

Arqueología y *Téchne*

Métodos formales, nuevos enfoques

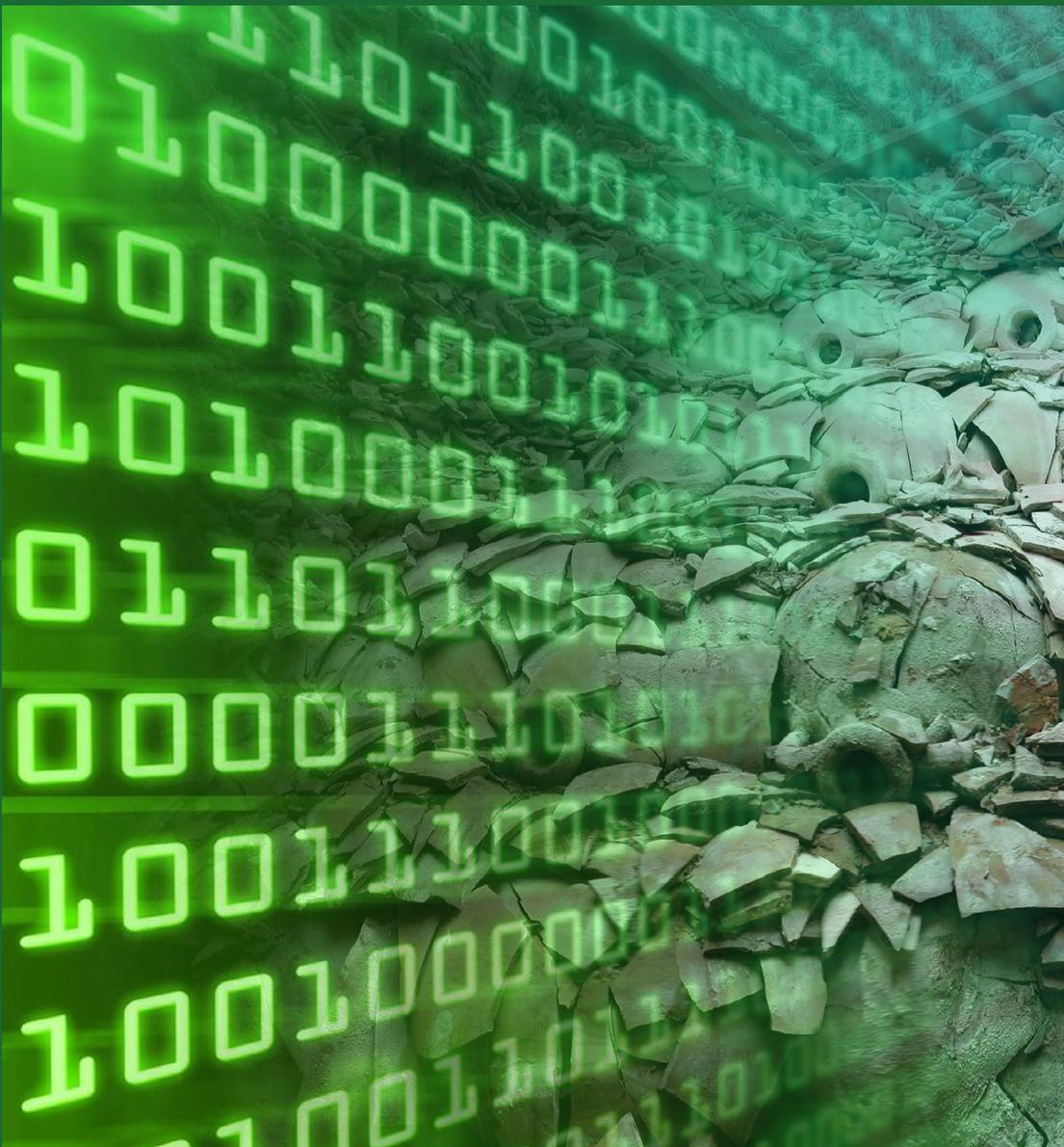
Archaeology and *Techne*

Formal methods, new approaches

Editado por

José Remesal Rodríguez

Jordi Pérez González



Access Archaeology



About Access Archaeology

Access Archaeology offers a different publishing model for specialist academic material that might traditionally prove commercially unviable, perhaps due to its sheer extent or volume of colour content, or simply due to its relatively niche field of interest. This could apply, for example, to a PhD dissertation or a catalogue of archaeological data.

All *Access Archaeology* publications are available as a free-to-download pdf eBook and in print format. The free pdf download model supports dissemination in areas of the world where budgets are more severely limited, and also allows individual academics from all over the world the opportunity to access the material privately, rather than relying solely on their university or public library. Print copies, nevertheless, remain available to individuals and institutions who need or prefer them.

The material is refereed and/or peer reviewed. Copy-editing takes place prior to submission of the work for publication and is the responsibility of the author. Academics who are able to supply print-ready material are not charged any fee to publish (including making the material available as a free-to-download pdf). In some instances the material is type-set in-house and in these cases a small charge is passed on for layout work.

Our principal effort goes into promoting the material, both the free-to-download pdf and print edition, where *Access Archaeology* books get the same level of attention as all of our publications which are marketed through e-alerts, print catalogues, displays at academic conferences, and are supported by professional distribution worldwide.

The free pdf download allows for greater dissemination of academic work than traditional print models could ever hope to support. It is common for a free-to-download pdf to be downloaded hundreds or sometimes thousands of times when it first appears on our website. Print sales of such specialist material would take years to match this figure, if indeed they ever would.

This model may well evolve over time, but its ambition will always remain to publish archaeological material that would prove commercially unviable in traditional publishing models, without passing the expense on to the academic (author or reader).



Arqueología y *Téchne* **Métodos formales, nuevos enfoques**

Archaeology and *Techne* **Formal methods, new approaches**

Editado por
José Remesal Rodríguez
Jordi Pérez González

Access Archaeology





ARCHAEOPRESS PUBLISHING LTD

Summertown Pavilion
18-24 Middle Way
Summertown
Oxford OX2 7LG
www.archaeopress.com

ISBN 978-1-80327-181-1
ISBN 978-1-80327-182-8 (e-Pdf)

© the individual authors and Archaeopress 2022

Cover image: Muro de ánforas de la campaña arqueológica del año 2002 en el Monte Testaccio (Roma). Directores de la excavación: J.M.^a Blázquez Martínez y José Remesal Rodríguez. Fotografía y edición gráfica de la cubierta: Juan Moros Díaz (CEIPAC). *Arqueología y Téchne. Métodos formales, nuevos enfoques / Archaeology and Techne. Formal methods, new approaches* has been funded within the research projects “Economic and Political Networks (EPNet Project), ERC-2013-ADG-340828 (<http://www.roman-ep.net>); Relaciones Interprovinciales en el Imperio Romano. Producción y comercio de alimentos hispanos (*Provinciae Baetica et Tarraconensis*) (HAR2017-85635-P; AEI/FEDER, UE); Centro para el Estudio de la Interdependencia Provincial en la Antigüedad Clásica (CEIPAC) (2017 SGR 512); Dinàmiques socioeconòmiques del món rural romà: formes d’habitat i cultura material al litoral central català (CLT009/18/00045). Jordi Pérez González funded by Juan de la Cierva-Formación-2019 (Agencia Estatal de Investigación (Ministerio de Ciencia e Innovación) REF: FJC2019-040688-I.



UNIVERSITAT DE
BARCELONA

Universitat
de Girona



All rights reserved. No part of this book may be reproduced, stored in retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying or otherwise, without the prior written permission of the copyright owners.

This book is available direct from Archaeopress or from our website www.archaeopress.com

Contents

Contributors	ii
Arqueología y Técnica José Remesal Rodríguez	1
Interacciones que dejan huella. Hacia una ciencia de redes de los objetos que quedaron Luce Prignano, Ignacio Morer Zapata and Albert Díaz-Guilera	5
RomanOpenData: A semantic based Data Visualization & Exploratory Interface Xavi Giménez Baqués, Alessandro Mosca, Bernardo Rondelli and Guillem Rull Fort	18
From Counting Pots to Reconstructing Economy: Computational Tools Developed in the EPNet Project Iza Romanowska, Simon Carrignon, María Coto-Sarmiento, Jean-Marc Montanier and Xavi Rubio-Campillo	27
Similarity Analysis in Epigraphy. Syntactic Clustering of <i>Tituli Picti</i> on the PO8 Amphoras Daniel J. Martín Arroyo-Sánchez and Iza Romanowska	34
La economía de <i>Mauretania Tingitana</i> en el Alto Imperio (2009-2019). Colaboraciones internacionales y aplicación de nuevas técnicas Lluís Pons Pujol	55
Analysis tools for the study of the <i>amphorae</i> productions from the northeast of <i>Hispania Citerior Tarraconensis</i>. A first approach from EPNet Project Antoni Martín i Oliveras, Carlos Palacín Copado and Jordi Pérez González	67
Esquemas organizativos de la producción anfórica olearia bética (Dressel 20, ca. 30-270 d.C.) ... Juan Moros Díaz	98
La importancia de una provincia bisagra en el concierto del Imperio romano. El caso de <i>Raetia</i> Juan Manuel Bermúdez Lorenzo	113
Olive Oil Beyond the Wall: Stamps on Baetican Dressel 20 found on the Forth-Clyde Isthmus Jordi Pérez González	121
Approaches to the Roman Food Economy: GIS Agricultural Modelling in Baetica and Amphorae Epigraphy from <i>Pompeii</i> Daniel J. Martín Arroyo-Sánchez	146
Bibliography of the EPNet Project Leticia Sierra Díaz and Jordi Pérez González	173

Roman Open Data: A semantic based Data Visualization & Exploratory Interface

Xavier Gimenez Baqués

Alessandro Mosca

Bernardo Rondelli

Guillem Rull Fort

1. From observation to explanation

Ancient History is based on abductive reasoning: it starts with a set of observations then seeks to find the most likely explanation for the observations. This process, unlike deductive reasoning, yields plausible conclusions but does not always verify it. Abductive conclusions can generate uncertainty or doubt, which is expressed in retreat terms such as “best available” or “most likely” explanation. In this framework, Ancient History strongly depends on the capacity of exploring observations, building the explanation and justifying it as the “most likely”.

The EPNNet project (Remesal *et al.* 2014; Remesal and Rondelli 2017) was born exactly to overtake this boundary. The main goal of EPNNet was the use of formal tools for falsifying existing hypotheses concerning the roman economy. Over the last couple of centuries, several historians have developed a variety of hypotheses to explain the organization of the Roman Empire trade system. These hypotheses continue to be based on “the most likely” approach and are difficult to falsify.

Roman amphorae can be considered one of the best archaeological markers of Roman Empire economic production and transactions. They are a unique indicator of social relationships and chronological dynamics, because of their distribution throughout the Empire, the precise information contained in their epigraphy (similar to modern trademarks and labels) and their diachronic persistence.

Here we describe our contribution, within the EPNNet Project, in providing historians with computational tools to compare, aggregate, measure, geo-localise, and search data about Latin inscriptions on amphoras for food transportation.

2. open data, semantics and data visualization

At the end of the eighties, the CEIPAC research group began compiling a database of amphorae epigraphy. The database is based on a systematic analytical approach to the epigraphy of amphorae, initiated by H. Dressel, neglected for a long time, and then updated and improved first by E. Rodríguez Almeida and later by the CEIPAC group. This database currently contains more than 54,000 entries and represents an extraordinary source for providing observations about roman empire and supporting the generation of historical hypotheses.

A specific research activity within the EPNet project has been dedicated to the development of an open and semantic based new infrastructure of the CEIPAC database, integrating a series of other existing datasets (ADS – Roman Amphorae, Epigraphic database of Heidelberg, Pleiades graph of ancient places) using the ontology-based data access and integration approach (Calvanese *et al.* 2015; 2016; Mosca *et al.* 2015). A Data Visualization & Exploratory Interface has been then developed to explore in a more powerful and intuitive way all the data integrated.

We developed an ontology that models the domain of Roman amphora epigraphy. This ontology is defined according to the state-of-the-art formal ontological models and standards for representing cultural heritage objects; in particular, it is a specialization/extension of the CIDOC CRM.¹ Other standards that have been taken into account and served as inspiration are: FaBiO² for the bibliographic references, the Europeana Data Model,³ and the EAGLE Metadata Model and domain-centered vocabularies.⁴ This knowledge representation effort addresses two main problems. Firstly, the ontology provides a formally defined, unambiguous, framework for exporting the data in a way that can be manipulated by computer-based simulations and complex-network analysis tools (*see the other papers in this volume*). And secondly, it allows the publication of the data on the web in an interoperable, standard-compliant way, following the Linked Data approach, which makes the project's data integrable with other data sources.

All the documentation for the ontology developed can be found in <http://romanopendata.eu/sparql/doc>. A technical and exhaustive description of the approach can be found in (Calvanese *et al.* 2015; 2016).

2.1 Data integration & management

We integrated the CEIPAC database with three complementary datasets:

- the ADS (Archeology Data Service) Roman Amphorae⁵
- the Epigraphic Database Heidelberg⁶
- the Pleiades gazetteer of ancient places⁷

The integrated system follows the Ontology-Based Data Access (OBDA) approach, in which the integrated data sources are mapped into a common domain ontology, so users can pose their queries on the system in terms of the homogeneous vocabulary defined by the ontology. In this way, users can access the data from all the integrated datasets without having to know the individual structure of each of them. To the users, all the data is structured according to the domain ontology of the integrated system. Data access is done via the query language SPARQL, a standard defined by the W3C, and one of the core components of the Linked Data framework. A SPARQL endpoint for the project is available at <http://romanopendata.eu/sparql>.

To facilitate the addition of new data to the system, a web-based data entry interface has been designed and is available at <http://romanopendata.eu/dataentry>. Authorized users can login into this web

¹ <http://www.cidoc-crm.org>.

² <http://vocab.ox.ac.uk/fabio>.

³ <http://pro.europeana.eu/edm-documentation>.

⁴ <http://www.eagle-network.eu/about/documents-deliverables>.

⁵ http://archaeologydataservice.ac.uk/archives/view/amphora_ahrb_2005.

⁶ <https://edh-www.adw.uni-heidelberg.de>.

⁷ <https://pleiades.stoa.org>.

application, fill the corresponding forms with the new epigraphic data, and submit it for review. An administrator user that is also an expert in the domain will later verify the quality of the data and, if it qualifies, accept it for publication. New records are added as part of the CEIPAC dataset for Roman amphora epigraphy.

2.2 Data visualization

Data visualisations and graphics are fundamental to studying complex subject matter. However, beyond acknowledging this value, scientists rarely consider how visualisations can enable discovery, create engaging and robust reporting, or support online resources. Producing accessible and unbiased visualisations from complicated, uncertain data requires expertise and knowledge from science, computing, and design (McInerny *et al.* 2014), however it represents a powerful solution for supporting scientific research both in terms of data explorations and argumentation. In the last years the use of data visualizations to support scientific research has increased, as well as the literature that emphasizes the research about how people perceive, read (and misread) and interprets different kinds of visualizations and which types of plots are most effective and easiest to decipher. Nowadays, there is a good amount of literature depicting best practices about how to create effective visualization ('Visualization Analysis & Design', Tamara Munzner 2014). On this domain, and as a result of the growth of influence that Data Visualization is acquiring, it is noteworthy to mention some examples, such as the publication of *The Atlas of Science: Visualizing what we know* (Porter and Börner 2011), which offers a compendium of the evolution of science through the use of visual maps.

In the context of the EPNNet project, in order to facilitate the exploration and navigation of the integrated data for users that are not proficient with the SPARQL query language, we have developed an interactive, online data visualization tool, which is available at <http://romanopendata.eu>.

This tool innovates by coupling the semantic web approach with data visualization technologies, for supporting data-driven explorations and pattern detections. This means that all data shown in the tool is obtained in real time by querying the integrated system via the provided SPARQL endpoint. The generated queries depend on the user's selections within the tool: users are indirectly making their own queries by clicking on items and selecting filters, without having to be aware of the SPARQL query language at all.

In order to offer the maximum flexibility in terms of search capabilities, the user interface offers a set of filtering options that covers most of the relevant attributes that characterizes the data. The filters are divided into two categories (epigraphic and bibliographic attributes) that can be combined when performing an advanced search, as shown in the figure below:

New Search «

Search for inscriptions... (like pnn) **Advanced search**

match exact text ⓘ

Epigraphic search ⓘ

Inscription Types	Places	Finding Places
Countries	Amphora Types	Place Functions
Inscription Direction	Inscription Position	Reliefs

Select inscription types to look for:

Titulus pictus

Stamp

Sketch

Bibliographic search ⓘ

Search by author:

Bibliography title:

Search by publication year:

Search

New Search «

Search for inscriptions... (like pnn) **Advanced search**

match exact text ⓘ

Epigraphic search ⓘ

Bibliographic search ⓘ

Search by author:

Bibliography title:

Search by publication year:

Search

The interface relies on some well-know foundational visualization design patterns, by offering an interconnected set of views (linking and brushing paradigm), each of which facets the data in different dimensions, and at the same time, all reacting to possible filter or brushing actions done by users from any view.

The tool also promotes ‘progressive disclosure’, by showing first aggregated views of the data and then providing a clear path to unveiling details, accessible on demand:



Example of the user interface showing aggregated data from 3 different query results, and then providing affordances to drill down to more granular data.

The tool is equipped also with more advanced filtering systems, like Elastic lists, which offers mechanisms for browsing multi-faceted data, by enabling the selection of different attributes on the data and generating combinations of attributes that shape the data according the selection.

AmphoricTypeTitle	AreaTitle	FindingPlaceTitle	SimplifiedTranscription
Dressel 20	Abernethy	13.8.1960.	PNN
	Agde	89 N 01 (090-110)	
	Alcolea del Rio	89 N 01-02 (030-040)	
	Amiens	89 N 02 (120-140)	
	Angers	89 N/S (01-02, 050-060)	
	Ardèche (Alba)	89 S 01 (120)	
	August	89 S 01-02 (030-050)	
	Banasa	89 S 09 (100-120)	
	Bardon Mill	89 S 09 (120-140)	
	Biesheim	89 S 09 (160-180)	
	Bonn	89 S 10 (030-040)	
Dressel 23	Caerleon	89 S 10 (065)	

Example of an elastic list showing the distribution of occurrences by 4 different attributes and enabling combinations of selection in order to filter the data by multiple criteria.

Access to individual datum is always available from the different existing views (geographically, timeline, and tabular representations), as shown below:

Browsing 11 images so far.
Scroll down to fetch more images, if any

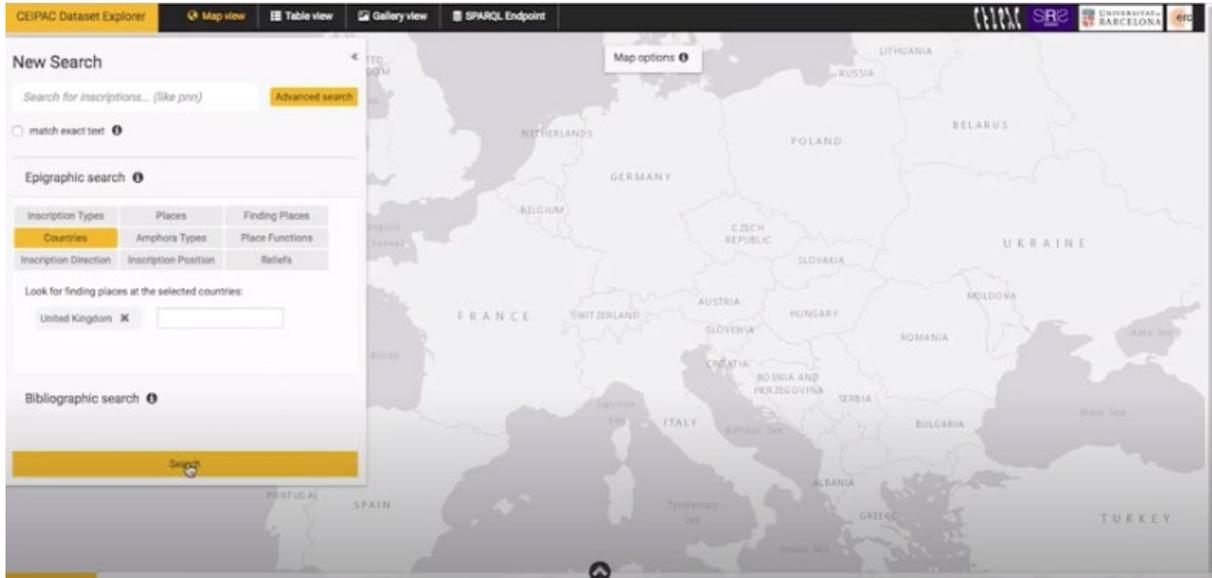
The grid displays 11 individual datum cards. Each card contains an image of an artifact, a 'SimplifiedTranscription' (e.g., PNN, P·N·N·), and a 'FullTranscription' (e.g., |PNN|, |P·N·N·|). Some cards also include a small map or a scale bar. The artifacts are: 1. PNN (Elsenne-Allyet 2004), 2. PNN (Elsenne-Allyet 2004), 3. PNN (Elsenne-Allyet 2004), 4. Ponsich (1974) *, 5. PNN (Ponsich 1982), 6. PNN (Ponsich 1974) *, 7. P·N·N· (Ponsich 1982), 8. PNN (Chic 1985) *, 9. PNN (Elsenne-Allyet 2004), 10. P·N·N· (Ponsich 1982), 11. P·N·N· (Ponsich 1982).

In order to facilitate the use of the visualization tool, a series of demos and tutorials have been realized and accessible from the home page.

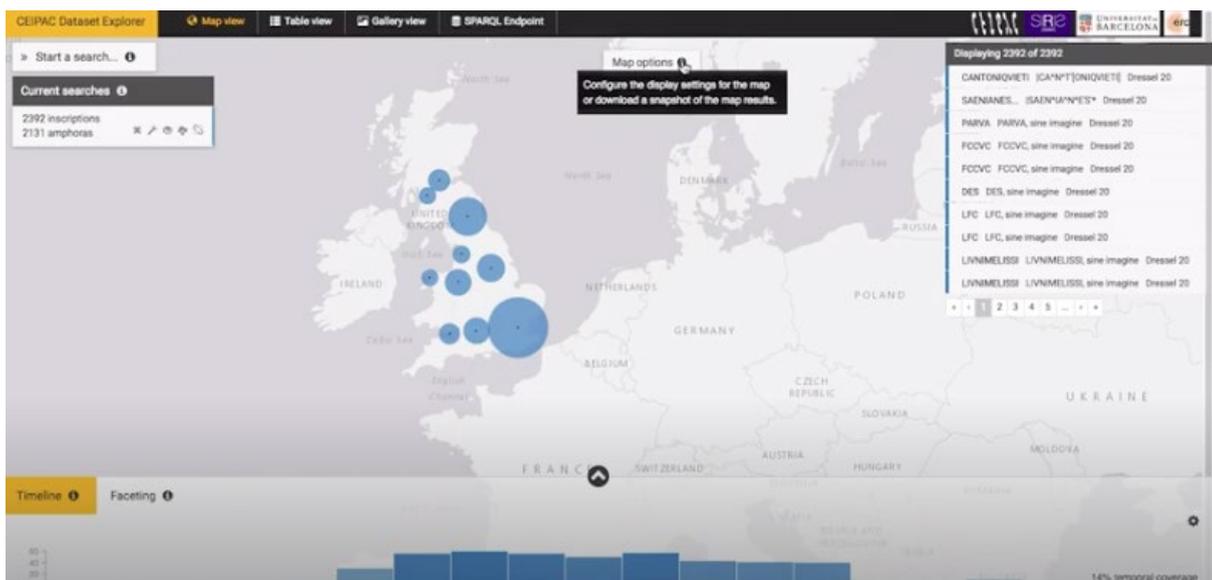
A use-case example:

To demonstrate some of the capabilities of the tool, here the series of actions performed by a user aiming at downloading data from “*searching amphoras nearby the Hadrian’s Wall*”:

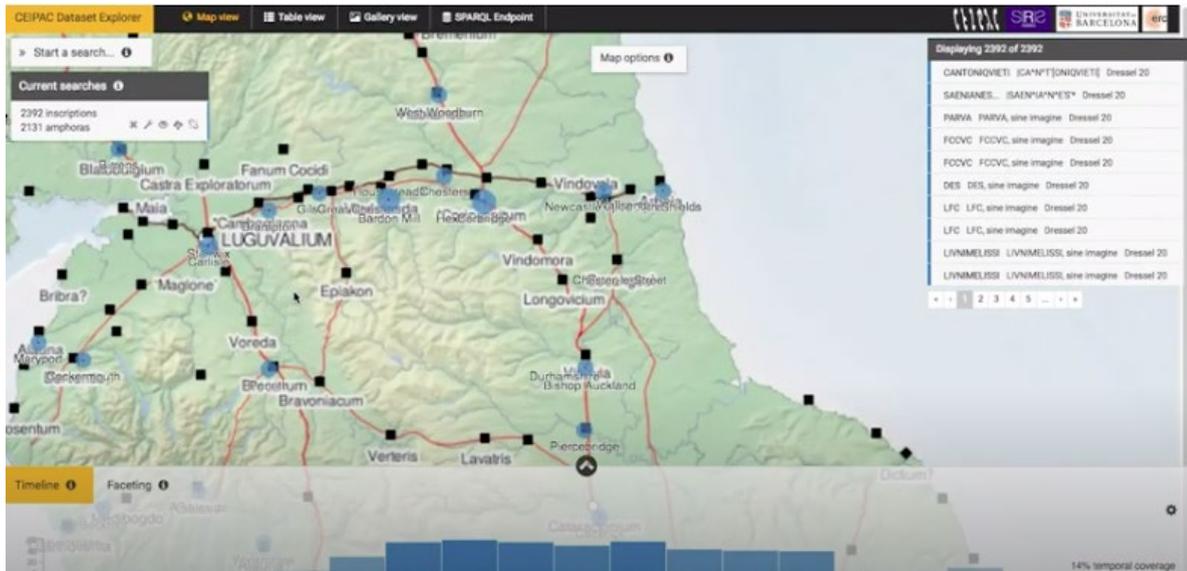
1. By using the advanced search, filter results only for amphoras found within United Kingdom:



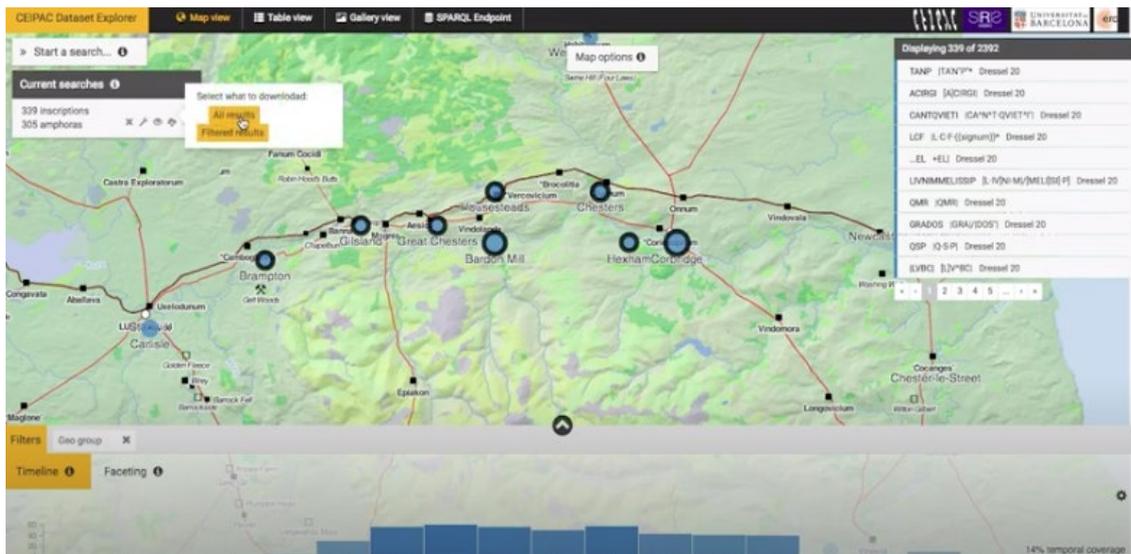
2. Once the results are obtained, the user can change the type of map by selecting the roman Empire ancient map.



3. The user performs zooming and panning option to frame the area of interest (North of Britain)



4. After selecting places of interest around the Hadrian's Wall (MaryPort, Carlisle, Bardon Mill (Vindolanda), Corbridge, ...), the user can download the data filtered by the geographical selection performed by the user.



3. From technology change to research-culture change

Over the last ten years, the amount of data available in the so-called linked data cloud has continued to grow very rapidly. Although a huge amount of data is now also available in the cultural heritage and digital humanities sectors, there are still few examples of reuse of such data, especially in the field of ancient history and archaeology. In other words, there exists an increasing number of research projects working for developing innovative technological solutions for data integration and access, however few scholars and research groups within the humanities are really using them.

This shortage requires more attention and understanding in order to take advantage from the research efforts realised in the last years. For fully exploiting the enormous potential offered by semantic technologies and linked data it will be necessary to face the challenge of contributing not only to a technological change but also to a new academic and scientific culture: a culture that needs to prioritize the processes of data management and data sharing. In sectors such as digital humanities, still dominated by a strong tradition from the two-dimensional world of print, facing these challenges will be more difficult and more urgent than in other sectors.

In this perspective, the first lesson learnt from the EPNet is that publishing data in open format, using the most advanced semantic technologies is not enough if you don't train scholars and students in a) learning how to use these technologies and b) how to contribute to publishing their data in the same format.

The second lesson is the urgent need to change from a culture of controlling data for individual benefit to the capacity of linking and contributing to data sharing and integration. The questionable practice of publishing the explanation before the access to the full datasets of observations is an issue that the humanities in general, and ancient history in particular, should seriously deal with. The attitude of not publishing full datasets for not creating the conditions (full access to the data) for someone else to be able to propose a better or before explanation should be strongly fought in the academic environment, as it happens in other scientific disciplines.

The tension between promoting scientific goals in spite of protecting the privacy interests of study participants should be overcome, as this happened and is happening in other scientific fields during the last decades. Science is becoming more and more collaborative and without embracing the open science paradigm, humanities and ancient history will be even more relegated to the periphery.

Projects like EPNet demonstrate that it is much easier to dedicate effort to improve technological solutions (e.g from a classical relational database to a fully integrated semantic platform), but definitely harder changing the research culture in which these solutions are built and exploited. The challenge is how to facilitate this research-culture change and for this reason future investments should be made in that direction.

References

Calvanese *et al.* 2015 = Calvanese, D., A. Mosca, J. Remeseal, M. Rezk and G. Rull. 2015. A 'Historical Case' of Ontology-Based Data Access, in *Proceedings of the 2015 Digital Heritage International Congress 2*: 291-298. IEEE, Granada.

Calvanese *et al.* 2016 = Calvanese, D., P. Liuzzo, A. Mosca, J. Remesal, M. Rezk and G. Rull. 2016. Ontology-based data integration in EPNet: Production and distribution of food during the Roman Empire. *Engineering Applications of Artificial Intelligence* 51: 212-229.

McInerney *et al.* 2014 = McInerney, G.J., M. Chen, R. Freeman, D. Gavaghan, M. Meyer, F. Rowland, D.J. Spiegelhalter, M. Stefaner, G. Tessarolo and J. Hortal. 2014. Information visualisation for science and policy: engaging users and avoiding bias. *Trends in ecology & evolution* 29(3): 148-157.

Mosca *et al.* 2015 = Mosca, A., J. Remesal, M. Rezk and G. Rull. 2015. Knowledge Representation in EPNet, in T. Morzy, P. Valduriez and L. Bellatreche (eds.) *New Trends in Databases and Information Systems: ADBIS 2015. Poitiers, France, September 8-11, 2015. Communications in Computer and Information Science Proceedings* 539: 27-437. Berlin: Springer.

Porter, M.A. and K. Boerner. 2011. Atlas of Science Visualizing What We Know. *Science* 331(6018).

Remesal *et al.* 2014 = Remesal, J., A. Díaz-Guilera, B. Rondelli, X. Rubio, A. Aguilera, D.J. Martín-Arroyo, A. Mosca and G. Rull. 2014. The EPNet Project. In: *Information Technologies for Epigraphy and Cultural Heritage. Proceedings of the First EAGLE International Conference (Paris, September 29-30 - October, 1, 2014)*: 455-464. Roma.

Remesal, J. and B. Rondelli. 2017. The EPNet Project: a non-conventional framework for falsifying historical theories, in J. Remesal Rodríguez (ed.) *Economía romana. Nuevas perspectivas / The Roman economy. New perspectives* (Instrumenta 55): 119-124. Barcelona: Publicacions de la Universitat de Barcelona.