

# The Semantics of Historical Knowledge. Labelling Strategies for Interdisciplinary and Digital Research in History

Esther Travé Allepuz<sup>1</sup> and Pablo del Fresno Bernal<sup>2</sup> and Alfred Mauri Martí<sup>3</sup> and Sonia Medina Gordo<sup>1</sup>

**Abstract.** This short paper aims at introducing some labelling concepts developed and used by the authors. Indeed, developing a conceptual framework for interdisciplinary research in history is a much-needed strategy in order to ensure that historians use all vestiges from the past regardless their origin or support for the construction of historical discourse. Fixing the semantics of historical knowledge is the first unavoidable step to build a new scenario in which NLP tasks enable more efficient data gathering and exploitation.

## 1 INTRODUCTION

Historical knowledge is a construction of the past built from its vestiges, thoroughly examined and assessed in order to purify them from its potential bias. In vestiges left from former époques, historians search for information related to two key interdependent concepts: time and change. These are ontological concepts for past construction –in K. Thibodeau’s terms [28]– because the particular coordinates of both determine the existence of events. In other words, historical facts are a continuum of elements balanced between permanence and change, what we call Historical Time [3].

Data related to time and change might be present in any written, material or immaterial vestige. Hence, data gathering and exploitation must surmount the academic fragmentation of information sources in order to build an *integrated* discourse. The spatial and material turns in history have led historians to a more complete and accurate reflection of the past. Nonetheless, the digital turn occurred in many Social and Human Sciences still finds an unreceptive reaction when coming to History, and data managing strategies have been widely discussed [11].

Far from being overwhelmed by the unknowns of this domain, a few exceptions deal with different ways of representing historical information [18-19] and the semantic definition of historical ontology building [12, 24, 29]. Recent experiences focus on quantitative data analyses [5] and, predominantly, on written historical texts [1, 16]. Some of them struggle to find the best ways to deal with bias [9] and uncertainty [25]. Despite this, a normalized user-friendly code to exploit vestiges of different nature and support is still missing and historical knowledge seems to be restricted to its written apparel. We acknowledge the capability of hybrid intelligence for natural language processing as

a must-use tool to speed up data gathering and exploitation processes and to open a brand-new field for historical research in which new and more complex questions can be asked to past vestiges. In doing so, research itself acquires a FAIR character [32], ensuring the reliability and traceability of past construction. Furthermore, available tools should allow us to deal with massive datasets, some of which have been disregarded until recently as marginal or non-significant.

This new scenario requires an effort from different disciplines in order to explore common languages and codes which become able to identify, register and exploit common and exchangeable units of information regardless of the specificity of our areas of expertise, scientific domain or sources used. According to these needs, we offer a brief layout of interdisciplinary Semantics of Historical Knowledge and the main concepts that have proved to be operational in our domain in order to develop an integrated historical approach.

## 2 DATA MODELLING AND LABELLING CATEGORIES

One of the most common practices in History when approaching archival capital is to read ancient documents endless times until you get an exact idea of their content and implications. Frequently, historians take brief notes about the information discovered or write down the archival reference of the set of files concerning the researcher’s specific field of interest. Apparently, there is nothing wrong in doing so and, definitely, accurate dissertations and essays have been written through this method.

Unfortunately, as time passes by, references and notes are no longer used and successive generations of historians need to go back again to the original file in order to increase our knowledge of past societies, or to review historical discourses under the perspective of a new state-of-the-art, or to address new questions to written vestiges. In addition, non-normalized data obtained through this procedure are hardly ever comparable to other sources of information, particularly if a published reflection is missing. According to M. D. Wilkinson and colleagues, historical science as such would certainly not be FAIR [32].

Archaeological method [8] as performed nowadays forces archaeologists to keep a standard register [15] of what they excavate, since the very same act of digging out a site destroys its materiality. Archaeologists will never be able to *read* the site again, as it will cease existing after the fieldwork is completed. As archaeologists and historians ourselves, we are concerned about

<sup>1</sup> History and Archaeology department, Universitat de Barcelona, Spain, [esther.trave@ub.edu](mailto:esther.trave@ub.edu); [smedingo12@alumnes.ub.edu](mailto:smedingo12@alumnes.ub.edu)

<sup>2</sup> Sistemes de Gestió de Patrimoni SCCL; Spain, [pdfsgp@gmail.com](mailto:pdfsgp@gmail.com)

<sup>3</sup> Centre d’Estudis Martorellencs, Spain, [bnn@heraclit.net](mailto:bnn@heraclit.net)

\* Corresponding author: Esther Travé Allepuz

having common codes and units of information not only to get the right balance between written and material evidences when building the historical discourse, but also to include many other sources of historical data. Indeed, iconography, linguistics, journalism, literature... deal with vestiges of the past and specialists on these domains contribute to the development of historical knowledge within a co-creation scenario.

As a result, we used several semantic concepts that define the minimum informative data present in vestiges regardless of their origin, purpose, nature, and support. These are *Unit of Topography (UT)* and *Actor (Ac)*. Both are ontological concepts to identify historical facts: entities existing and changing all the time, and events as an expression of what alters permanence. Even though the concept of UT owes so much to Harris' *Unit of Stratigraphy (US)* [15] (p. 42), researchers from all kinds of Social and Human Sciences could identify and label these in any source of information whether textual, material or audiovisual. That would lead to the opening of promising interdisciplinary challenges.

## 2.1 Units of Topography and Actors

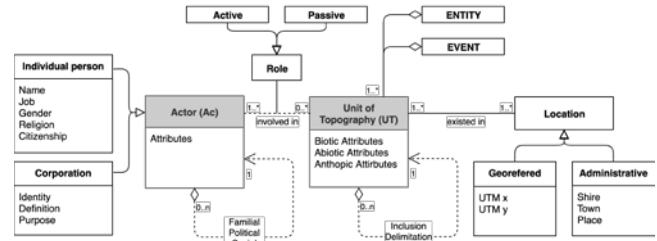
As defined by K. Thibodeau [28] (p. 7), an Entity is something that existed and an Event is something that happened or was done. Entities and Events have relationship of involvement, as every event involves at least one entity that might be the participant in the event, its observer, the mechanism for the event to happen, or the object altered by the event itself. In terms of data-labelling, the categories *Unit of Topography* and *Actor*, as defined by A. Mauri [22] (p. 45), and their relations, provide the unique and univocal identifiers for historical facts regardless of their link to permanence (Entity) or change (Event), or the nature and support of the vestige.

- **Unit of Topography (UT):** It is the evidence of an action or situation that can be located in space and time, regardless of the specificity of the information source and its biotic, non-biotic or anthropic attributes. Each UT has a specific location and date. Location can be expressed as a UTM coordinate or as an administrative delimitation that might have changed through time.
- **Actor (Ac):** It is the individual or corporative, active or passive, protagonist of an action identified as a UT. If being an individual, its attributes are their name, gender, religion, citizenship, date of birth and death, etc. Different individual actors gathered for a given period of time with a particular purpose and under determinate conditions can act as corporative actors.

Several types of relationships can be set between UT and Ac. A UT can include, link or delimitate another UT. Hence, Inclusion, Delimitation and Link are classes of the UT-UT relation. An Actor always plays an active or passive role within a UT, so Role is the only class of Ac-UT. Actors can relate to other actors through familial, political, social or economic Ac-Ac relationships. Some of them can turn an assemblage of individual actors into a corporative one. Being the two the main labelling categories, written sources in particular can provide information about values or prices that are labelled accordingly by means of a Value (V) label. Values usually are mechanisms for the Actors to perform new UTs.

According to our data modelling, the UML diagram shown in figure 1 expresses the ontological concepts and their relations [13] as classes, which does not get into contradiction with other existing

proposals [22] (p. 203). Our labelling proposal owes considerably to the interpretation of the analyst, which might seem paradoxical, due to the existence of automatized tools such as XML text encoding [6]. Nevertheless, only through the identification and registration of UT, Ac and their attributes, we are capable of exploiting historical data regardless their written, material or non-material character.



**Figure 1.** UML diagram of ontological concepts –UT, Ac and their relations– for integrated historical research.

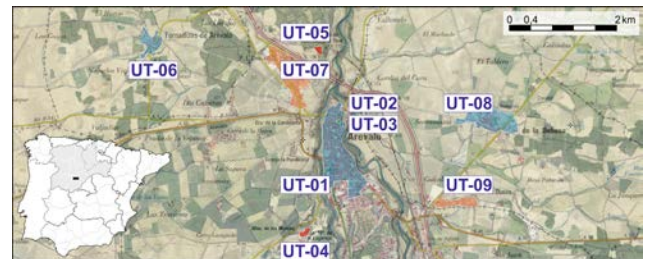
## 2.2 Procedure and examples

In the following section, we provide several examples of UT/Ac identification and labelling as a brief demonstration of methodological procedure and the potential of data exploitation. Examples include multiscale and interdisciplinary primary or secondary sources related to the site of Arévalo (Ávila, Spain). These examples have been selected in order to represent how different UT and Ac can be identified in different textual and non-textual sources and exploited accordingly.

### 2.2.1 Cartographic sources

The first example is the location map of the site and the immediate neighbourhood represented in figure 2 [17]. We identified and labelled UT through image processing software. Notice that UT identification can be as exhaustive as required by the sphere of interest determined by the researcher. In this case, only urban areas and historical buildings have been recorded.

#### Example 1 Cartography of Arévalo (Ávila, Spain)



**Figure 2.** Map 1:50.000 and aerial view transparency of Arévalo and its neighbourhood [17]. UT are identified and labelled as polygons.

### 2.2.2 Photographic vestiges

The church had some attached chapels, unfortunately demolished during the refurbishment works carried out in 1969 and 1970. Ancient pictures taken before the chapel's demolition are the last vestiges of their architectural features, and we selected them as the second example. Photographic record of the past and present building together with UT identification is shown in figure 3.

## Example 2 Photographs of the Church of Saint Mary



**Figure 3.** View of the Church of Sant Mary before 1969 –left [2] and centre [14] (p. 9)– and after the latest refurbishment work on the right [26].

### 2.2.3 Written primary files

Medieval documents provide information about the organization of the Ávila territory. In 1140, Pope Innocent II confirmed the possessions of the Bishop of Ávila, named Íñigo. He also gave him and his successors control over the churches in Ávila, Arévalo, Olmedo and Alcazarén<sup>2</sup> [4] (p. 3-4). Almost eighty years later, in 1179, Pope Alexander III confirmed the possessions of the new Bishop of Ávila –named Sancho– in a similar document<sup>3</sup>. Then, he gave him and his successors control over the churches in Ávila, Arévalo and Olmedo and over some other monasteries. We selected this second document as example 3 and labelled<sup>4</sup> it accordingly identifying UT and Ac on the excerpt transcript below.

**Example 3** <Ac01 Alexander> <Att-Ac01 episcopus>, <Att-Ac01 servus servorum Dei>, venerabili fratri<sup>5</sup> <Ac02 Sancio>, <Ac02-UT27 Abulensis episcopo>, <Ac02-Ac03 eiusque> <Ac03 successoribus> <Att-Ac03 cononice> substituendis in perpetuum. (...) Ea propter venerabilis in Christo frater episcopus tuis iustis postulationibus clementer annuimus et <UT18 ecclesiam abulensem>, et a Deo auctore, preesse dinosceris sub Beati Petri et <Ac01-UT19 nostra> <UT19 protectione suscipimus> et presentis <UT20 <Att-UT20 scripti> privilegio> <UT21 communi-mus>, <UT22 statuentes ut quascumque <UT23 possessiones>, quecumque bona eadem ecclesia in presentiarum iuste et canonice possidet aut in futurum concessione pontificum, largitione regum vel principum, oblatione fidelium seu aliis iustis modis Deo propitio, poterit adipisci, firma vobis vestrisque successoribus et illibata permaneant>, in quibus hec propriis duximus exprimenda vocabulis <UT24 monasteria Sancte Marie de Fundo>, <UT04 Sancte Marie de Gomez Roman> et <UT25 ecclesias>, <UT25-UT28; UT25-UT29; UT25-UT30 quas> <UT26 Abule>, <UT01 Arevali>, <UT27 Ulmeti> et in <UT28; UT29; UT30 terminis locorum ipsorum> habere dinosceris et libertatem omnium ecclesiarum tui episcopatus, quas pleno iure possidet ecclesia tua, et nulla alia in eis persona vel ratione patronatus vel quolibet alio

modo aliquid valeat vindicare. (...) <UT20-UT31 Datum> <UT31 Laterani> <Ac04-UT20 per manum> <Ac04 Alberti>, Sancte Romane ecclesie <Att-Ac04 presbyteri, cardinalis et cancellarii>, <Date-UT20 XI kalendas maii, indictione XII, Incarnationis Dominice anno MCLXXVIII, pontificatus vero domini Alexandri papae III, anno XX>. [4] (p. 13-15)

### 2.2.4 Bibliographic reflections

As many medieval buildings, the church has been object of detailed analyses from several perspectives, which are considered reflections in Thibodeau's terms [28] (p. 14). Example 4 is a short excerpt of an art-architectonic study of Saint Mary's church. We labelled data using the same code and exploited them accordingly to demonstrate the validity of UT and Ac as interdisciplinary ontological concepts.

**Example 4**<sup>6</sup> Probably, <UT03 Santa María la Mayor> was one of the first <UT25 churches> <UT32 built> in <UT01; UT29 Arévalo><sup>7</sup> <Date-UT32 during> the <UT33 repopulation> in the <Date-UT33 12<sup>th</sup> century AD>, and its <UT17 tower> was probably <UT34 built> <Date-UT34 at the same time>. The <UT03 church> <UT02-UT03 is located> in the <UT02 Plaza de la Villa> <UT02-UT35; UT02-UT36 together with> other buildings of great architectural value such as <UT35 the church of San Martín> and the <UT36 Casa de los Sexmos>. <UT03-UT17 Santa María's> <UT17 tower> <Ac05-UT17 belonged to> the <Ac05 Briceño> <Att-Ac05 family> and the <UT03 church> <UT03-UT37 was> <Ac05-UT37 its> <UT37 burial place>. The church <UT17 tower> <UT34 rises> <UT17-UT38 above> an <UT38arch> <UT38-UT39 through which> the <UT39 street> passes by, which is the most striking feature of the building. [20] (p. 4)

## 2.3 Data gathering and exploitation

We have shown several vestiges on different supports and identified the historical data contained within them. Vestige labelling is just a strategy to make UT/Ac identification easier, as the process has a strong historical interpretative component attributed to the analyst. Information is then included in a database built according to the basic research processes: Source register, data gathering and analysis or data exploitation. Figure 4 shows a screen view of the data gathering interface.

Obtained data are stored in tidy-structured tables with variables in columns and observations in rows [31]. Tables 1 and 2 below show the data extracted from the examples labelled above regardless of the support or nature of the source.

**Table 1.** Simplified Ac dataset gathered from examples 1 - 3.

Ac Id	Name	Type	Attributes	Related UT/Ac
01	Alexander III	Individual	Bishop, Pope	UT19; UT21; UT22
02	Sancho	Individual	Bishop	Ac03; UT23; UT26
03	Bishopric	Corporative	Bishop; canonical election	Ac02; UT23; UT26
04	Albert	Individual	Priest; Cardinal; Chancellor	UT20
05	Briceño	Corporative	Family; Lineage	UT17; UT37

<sup>2</sup> AC (Archive of the Cathedral of Avila), Section 'Documentos', num. 1, Original Document.

<sup>3</sup> AC (Archive of the Cathedral of Avila), Section 'Documentos', num. 6, Original Document.

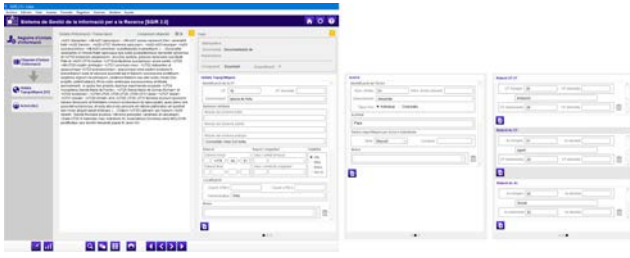
<sup>4</sup> Labelling code: <UT00> <Ac00> <Attribute-UT00> <Attribute-Ac00> <Date-UT00>. Relations are labelled in accordance with concepts related: <Ac00-UT00> <Ac00-Ac00> <UT00-UT00>. A semicolon separates different UT, Ac, Attributes, or Relations identified through the same word or syntagmatic expression.

<sup>5</sup> Notice that the word *fratri* here cannot be interpreted as a familial relationship between both Ac (Pope Alexander and Bishop Sancho), as it is a religious vocation. Hence, this is an illustrative example of the interpretative task of the analyst in order to understand the text carefully, and identify correctly these relations in order to complete the database.

<sup>6</sup> English translation from the Spanish reference by E. Travé.

<sup>7</sup> Texts can occasionally be ambiguous, particularly reflections. Here it is not clear if the term «Arévalo» refers to the urban nucleus or to the municipality.





**Figure 4.** SGIR 2.0 database screen view. Labelled example 3 is shown on the screen and forms for UT (left), Ac (centre) or Relation –UT-UT, Ac-UT and Ac-Ac– (right) can be displayed alternatively for data introduction.

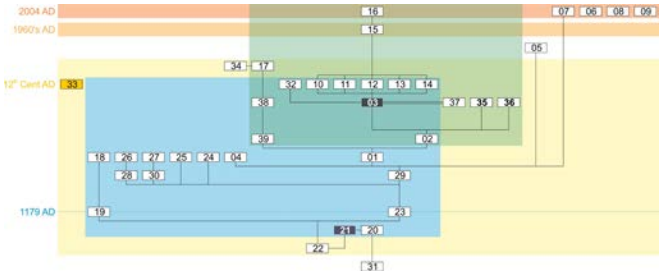
**Table 2.** Simplified UT dataset gathered from examples 1 - 3.

UT Id	Brief description	Related UT/Ac	Location	Attributes	Date
01	Arévalo (Urb. nucleus)	UT02, 03, 07, 09, 29, 39	= UT29	Urban area	
02	Plaza de la Villa	UT03, 35, 36	= UT01	Urban layout	
03	Church of Santa María	UT10, 11, 12, 13, 14, 16, 17, 32	Longitude Latitude <sup>8</sup>	Structure	1179
04	S. M <sup>a</sup> de la Lugareja	UT23	Ávila (Prov.)	Structure	1179
05	Santa María la Real	UT29	Ávila (Prov.)	Structure	
06	Tornadizos de Arévalo		Ávila (Prov.)	Place name	
07	Industrial Quarter	UT29	Ávila (Prov.)	Urban layout	
08	Martín Muñoz de la D.		Ávila (Prov.)	Place name	
09	Las Dunas (Urban area)		Ávila (Prov.)	Place name	
10	(SM) South Façade	UT03	=UT03	Structure	12 <sup>th</sup> C.
11	(SM) Chapel 1	UT03	=UT03	Structure	12 <sup>th</sup> C.
12	(SM) Chapel 2	UT03	=UT03	Structure	12 <sup>th</sup> C.
13	(SM) Chapel 3	UT03	=UT03	Structure	12 <sup>th</sup> C.
14	(SM) Bell Gable	UT03	=UT03	Structure	12 <sup>th</sup> C.
15	Chapel's demolition	UT11, 12, 13, 16	=UT03	Destruction	1960's
16	Building refurbishment	UT03, 15, 17	=UT03	Construction	2004
17	(SM) Tower	Ac05 UT03, 16, 38	=UT03	Structure	12 <sup>th</sup> C.
18	Ávila's Church	UT19	Ávila (Prov.)	Entity	1179
19	Pope's protection	UT18	=UT28	Political action	1179
20	Privilege scripture	UT31	=UT31	Scripture	1179
21	Privilege concession	UT19, 23	=UT31	Political action	1179
22	Pope's privilege	UT23	=UT31	Political action	1179
23	Bishop's possessions	UT04, 22, 24, 25	Ávila (Prov.)	Ownership	1179
24	S. M <sup>a</sup> de Burghondo	UT23		Structure	1179
25	Churches	UT26; 27, 28	Ávila (Prov.)	Structure	1179
26	Ávila (Urban nucleus)	UT25, 28	=UT28	Urban area	1179
27	Olmedo (Urb. nucleus)	UT25, 30	=UT30	Urban area	1179
28	Ávila (Municipality)	UT25, 26	Ávila (Prov.)	Place name	1179
29	Arévalo (Municipality)	UT01, 05, 07, 25	Ávila (Prov.)	Place name	1179
30	Olmedo (Municipality)	UT27	Ávila (Prov.)	Place name	1179
31	Lateran	UT20	Rome	Place name	1179
32	Santa Maria's Building	UT03	=UT03	Construction	=UT33
33	Repopulation		Ávila (Prov.)	Political action	12 <sup>th</sup> C.
34	Tower's Building	UT35	=UT03	Construction	=UT33
35	Church of San Martín	UT02, 34	=UT02	Structure	=UT34
36	Casa de los Sexmos	UT02	=UT02	Structure	
37	Mausoleum	Ac05	=UT03	Burial	
38	Arch	UT17, 39	=UT17	Structure	=UT34
39	Street	UT01, 38	=UT01	Road	

When gathering and storing data in this way, we can represent relations quite easily through flux diagrams and matrices to establish the temporal sequence of activities, and their permanence or transformation, in a visual way. Figure 5 shows an extract of a

<sup>8</sup> 41° 03' 58.61" N; 4° 43' 11.69" W

historical Harris-like [15] matrix created for the Church of Saint Mary, in Arévalo (Ávila, Spain), which has been the main object of our example selection.



**Figure 5.** Historical matrix of the Church of Saint Mary in Arévalo Ávila. It can be completed and enlarged through further research [7].

### 3 DISCUSSION, FUTURE DEVELOPMENT AND CONCLUDING REMARKS

The most striking point of using Units of Topography and Actor as ontological concepts of Historical semantics is that they allow for a truly interdisciplinary and integrated construction of the past. In recent years, data modelling and database construction has allowed us accordingly to develop integrated approaches [21, 30] and software [10] overcoming the traditional inconveniences arising from the fragmentation of sources of information.

The proposal of UT/Ac gathering is an adequate compromise solution in order to develop an ontology for past construction in which entities and events are located within precise spatiotemporal coordinates. This actually implies more interpretative knowledge on the historians' part, as it is not always possible to detect these data units through mere automatic data labelling applications yet.

Despite TEI [27] being one of the most successful XML experiences [6] in the linguistics domain, the process does not seem to be proficient enough in the identification of entity and events as required by historical knowledge. Units of Topography and actors are represented in too many different shapes, and all supports must be considered, not only textual –even if textual sources are the most abundant.

Hybrid intelligence would be, to our perception, a challenging field to explore the possibilities of historical knowledge to become digital and interdisciplinary, and to develop appropriate UT/Ac recognition patterns. Ontology-mediated databases are key to ensure data exchange. Nowadays, our research team is working on SGIR 2.0 development, a database for UT/Ac gathering and management. The short summary we offered aimed at introducing the main ontological concepts in use and showing data gathering procedure according to the wide variety of sources available to us.

### ACKNOWLEDGEMENTS

This study is part of the current research tasks carried out by the Medieval and Postmedieval Archaeology Research Group GRAMP.-UB (2017-SGR-833-GRC) of the University of Barcelona, to which the authors belong, and it is included in our Landscape Archaeology research line. Theoretical development, conceptualization, and applied methods are part of a funded research archaeological project developed by the Centre d'Estudis Martorellencs (CLT009/18/00036 – DGPC/exp. 35). The authors wish to thank Ms Noemí Travé for language edition and review.

## REFERENCES

- [1] A. Antonacopoulos, D. Karatzas. 'Semantics-Based Content Extraction in Typewritten Historical Documents', In *Proceedings of the 8<sup>th</sup> International Conference on Document Analysis and Recognition (ICDAR2005)*, IEEE Digital Library, 2005, p. 48-53
- [2] Arévalo. *Plaza de la Villa e Iglesia de Santa María*. Available online: [https://www.todocoleccion.net/postales-andalucia/arevalo-avila-plaza-villa-e-iglesia-santa-maria~x47348260#formas\\_de\\_pago\\_y\\_envio](https://www.todocoleccion.net/postales-andalucia/arevalo-avila-plaza-villa-e-iglesia-santa-maria~x47348260#formas_de_pago_y_envio) (accessed on 27<sup>th</sup> January 2020).
- [3] J. Arostegui. *La investigación histórica: teoría y método*, Crítica, Barcelona, 1995.
- [4] Á. Barrios. *Documentación medieval de la Catedral de Ávila*, Universidad de Salamanca, Salamanca, 1981.
- [5] H. Berger. 'Microhistory and quantitative data analysis', In *Humanities, Computers and Cultural Heritage* (op. cit), p. 39-42.
- [6] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, F. Yergau (Eds.). *Extensible Markup Language (XML) Version 1.0 (Fourth edition)*, 16 August 2006. Available online: <http://www.w3.org/TR/REC-xml/> (accessed on 3<sup>rd</sup> February 2020).
- [7] J. A. Calvo. 'La restauración de la diócesis de Ávila y el episcopado de la repoblación (1085 – 1198)', *Studia Historica. Historia Medieval*, **37/2**, 155-178 (2019).
- [8] A. Carandini. *Historias en la tierra: Manual de excavación arqueológica*, Crítica, Barcelona, 1997.
- [9] C. Coffin. 'The voices of History: Theorising the interpersonal semantics of historical discourses', *Text*, **22/4**, 503-508 (2002).
- [10] P. Del Fresno. *Sistema de Información Arqueológica: propuesta de normalización, desarrollo conceptual e informático*. PhD Dissertation. Euskal Herriko Unibertsitatea – University of the Basque Country, Vitoria, 2016. Available online: <https://www.academia.edu/23836202/> (accessed on 30<sup>th</sup> September 2020).
- [11] J. Delve, R. G. Healey. 'Is there a role for data warehousing technology in historical research? In *Humanities, Computers and Cultural Heritage* (op. cit), p. 106-111.
- [12] S. Eriksson, H. Skånes. 'Addressing semantics and historical data heterogeneities in cross-temporal landscape analysis', *Agriculture, Ecosystems and Environment*, **139**, 516-521 (2010).
- [13] C. González-Pérez. 'A Conceptual Modelling Language for the Humanities and Social Sciences', In *Sixth International Conference on Research Challenges in Information Science (RCIS)*, IEEE Digital Library, 2012.
- [14] J. L. Gutiérrez. 'Memoria Mudéjar en La Moraña: la arquitectura', In *Memoria Mudéjar en La Moraña*, ASODEMA/ Proyecto Leal, Ávila, 2011, p. 7-122.
- [15] E. C. Harris. *Principles of Archaeological Stratigraphy*, Academic Press Inc., San Diego, 1989.
- [16] M. Heller, G. Vogeler. 'Modern information retrieval technology for historical documents', In *Humanities, Computers and Cultural Heritage. Proceedings of the XVI International Conference of the Association of Arts and Sciences*, Royal Netherlands Academy of Arts and Sciences, Amsterdam, 2005, p. 143-148.
- [17] Instituto Geográfico Nacional. *Visor. Iberpix 4*. Available online: <ign.es/iberpix2/visor/> (accessed on 27<sup>th</sup> January 2020).
- [18] A. Ivanovs, A. Varfolomeyev. 'Editing an exploratory analysis of medieval documents by means of XML technologies', In *Humanities, Computers and Cultural Heritage* (op. cit), p. 155-160.
- [19] F. de Jong, H. Rode, D. Hiemstra. 'Temporal language models for the disclosure of historical text', In *Humanities, Computers and Cultural Heritage* (op. cit), p. 161-168.
- [20] E. Martín. 'Arquitectura religiosa en Arévalo: Del Románico al Barroco', In *I Jornadas de Introducción a la Historia del Arte a través del Patrimonio de Arévalo*, 2013. Available online: <https://www.academia.edu/4251068/> (accessed on 27<sup>th</sup> January 2020).
- [21] A. Mauri. *La configuració del paisatge medieval: el comtat de Barcelona fins al segle XI*. PhD Dissertation. Universitat de Barcelona, Barcelona, 2006. Available online: <https://www.tesisenred.net/handle/10803/2065> (accessed on 27<sup>th</sup> January 2020).
- [22] A. Mauri, E. Travé, and P. Del Fresno. 'An Integrated Implementation of Written and Material Sources – Conceptual Challenge and Technological Resources', In I. Ollich, Ed. *Archaeology. New Approaches in Theory and Techniques*, InTech, Rijeka, 2012, p. 41-64.
- [23] V. Mirzaee, L. Iverson, B. Hamidzadeh. 'Computational representation of semantics in historical documents', In *Humanities, Computers and Cultural Heritage* (op. cit), p. 199-206.
- [24] G. Nagypál. 'History ontology building: The technical view', In *Humanities, Computers and Cultural Heritage* (op. cit), pp. 207-214.
- [25] G. Nagypál, B. Motik. 'A Fuzzy Model for Representing Uncertain, Subjective, and Vague Temporal Knowledge in Ontologies', In R. Meersman, Z. Tari, D. C. Schmidt (Eds.). *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. OTM 2003. Lecture Notes in Computer Science*, vol. 2888, Springer, Berlin / Heidelberg, 2003.
- [26] Rowanwindwhistler. *Iglesia de Santa María la Mayor, Arévalo, provincia de Ávila, románico mudéjar*, 22, sept, 2010. CC BY-SA 3.0, Available online: <https://commons.wikimedia.org/w/index.php?curid=11569761> (accessed on 27<sup>th</sup> January 2020).
- [27] C. M. Sperberg-McQueen, L. Burnard *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Text Encoding Initiative Consortium, 2019. Available online: <https://tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf> (accessed on 3<sup>rd</sup> February 2020).
- [28] K. Thibodeau. 'The Construction of the Past: Towards a Theory for Knowing the Past', *Information*, **10**, 332, (2019).
- [29] B. Tomaszewski. 'Local Model Semantics, Categories, and External Representation: Towards a Model for Geo-historical Context', In B. Kokinov, D. C. Richardson, Th. R. Roth-Berghofer, L. Vieu (Eds.). *Modeling and Using Context. 6th International and Interdisciplinary Conference, CONTEXT 2007, Roskilde, Denmark, August 20-24, 2007, Proceedings. Lecture Notes in Artificial Intelligence*, vol. 4635, Springer, Berlin / Heidelberg, 2007.
- [30] E. Travé, 'Los hornos medievales y la cerámica negra' In *Actas del IX Congreso de la Asociación de Ceramología. Obra negra y alfarerías de cocina*, Asociación de Ceramología / Ajuntament de Quart, Quart (Spain), 2017, p.123-147.
- [31] H. Wickham, 'Tidy Data', *Journal of Statistical Software*, **59/10**, 1-10, (2004).
- [32] M. D. Wilkinson et al. 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, **3**, 160018, (2016).