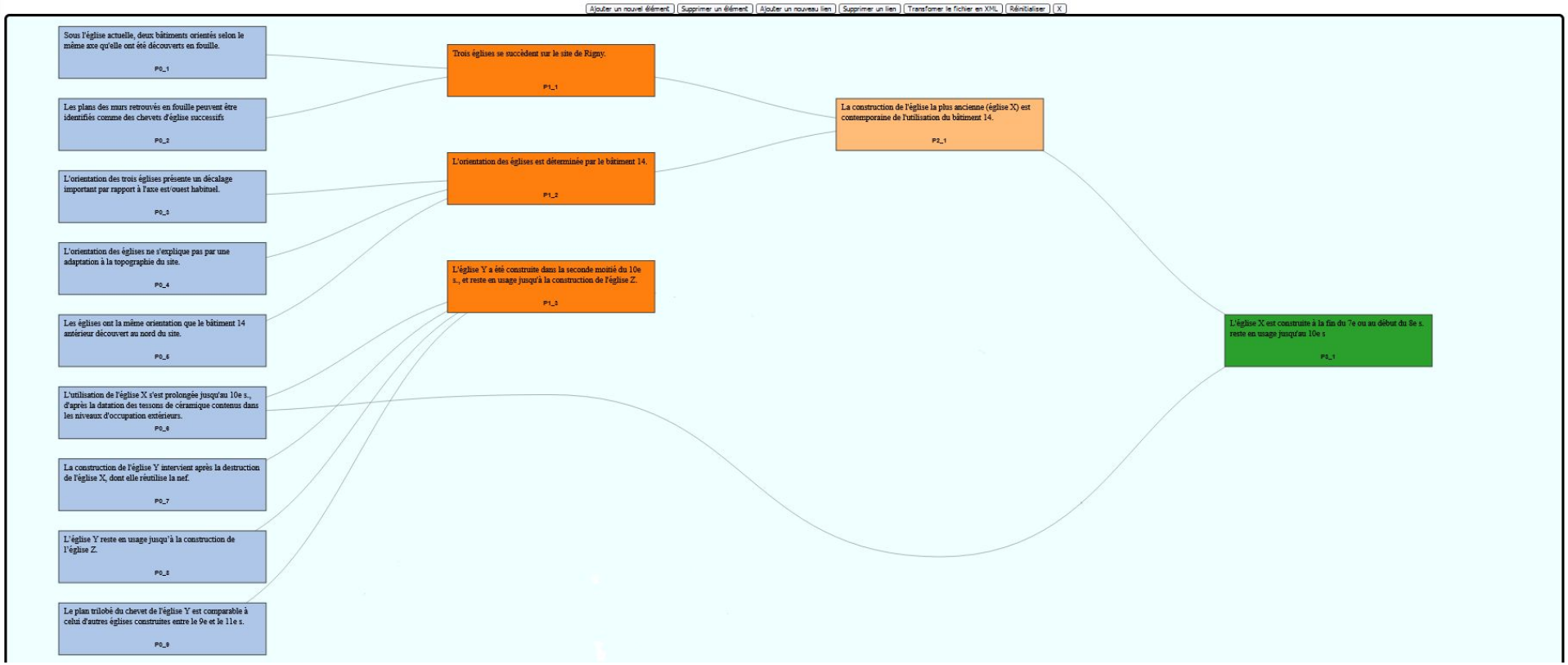
Orientation :	90	Contenant :	tombe construite
Longueur :	2.52	Remploi contenant :	pas de trace de remploi
Largeur :	1.5	Linceul :	

74 PHOTOS 1 2 3 4

LOGICISTWRITER

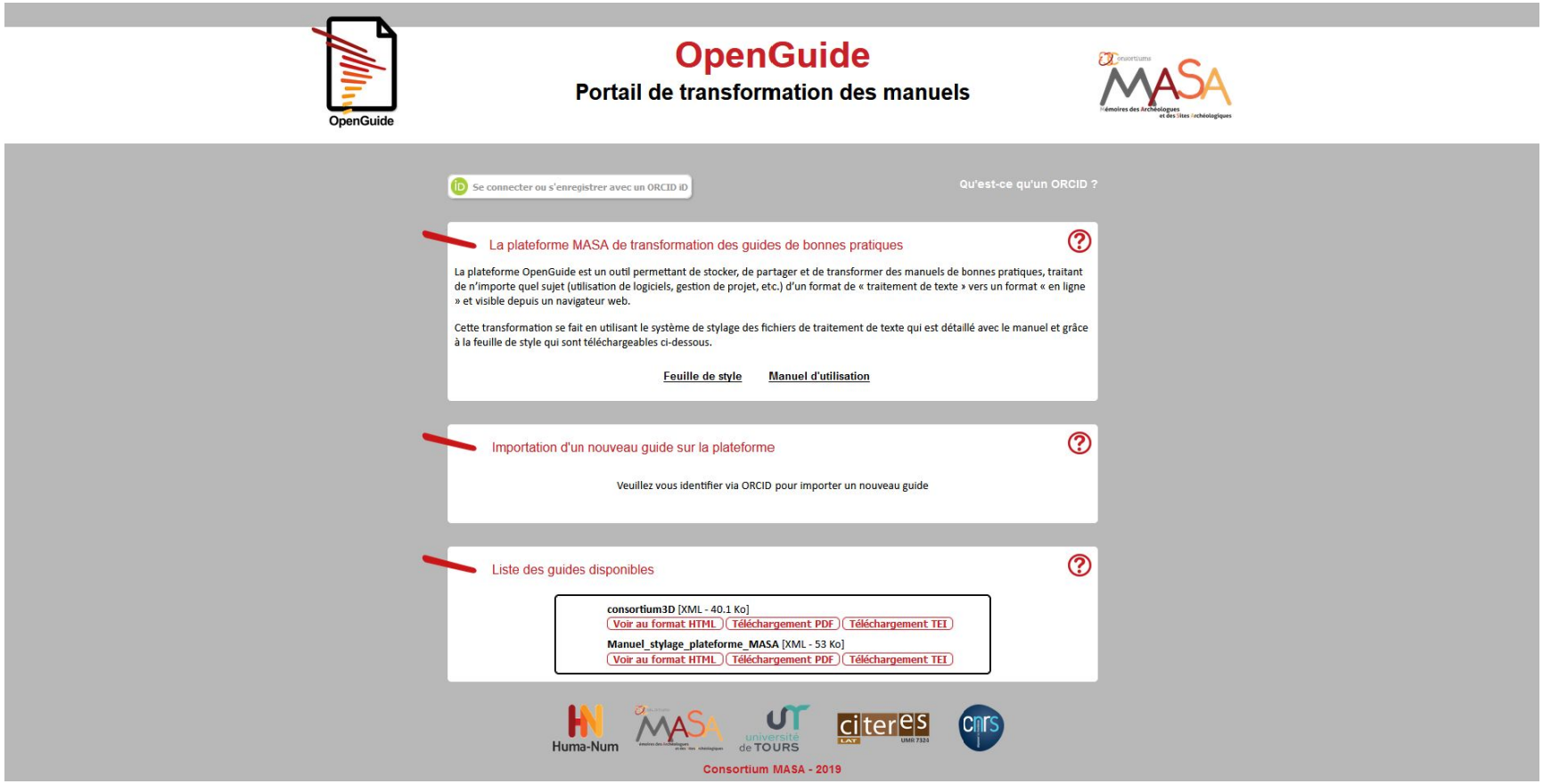
Application of assistance for logicist writing

as defined by Jean-Claude Gardin

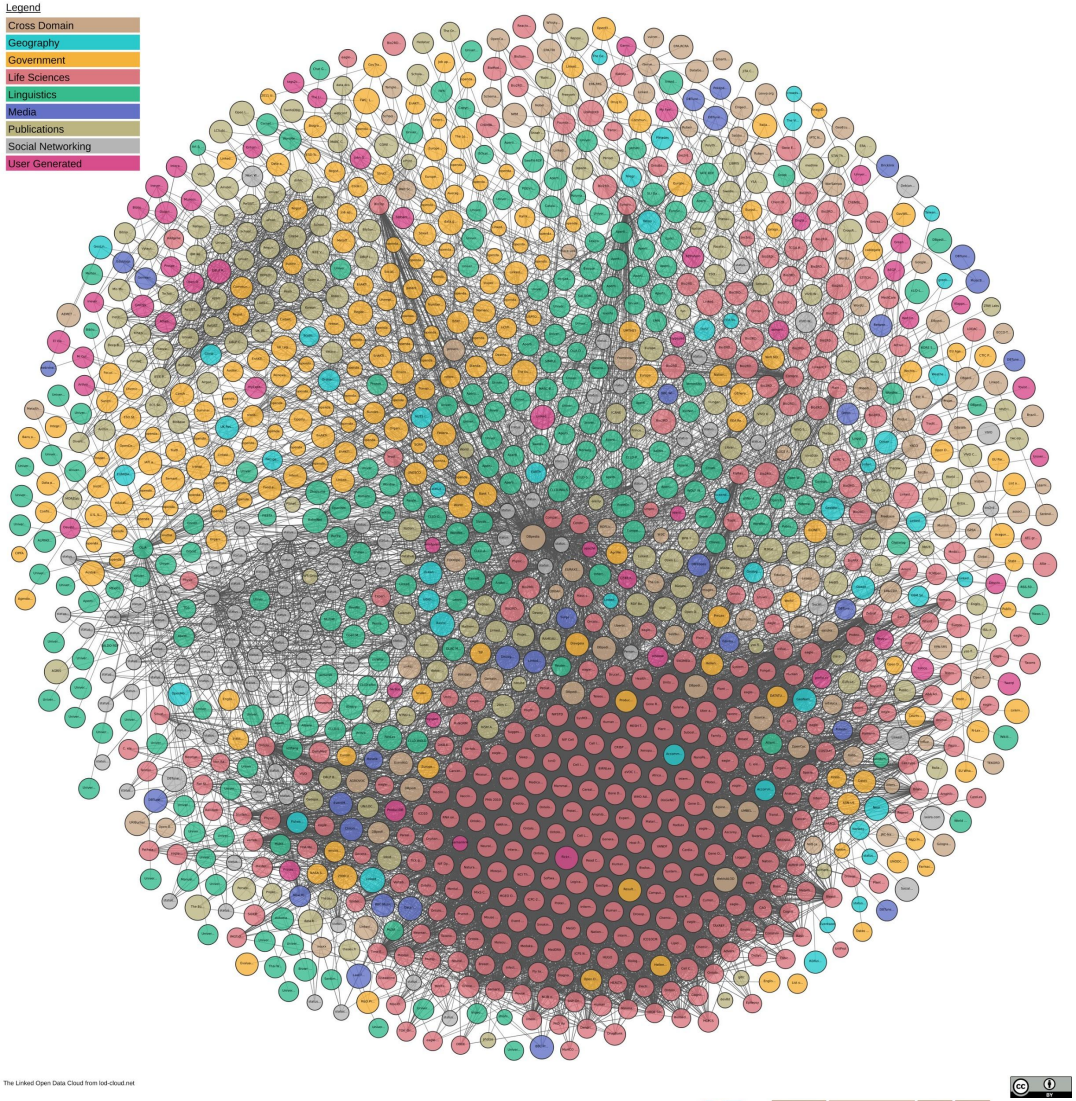


OPENGUIDE

Application to submit and share guidelines



THE LINKED OPEN DATA CLOUD



Wednesday **17th June** to Thursday **18th June 2020**

VIRTUAL MEETING

Organized by

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

Contact

registration@lestudium-ias.fr

The FAIR Heritage conference

Purpose

Share experiences and discuss about:

- State-of-the-art approaches and technologies for cultural/natural heritage data management;
- The challenges that scholars are **currently** facing and how to deal with them;
- Envision the challenges that we will need to face **in the next future**, including, e.g., environmental concerns about data storage.

Ideally,

- To meet fellows;
- To come to know about interesting research projects and initiatives;
- To contribute to the strengthening of the research community.

Some questions

- What are the data management challenges that you face in your research?
- Are the FAIR principles relevant for your work?
- How do you implement the FAIR principles?
- How do you measure the FAIRness of your digital objects?
- Do you make your data available for others to be explored and reused?
- How do you maintain your digital objects beyond specific projects?
- Do you use existing metadata vocabularies? Do you need to adapt them to your own modeling purposes?
- What are the good and bad lessons learned from the use of metadata vocabularies and/or semantic technologies?

Thank you!

Emilio M. Sanfilippo
emilio.sanfilippo@univ-tours.fr
emiliosanfilippo@gmail.com (permanent)

Xavier Rodier
xavier.rodier@univ-tours.fr