

Reference list and tool/learning source registry for excavation data modeling

Working document: *Draft version 0.6 [24/11/2022]*

Table of Contents

1. Excavation data modeling reference list	2
2. Excavation data modeling PPTs	5
3. Data modeling tools and learning facilities registry	5
A. Tools useful in data modelling	5
B. Learning facilities	7

1. Excavation data modeling reference list

- Amico, N. & Felicetti, A. 2021. Ontological Entities for Planning and Describing Cultural Heritage 3D Models Creation. *ArXiv* abs/2106.07277 (2021): n. Pag. <https://arxiv.org/ftp/arxiv/papers/2106/2106.07277.pdf>
- Ashley, M., Tringham, R. and Perlingieri, C. 2011. Last House on the Hill: Digitally remediating data and media for preservation and access. *J. Comput. Cult. Herit.* 4, 4, Article 13 (December 2011), 26 pages. <https://doi.org/10.1145/2050096.2050098>.
- Binding, C., May, K., Tudhope, D. 2008. Semantic Interoperability in Archaeological Datasets: Data Mapping and Extraction Via the CIDOC CRM. In: Christensen-Dalsgaard, B., Castelli, D., Ammitzbøll Jurik, B., Lippincott, J. (eds) *Research and Advanced Technology for Digital Libraries. ECDL 2008. Lecture Notes in Computer Science*, vol 5173. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-87599-4_30.
- Castelli, L., Felicetti, A. & Proietti, F. Heritage Science and Cultural Heritage: standards and tools for establishing cross-domain data interoperability. *Int J Digit Libr* **22**, 279–287 (2021). <https://doi.org/10.1007/s00799-019-00275-2>
- Christaki, E., Doerr, M., Bekiari, C., & Bruseker, G. 2017. Μεθοδολογία μοντελοποίησης ανασκαφικής διαδικασίας με σωζόμενα αρχιτεκτονικά λείψανα. In G. Vavouranakis, M. Katsianis, Y. Papadatos, M. Mouliou, & P. Petridis (Eds.), *Digital Pasts for the present. Proceedings of the 2nd conference on Computer Applications and Quantitative Methods in Archaeology Greek Chapter (CAA-GR)*. Athens, 22-23/12/2016: 26–35. Athens: Department of History and Archaeology. National and Kapodistrian University of Athens (*in Greek*)
- Cripps, P. & May, K. 2010. To OO or not to OO? Revelations from Ontological Modelling of an Archaeological Information System, in: Nicolucci, F. and S. Hermon (eds.), *Beyond the Artifact. Digital Interpretation of the Past. Proceedings of CAA2004*, Prato 13–17 April 2004. Archaeolingua, Budapest, pp. 59-63. https://proceedings.caaconference.org/files/2004/08_Cripps_May_CAA_2004.pdf
- Cripps, P., Greenhalgh, A., Fellows, D., May, K., & Robinson, D. 2004. *Ontological Modelling of the work of the Centre for Archaeology*. CIDOC-CRM Technical Paper. http://old.cidoc-crm.org/docs/Ontological_Modelling_Project_Report_%20Sep2004.pdf.
- D'Andrea, A. Marchese, G. and Zoppi, T. 2006. Ontological modelling for archaeological data. In *Proceedings of the 7th International conference on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST'06)*. Eurographics Association, Goslar, DEU, 211–218.
- Doerr M., Hermon S., Hiebel G., Kritsotaki A., Masur A., May K., Schmidle W., Theodoridou M., Tsiadaki D., 2013. *CRMarchaeo: Modelling Context, Stratigraphic Unit, Excavated Matter*, Paper presented at the 29th CRM-SIG Meeting, Heraklion, Greece, 21 – 25 October, 2013.
- Felicetti, A., Castelli, L., Fedi, M., Aspöck, E., Kreiter, A., Vassallo, V., Hermon, S., Correia, M.J.R.M & Silva, A. S. 2020. *Application Profile for Scientific Data. A tentative list of requirements*. <https://tinyurl.com/yyvlwor8>.

- Felicetti, A., Meghini, C., Richards, J. & Theodoridou, M. 2020. *Towards the AO-Cat Ontology. Version 1.1.* [D4Science > Workspace > VRE Folders > ARIADNEplus_Project > Internal Documents]
- Felicetti, A., Niccolucci, F. & Castelli, L. 2020. *CRMhs - A Semantic Model for Heritage Science. Version 0.8* [D4Science > Workspace > VRE Folders > ARIADNEplus_Project > Internal Documents]
- Felicetti, A., Scarselli, T., Mancinelli, M.L., & Niccolucci, F. 2013. Mapping ICCD archaeological data to CIDOC-CRM: the RA Schema. *CRMEX@TPDL*. <https://ceur-ws.org/Vol-1117/paper2.pdf>.
- Giagkoudi, E., Tsiafaki, D. & Papatheodorou, C. 2018. Describing and revealing the semantics of excavation notebooks. *CIDOC 2018 Heraklion, Crete, Greece*.
http://www.cidoc2018.com/sites/default/files/CIDOC2018_paper_71_0.pdf.
- Hiebel, G., Doerr, M., Hanke, K. & Masur, A. 2014. How to Put Archaeological Geometric Data into Context? Representing Mining History Research with CIDOC CRM and Extensions. *International Journal of Heritage in the Digital Era* 3 (3): 557–77. <https://doi.org/10.1260/2047-4970.3.3.557>.
- Hiebel, G., Aspöck, E. and Kopetzky, K. 2021. Ontological Modeling for Excavation Documentation and Virtual Reconstruction of an Ancient Egyptian Site. *J. Comput. Cult. Herit.* 14, 3, Article 32 (July 2021), 14 pages. <https://doi.org/10.1145/3439735>.
- Jones, A. 2001. *Archaeological Theory and Scientific Practice* (Topics in Contemporary Archaeology). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511606069.
- Katsianis, M. 2009. *Ανασκαφική Μεθοδολογία και Σχεδιασμός Πληροφοριακού Συστήματος για τη Διαχείριση Αρχαιολογικών Τεκμηρίων* (Excavation Methodology and Information System Design for Archaeological Data Management), Aristotle University Thessaloniki. <http://ikee.lib.auth.gr/record/112182/files/KatsianisPhD.pdf> (in Greek)
- LeBoeuf, P., Doerr, M., Ore, C.E. & Stead, S. 2015. *Definition of the CIDOC Conceptual Reference Model. Version 6.2*. http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_6.2.pdf.
- Lucas, G. 2002. *Critical approaches to fieldwork: contemporary and historical archaeological practice*. Routledge.
- Lukas, D. Engel, C & Mazzucato, C. 2018. Towards a Living Archive: Making Multi Layered Research Data and Knowledge Generation Transparent, *Journal of Field Archaeology*, 43:sup1, S19-S30, <https://doi.org/10.1080/00934690.2018.1516110>.
- Marlet, O., Zadora-Rio, E., Buard, P.-Y., Markhoff, B. & Rodier, X. 2019. The Archaeological Excavation Report of Rigny: An Example of an Interoperable Logician Publication. *Heritage*, 2, 761-773. <https://doi.org/10.3390/heritage2010049>.
- Marlet, O., Francart, T., Markhoff, B. & Rodier, X. 2019. OpenArchaeo for Usable Semantic Interoperability. *ODOCH 2019 @CAiSE 2019*. Jun 2019, Rome, Italy. <https://hal.archives-ouvertes.fr/hal-02389929>.
- May, K. 2020. The Matrix: Connecting Time and Space in archaeological stratigraphic records and archives, *Internet Archaeology* 55. <https://doi.org/10.11141/ia.55.8>.
- Nussbaumer, P. Haslhofer, B. & Klas, W. 2010. Towards Model Implementation Guidelines for the CIDOC Conceptual Reference Model. Technical Report. Multimedia Information Systems. University of Vienna. http://eprints.cs.univie.ac.at/58/1/nussbaumer10_cidoc_crm.pdf.
- Paveprime and collaborators 2019. *CRMinf: the Argumentation Model. An extension of CIDOC-CRM to support argumentation*. Version 0.10.1. <https://cidoc-crm.org/crminf/sites/default/files/CRMinf%20ver%2010.1.pdf>

- Ronzino, P., Niccolucci, F., Felicetti, A. *et al.* CRMba a CRM extension for the documentation of standing buildings. *Int J Digit Libr* **17**, 71–78 (2016). <https://doi.org/10.1007/s00799-015-0160-4>.
- Sitara, M., Poulos, M. & Papavasopoulos, S. 2017. Documenting a fruitstand from Karabournaki with CIDOC/CRM. *Qualitative and Quantitative Methods in Libraries*, 4 (2): 363-371. Available at: <http://qqml-journal.net/index.php/qqml/article/view/253>.
- Tudhope, D., Binding, C., May, K., & Heritage, E. (2008). Semantic interoperability issues from a case study in archaeology. In *Semantic Interoperability in the European Digital Library, Proceedings of the First International Workshop SIEDL* (pp. 88-99). https://pure.southwales.ac.uk/ws/portalfiles/portal/2198731/SIEDL08_Tudhope_v3.pdf.
- Wright, H. 2011. *Seeing Triple: Archaeology, Field Drawing and the Semantic Web*. PhD thesis, University of York. <https://etheses.whiterose.ac.uk/2194/1/WrightThesis.pdf>.
- Vlachidis, A., & Tudhope, D. (2012). A pilot investigation of information extraction in the semantic annotation of archaeological reports. *International Journal of Metadata, Semantics and Ontologies*, 7(3), 222-235. https://discovery.ucl.ac.uk/id/eprint/1556223/1/Vlachidis_ijmso-vlachidis-pilot-investigation-of-IE.pdf.

2. Excavation data modeling PPTs

- Aspöck, E, Hiebel, G. & Ďurčo, M. 2016. [Developing an archaeological data archive at the Austrian Academy of Sciences: the 'A puzzle in 4D' project](#). 37th CIDOC CRM and 30th FRBR CRM
- Bekiari, C. 2014. [CRMarchaeo Modelling Context, Stratigraphic Unit, Excavated Matter](#). Tutorial.
- Felicetti, A. 2016. [CRMarchaeo Updates](#). 36th CIDOC CRM and 29th FRBR CRM.
- Felicetti, A. 2020. [CRMarchaeo-editorial issues](#). 46th CIDOC CRM and 39th FRBR CRM
- Fiorani, D. & Acierno, M. 2017. [Architectural Conservation Process Model \(CPM\)](#). 39th CIDOC CRM and 32nd FRBR CRM
- Hiebel, G. 2016. [Information integration in prehistoric mining archaeology](#). 37th CIDOC CRM and 30th FRBR CRM.
- Stead, S. 2019. [CIDOC CRM and CRMarchaeo: A vision of use for the 'future'](#)
- Ronzino, P. & Felicetti, A. 2015. [CRMBA and CRMarchaeo models harmonization](#), 34th CIDOC CRM and 27th FRBR CRM.

3. Data modeling tools and learning facilities registry

A. Tools useful in data modelling

- GraphDB - <https://graphdb.ontotext.com/>

GraphDB is an enterprise ready Semantic Graph Database, compliant with W3C Standards. Semantic graph databases or RDF triplestores, provide the core infrastructure for data modelling. Graph DB can be accessed via the ARIADNEplus VRE providing access to the ARIADNEplus Knowledge Base, including all the partner data as a Linked Open Data set and modelled according to the ARIADNE ontology. Using GraphDB, researchers can explore the knowledge base with the available web GUI or programmatically with SPARQL queries. Documentation (KB-guide-v1.0) is provided as a python notebook, a PDF file or in Markdown format at <https://data.d4science.net/EvVX>. The python notebook can be run on the JupyterHub tool also available in the ARIADNEplus Lab VRE at https://ariadne.d4science.org/group/ariadneplus_lab/jupyterhub. The guide also contains several examples of queries that can be run on the ARIADNE Knowledge base. Users should copy the notebook to a Jupyter instance in order to play with it and to create their own notebooks for their analysis of the ARIADNEplus Knowledge Base.

- OntoRefine - <https://disc-semantic.uibk.ac.at/ontorefine>

OntoRefine is a data transformation tool, based on OpenRefine, that is integrated into GraphDB by OntoText. The environment adds support for direct RDF conversion through a SPARQL endpoint.

- Sparnatural - <https://sparnatural.eu/>

A Javascript based system allowing the user to query an RDF graph with a graphic interface and without any SPARQL to write.

- Protégé-Ontop - <https://protege.stanford.edu/> & <https://ontop-vkg.org/>

Protégé is an open source ontology editor and knowledge management system. Ontop-Protégé is a plugin for designing and testing a Virtual Knowledge Graph (VKG) specification. (See also <https://halshs.archives-ouvertes.fr/halshs-03561376v2/file/MappingOntop.pdf>.

- Karma - <https://usc-isi-i2.github.io/karma/>

Karma is an information integration tool that enables users to quickly and easily integrate data from a variety of data sources including databases, spreadsheets, delimited text files, XML, JSON, KML and Web APIs. Users integrate information by modeling it according to an ontology of their choice using a graphical user interface that automates much of the process. Karma learns to recognize the mapping of data to ontology classes and then uses the ontology to propose a model that ties together these classes.

- SHACL - <https://shacl-play.sparna.fr/play/>

A tool for verifying the conformity of the knowledge graphs with the generic models used.

- X3ML Toolkit (FORTH)

Consists of a set of software components that assist the data provisioning process for information integration. The key components of the toolkit are:

a) X3ML Mapping Definition Language

(<https://github.com/isl/x3ml/blob/master/docs/x3ml-language.md>) is an XML based language that describes schema mappings,

b) 3M Mapping Memory Manager (<https://demos.isl.ics.forth.gr/3m>) is a web application suite containing several software sub-components and exploits several external services.

c) X3ML Engine (<https://github.com/isl/x3ml>) realises the transformation of the source records to the target format

d) RDF Visualiser (<https://demos.isl.ics.forth.gr/RDFV-Demo/>) is a generic browsing mechanism that gives the user a flexible, highly configurable, detailed overview of a dataset/database encoded in RDF.

- CRITERIA - <https://github.com/chin-rcip/CRITERIA> & Mermaid - <https://mermaid-js.github.io/mermaid/#/>

A tool for the automatic transform of RDF data into visualised graphs

- diagrams.net Libraries - https://github.com/chin-rcip/diagrams.net_libraries

A library created by the Canadian Heritage Information Network (CHIN) to facilitate the visualisation of data models based on CIDOC CRM and its extensions when using the diagrams.net tool (<https://app.diagrams.net/>).

- AIRtable - <https://www.airtable.com/>

A spreadsheet-database hybrid cloud collaboration service, with the features of a database applied onto a spreadsheet.

- Vocabulary Matching Tool (VMT) - <https://heritagedata.org/vocabularyMatchingTool/>

The Vocabulary Matching Tool developed as part of the ARIADNE H2020 project builds on a tool originally developed during the ARIADNE FP7 Project by the University of Wales. It can be used for the creation of mappings from locally used terms/concepts to the Getty Art & Architecture Thesaurus (AAT). The aim is to identify subject mappings from source terms/concepts to AAT concepts that are likely to be useful to assist subsequent data searching and browsing. The creation of mappings to a common spine vocabulary enables improved opportunities for multilingual subject access and cross search, by aggregating mappings from multiple data partners. The set of mappings created may be exported to JSON or delimited text (CSV) format for use in other applications. The tool is accessible also via the ARIADNEplus VRE.

- PeriodO - <https://perio.do/>

The PeriodO gazetteer service allows the definition of temporal intervals as web resources, using a label and two absolute dates giving the earliest start and the latest stop of each interval. The service also allows clustering period resources in collections, thus facilitating their exploration.

- DARIAH Back Bone Thesaurus - <https://www.backbonethesaurus.eu/>

A model for sustainable interoperable thesauri maintenance, developed in the European Union Digital Research Infrastructure for the Arts and Humanities (DARIAH). Its development aims at designing and establishing an overarching thesaurus for the humanities (BackBone Thesaurus, BBT for short). Relevant work focuses on identifying the top-level-concepts (facets and hierarchies) that will become a common basis for thesaurus building in an effort to meet the demands for objectivity and interdisciplinarity.

B. Learning facilities

- CIDOC CRM Game Digital - <https://www.cidoc-crm-game.org/node/39> (downloadable version), <https://ontomatchgame.huma-num.fr/> (online version)

CIDOC CRM the Game is an open source initiative that aims to promote wider adoption of the CIDOC CRM standard through facilitating the learning process for beginning users/learners. The Game provides a fun way to visualise and play with the ontology in relation to real world documentation scenarios.

- CIDOC CRM tutorials. <https://cidoc-crm.org/tutorialPageRes>

CIDOC CRM the Game is an open source initiative that aims to promote wider adoption of the CIDOC CRM standard through facilitating the learning process for beginning users/learners. The Game provides a fun way to visualise and play with the ontology in relation to real world documentation scenarios.