ISSN: 2687-8402

DOI: 10.33552/OAJAA.2019.02.000526

Open Access Journal of Archaeology and Anthropology

Iris Publishers

Research Article

Digital Approaches to Historical Archaeology: Exploring the Geographies of 16th Century New Spain

Liceras-Garrido, Raquel[1]; Favila-Vázquez, Mariana[2]; Bellamy, Katherine[1]; Murrieta-Flores, Patricia[1]; Jiménez-Badillo, Diego[2]; Martins, Bruno[3]

[1]*Digital Humanities Hub - History Department, Lancaster University, United Kingdom*

[2]*Museo del Templo Mayor, Instituto Nacional de Antropología e Historia, México*

[3]*Instituto de Engenharia de Sistemas e Computadores - Investigação e Desenvolvimento and Instituto Superior Técnico, Universidade de Lisboa, Portugal*

***Corresponding author:** Raquel Liceras-Garrido, Lancaster University, History Department, B005 Bowland Main, Lancaster, LA1 4YT, United Kingdom.

Received Date: November 13, 2019

Published Date: November 21, 2019

Abstract

The humanities have always been concerned with ideas of space, place and time. However, in the past few years, and with the emergence of Digital Humanities and Computational Archaeology, researchers have started to apply an array of computational methods and geographical analysis tools in order to understand the role that space plays in the historical processes of human societies. As a result, historians and archaeologists, together with computer scientists, are currently developing digital approaches that can be used to address questions and solve problems regarding the geographies contained in documentary sources such as texts and historical maps. Digging into Early Colonial Mexico is an interdisciplinary project that applies a Data Science/Big Data approach to historical archaeology, focusing on the analysis of one of the most important historical sources of the 16th century in Latin America, called the Geographic Reports of New Spain. The purpose of this paper is to:

a) describe the nature of the historical corpus,

b) introduce the methodologies and preliminary results produced so far by the project, and

c) explain some of the theoretical and technical challenges faced throughout the development of the methods and techniques that supported the analysis of the historical corpus.

Keywords: Historical Archaeology; Digital Humanities; Spatial Humanities; New Spain; Geographic Information Sciences; Machine Learning; Natural Language Processing; Corpus Linguistics; Geographical Text Analysis; FAIR data

Abbreviations : AI: Artificial Intelligence; CA: Collocation Analysis; CONABIO: *Comisión Nacional para el Conocimiento y Uso de la Biodiversidad*; CL: Corpus Linguistics; DECM: Digging into Early Colonial Mexico; ENAH: *Escuela Nacional de Antropología e Historia*; FAIR: Findable, Accessible, Interoperable, and Reusable Data; GISc: Geographic Information Sciences; GIS: Geographic Information Systems; GTA: Geographical Text Analysis; INAH: *Instituto Nacional de Antropología Historia*; INEGI: *Instituto Nacional; de Estadística Geografía e Informatic*; KML: Keyhole Markup Language; LOD: Linked Open Data; ML: Machine Learning; NER: Named Entity Recognition; NLP: Natural Language Processing; OCR: Optical Character Recognition; RGs: *Relaciones Geográficas de Nueva España*; UNAM: *Universidad Nacional Autónoma de México*

Introduction

During the last few decades, quantitative and computational approaches have become increasingly popular in the different disciplines within the Humanities. Archaeology is one of the pioneering fields where quantitative methods have been increasingly in development since the 1960s. Following the post-processual approaches of the 1980s, in the last 40 years, one of the most commonly used digital tools in Archaeology has been Geographic Information Systems. Despite its popularity, archaeologists have had to deal with the problems derived from applying GIS. Some of these issues, according to Huggett [1], include the difficulties in managing the category of time, and the apparent reductionism in the ways data is categorised and complexity handled, amongst others. Interdisciplinary approaches aim to tackle these issues, with the development of tools to create, recover, and store large amounts of data going together with the textual emphasis of the Digital Humanities [2-4]. In this paper, we focus on the geographical data that we can recover from historical sources (both text and image), demonstrating that historical archaeologists can benefit greatly from the methodologies provided by this field.

Focusing on 16th century colonial sources, the 'Digging into Early Colonial Mexico' (DECM) project has three main and interconnected objectives. The first one is the development and improvement of computational methods and tools that facilitate access to, management, and analysis of historical sources, whether textual or cartographic. To do this, we combine methods and techniques from four subdisciplines of Computer Science, including Machine Learning (ML), Natural Language Processing (NLP), Corpus Linguistics (CL), and Geographic Information Sciences (GISc). ML is a branch of Artificial Intelligence that gives computers the ability to learn. Whilst CL focuses on how language is expressed within a corpus, NLP delves into the development of computational models for the analysis of human languages. Finally, within GISc, Geographic Information Systems (GIS) provide an environment to store, manipulate, visualise, and analyse spatial information. The second objective is the production of a range of digital datasets, derived from these historical sources, including:

i. a digital version of the RGs corpus annotated with an interoperable taxonomy;

ii. a digital atlas containing the physical, political, religious, and administrative geographies of New Spain, at the beginning of the 16th century; and

iii. a digital platform that connects the spatial information contained in the atlas with the text and image contents (i.e. paintings) of the RGs.

Our third and final objective, based on the digital methods, tools, and resulting data, is to answer research questions related, but not limited to: changes in settlement patterns between the Postclassic and Early Colonial period; progression and syncretism of Mesoamerican and Catholic cosmogonies; exploitation of resources and commercial networks; the role of indigenous and Spanish power institutions throughout the current territories of Mexico and Guatemala; and gender roles across class, societies, time and space. The paper begins with an introduction to the historical background of the colonial corpus of the Geographic Reports of New Spain (RGs), and the current location of this corpus. This is followed by a description of the methods and techniques developed in this project, including the creation of a digital gazetteer, the annotation process of textual and spatial information, and the creation of three 'StoryMaps' for public engagement. The final section provides the partial results and a discussion about the limitations that Corpus Linguistics and computational analysis techniques present when applied to documents in 16th century Spanish and indigenous languages, as opposed to the modern English sources that support most developments within the NLP field. Likewise, the difficulties that still present a challenge for the project are also explained at the end.

The Corpus: *Las Relaciones Geográficas de Nueva España*

The Geographic Reports of New Spain (RGs) are a set of textual accounts and graphic spatial representations that were compiled during the reign of Emperor Charles V and concluded under the reign of his son, Philip II. These aimed to gather all the available information of the kingdoms and territories under Spanish rule, including Castile, Aragon, Italy, Netherlands, Portugal (after 1581) and the four American viceroyalties: New Spain (which included Mexico, Guatemala and the Philippines), Nueva Granada, Perú and Río de La Plata [5] (Figure 1).

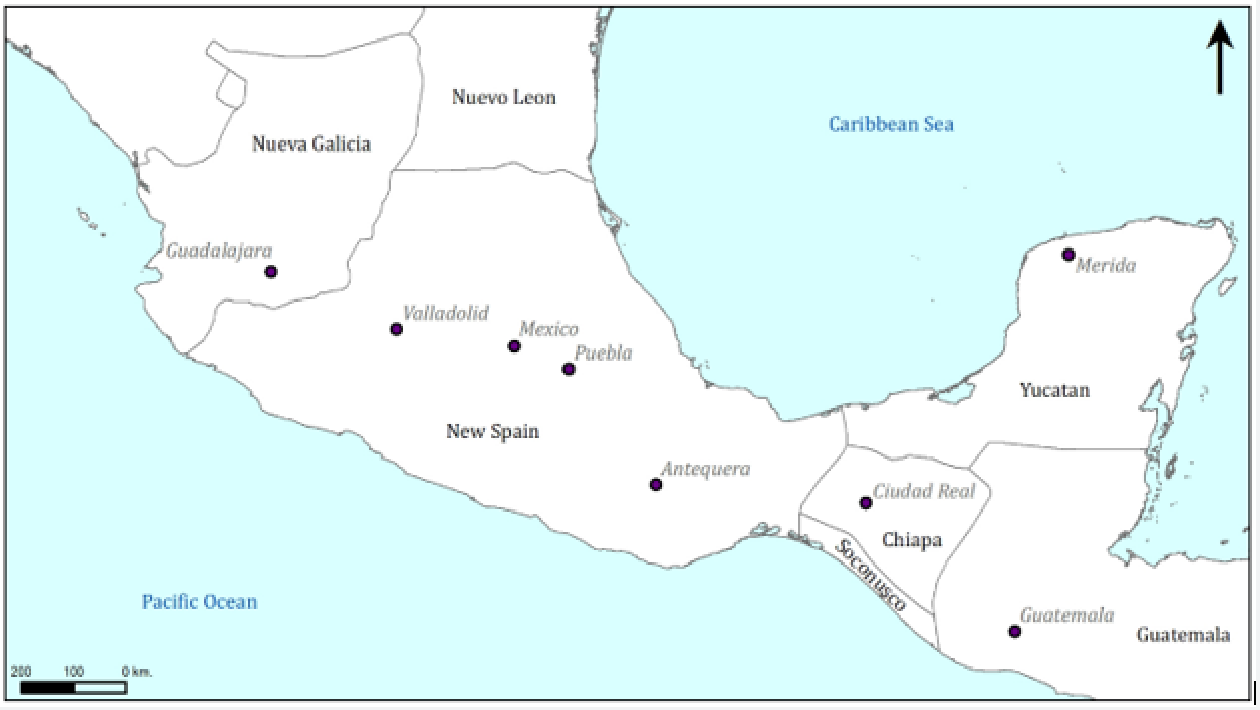


Figure 1: Gobiernos and main cities in 1580 [6: Figure 5].

Our corpus of RGs consists of reports compiled following Royal Cosmographer Juan López de Velasco's 1577 *Instrucción y Memoria*, which produced a more ordered series of reports than previous attempts [6,7]. The first reports to reach the Iberian Peninsula were those from Antequera (currently Oaxaca) and Yucatán between 1579 and 1581; later, those of the regions of Central Mexico, Michoacán, and Tlaxcala (1579-1582); and finally, those of Nueva Galicia (now Guadalajara) and Guatemala (1579-1585). The corpus includes a total of 2,800,000 words across 168 textual reports (26 of which are missing) in addition to 78 maps (16 of which are lost). These documents contain information on the 168 main *cabeceras* (town/district capital), the 248 subordinate *cabeceras*, around 414 towns, as well as smaller villages and farms [8]. Unfortunately, the spatial distribution is unequal, and according to H. Cline [9-10], lacking almost half of the existing civil jurisdictions in 1580. Despite this, we have 12 reports for Nueva Galicia, 17 for Michoacán, 34 for Mexico, 15 for Tlaxcala, 34 for Antequera, 54 for Tabasco and the Yucatán Peninsula, and 2 for Guatemala.

Composed of texts and pictorial maps, this corpus collects information primarily focused on the geography of New Spain, including rich data on rich data on the economy, religion, customs, and war activities of indigenous peoples a few years following the conquest. The corpus is, therefore, considered one of the most important sources of knowledge about the history of native groups, as well as their relationships with colonial Spanish officials.

Our project combines geospatial intelligence with the parsing of text, with the aim of facilitating the discovery of unsuspected data patterns and relationships among the entities mentioned in the collection of documents. The accumulation of information contained in this corpus has not been analysed jointly before, and can provide new interpretations of the social, spatial, economic, and ideological reconfigurations experienced by indigenous societies in the second half of the 16th century.

The original documents are mainly dispersed in four repositories: 80 reports are at the General Archive of the Indies (Seville, Spain); 46 at the Royal Academy of History (Madrid, Spain); 41 at the Latin American Collection of the Benson Library (University of Texas at Austin, USA) and 1 at the University of Glasgow (United Kingdom). The transcripts that form our corpus are the 54 RGs published by M. de la Garza [11] for Yucatán and Guatemala; the 114 RGs by R. Acuña [12] published in 10 volumes; and *'La Suma de la Vista de los Pueblos',* vol. 1 by F. del Paso and Troncoso[13], which provides additional information about the 16th century geographies. Alongside the texts, the corpus includes 78 maps preserved in the General Archive of the Indies (Seville, Spain), at the Royal Academy of History (Madrid, Spain), the University of Glasgow (Scotland), and the Benson Latin American Collection (University of Texas) [14]. The latest comprehensive analysis of these maps has been carried out by to Mundy [15]. Of these maps, 23 came from Mexico; 3 from Michoacan; 1 from Nueva Galicia; 21 from Oaxaca; 18 from Tlaxcala; 3 from Yucatan and 2 from Guatemala.

Our first task, therefore, was to obtain computer-readable files of the more comprehensive editions of the RGs. Thanks to the generosity of the National Autonomous University of Mexico (UNAM), we obtained René Acuña's 10 *Relaciones Geográficas* volumes [12], and Mercedes de la Garza's *Relaciones Histórico-geográficas de la Gobernación de Yucatán* [11]. To complement this, we also acquired Francisco del Paso y Troncoso´s *Papeles de Nueva España* in electronic format [13], though with significant errors. Consequently, this collection required pre-processing and re-applying of optical character recognition (OCR), as well as manual identification and correction of electronic transcription errors.

Developing Computational Methods for Historical Archaeology

Creating a digital dataset of 16th century geographies of New Spain: The DECM Gazetteer

The Geographic Reports of New Spain are rich in geographic references, therefore this corpus has been at the core of much of the historical and archaeological research of this period in the past century. Nevertheless, as can be expected, these geographies have substantially changed over time, and it was not until the end of the 1960's that scholars such as Peter Gerhard and Howard Cline started to compile more detailed information about these. Although the works in these (and other) authors have been widely published, they are usually only available in printed format. Before attempting to carry out spatial analyses or to ask questions, it was necessary to not only to create a digital version of this information, but also to conduct further research, investigating many of the toponyms that were not included in existing atlases created by other scholars. This work consists specifically of the linguistic and geographic disambiguation of toponyms.

Based on the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) Data and Linked Open Data (LOD), we are creating a digital directory, or gazetteer, that contains all the toponyms mentioned in the RGs related to settlements and geographic, political, or administrative features. To guarantee its interoperability with other systems and repositories, we have relied on the experience of previous projects such as the Alexandria Digital Library Gazetteer [16,17], and currently ongoing initiatives such as the World Historical Gazetteer. The well-established content standards from these projects offer the possibility to store and exchange not only geographic footprints (i.e. coordinates of latitude and longitude) of the place names, but also alternative spellings, the languages in which they are written, bibliographic references, and relationships, amongst others.

The first step in creating the gazetteer was to compile the information available of the 16th century place names contained in the RGs. To do so, we digitised the toponymy indexes of the aforementioned RGs editions: the 10 volumes of Acuña, the 2 volumes of De la Garza and the first volume of Del Paso y Troncoso. We standardised the indexes (as there was a considerable variety across all editions and volumes), and formatted them in plain text files, thereby ensuring they were in a computer-readable format. Due to name variations and inconsistencies, reconstructing 16th century toponymy poses many challenges. Although some place names have retained their original form, others have changed significantly over time as a result of linguistic evolution, including spelling variations, the use of Castilian transliterations of native names, or the addition of words such as the name of a patron saint or a natural resource typical of the area (e.g. Ixtapan de la Sal). Further challenges are encountered by one place having different names in different languages (e.g. Santiago Mitlantongo, Mictlantongo or Mitlantongo in Spanish/Nahua, is Dzandaya or Sandaya in Mixtec), or one place name corresponding to several locations (e.g. Acatlán may refer to a *pueblo* in the state of Guerrero, an *Alcaldía Mayor* and *cabecera* in the state of Puebla, a *cabecera* in Veracruz, or any of its numerous other occurrences).

In order to resolve this, we carried out the *linguistic disambiguation* of the place names. For this reason, we chose to reconstruct current toponymy as a starting point, using digital resources that were already available in GIS format, including:

I. The catalog of indigenous towns, produced by the *Instituto Nacional de Estadística Geografía e Informática* (INEGI) and the *Comisión Nacional para el Conocimiento y Uso de la Biodiversidad* (CONABIO) in 2010.

II. The geographic dataset created for the cartography of Mexico by INEGI, especially the toponym cartographic layers (scale 50k and 250k) and INEGI's general toponym locations [18].

III. The GeoNames geographical database [19].

IV. The Getty Thesaurus of Geographic Names [20].

V. The place names database from the National Geospatial-Intelligence Agency (NGA) of the United States [21].

As research on Mexican toponymy has a long tradition, we also compiled and digitised comprehensive studies in historical geographies to fill the gap between modern and 16th century toponymies. Some of the most relevant have been Gerhard [22-24], Cline [10], Moreno Toscano [25], and Tanck de Estrada *et al.* [26]. These provide indexes of towns, geographic feature names, and civil and ecclesiastical divisions as they were referred to in both 16th century documents and modern equivalents. Additionally, we have also drawn upon the valuable resources created by the 'HGIS de las Indias' project [27], which provided the location of place names in 18th century Spanish America.

After compiling these spatial references, we began with the process of *geographical disambiguation*, assigning coordinates to each and every toponym mentioned in the RGs, when possible. To do this, we designed a two-stage methodology. First, through a semi-automatic approach for detecting and merging duplicate gazetteer entries, the 16th century toponyms requiring disambiguation are matched against current toponymy and historical geographies. The result of these joins is reviewed manually to confirm or refute their accuracy. During this process, attributes are added manually to the geodatabase, such as the toponym type (city, town, mountain, river, etc.) based on a thesaurus, the coordinates confidence degree (1: definitive, 2: approximate), bibliographic references, and the relationships that place names share with other locations ('depends on', 'near', 'contains', etc.). The second stage focuses on those toponyms that cannot be disambiguated with this method, requiring a dedicated individual investigation process (Figure 2).

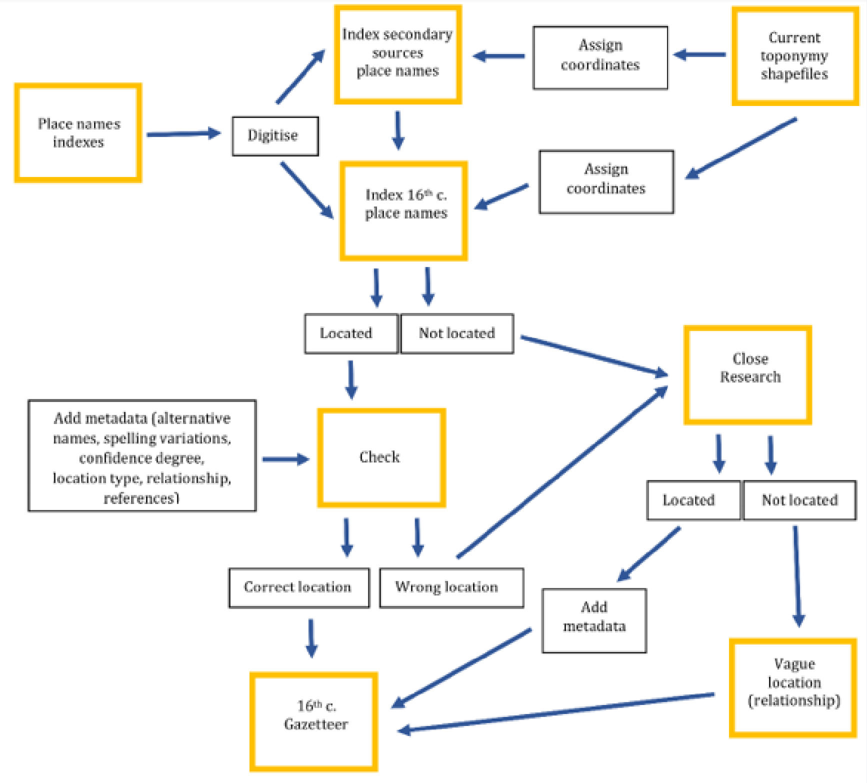

Index secondary sources place names
Assign coordinates
Current toponymy shapefiles
Place names indexes
Digitise
Assign coordinates
Index 16th c. place names
Located
Not located
Close Research
Add metadata (alternative names, spelling variations, confidence degree, location type, relationship, references)
Check
Located
Not located
Correct location
Wrong location
Add metadata
Vague location (relationship)
16th c. Gazetteer

Figure 2: Gazetteer workflow.

Many of the toponyms that could not be disambiguated in the first stage correspond to localities that no longer exist as a consequence of the colonial resettlement processes that occurred during the 16th, 17th, and 18th centuries. Others present orthographic variations that require consideration of linguistic normalisation of indigenous to Spanish orthography (e.g. Cuistlabaca, Cuestlaguaca and Coixtlahuaca). Homonyms and the complete modification of the toponyms are additional problems that need to be considered (e.g. Cintla or Zentla becoming Colonia Manuel González).

The methodology to disambiguate such problematic toponyms is based on the use of the digitised layers of the 16th century provinces of New Spain contained in the work of Peter Gerhard [23]. The GIS shapefile (polygon) layer delimiting these regions was transformed into a KML (Keyhole Markup Language) format that allows the visualisation of geographic data in a chosen 3D environment which was, in this case, the Google Earth platform. Subsequently, the modern localities belonging to each of the historical provinces were extracted through geoprocessing using CONABIO's layer. This new layer was also exported into the KML format for display in a geospatial interface. The next step was to identify the main *cabeceras* previously georeferenced, as well as historical information regarding the possible locations of problematic toponyms contained in Gerhard [23], Cline [10], Moreno Toscano [25], and numerous other geohistorical studies. In this way, it is possible to track, in most cases, the current location or the approximate coordinate that corresponds to the toponym in question. In addition, it was necessary to georeference other historical maps of the 16th century located in the General Archive of the Nation (Mexico) or in the General Archive of Indias (Seville, Spain), as well as to consult the Geonames database to assist the research process.

It is important to mention that the visualisation of layers of information in Google Earth or ArcMap is very useful given that, in some cases, corresponding places can be located a couple of kilometers outside the georeferenced polygon. This is because these polygons were obtained from the digitisation of the maps that Peter Gerhard published in 1972 [23] and do not, therefore, have precise limits. The construction of the gazetteer is an ongoing process and, thus far, we have been able to identify a total of 14,548 toponyms, including those from the RGs and secondary sources [23,25], with 73% disambiguated (Table 1).

Table 1: Geographic disambiguation process up to October 2019.

Index Publication	Total per vol.	Coord. assigned semiautom	To close research	Identified by close research	In progress	Total Disamb	% Total Disamb
Primary Sources / 16th C. RGs							
vol. 1 Guatemala Acuña	844	0	0	0	844	0	0
vol. 2 Antequera Acuña	541	379	162	0	162	379	70%
vol. 3 Antequera Acuña	700	561	139	0	139	561	80%
vol. 4 Tlaxcala Acuña	497	212	0	0	0	212	42%
vol. 5 Tlaxcala Acuña	891	401	0	0	0	401	45%
vol 6. México Acuña	794	423	371	255	116	678	85%
vol. 7 México Acuña	603	354	249	4	245	358	59%
vol. 8 México Acuña	215	157	58	0	58	157	73%
vol. 9 Michoacán Acuña	723	0	0	0	723	0	0
vol. 10 Nueva Galicia Acuña	455	0	0	0	455	0	0
RG Yucatán Garza Apéndices	166	166	0	0	0	166	100%
RG Yucatán Garza Toponym Index	729	406	323	323	0	729	100%
vol 1. Del Paso y Troncoso	929	658	271	0	271	658	70%
Secondary Sources							
Gerhard, New Spain	4,745	2,714	2,031	2,031	0	4,745	100%
Gerhard, SE Frontier	1,156	590	566	512	54	1,102	95%
Moreno Toscano	560	560	0	0	0	560	100%
Total	14,548	7,581	4,170	3,125	3,067	10,706	73%

Text mining: Geographical Text Analysis, annotation, and Machine Learning

Another important aspect of the project has been the development and advancement of text mining approaches for the exploration of historical corpora. With the textual corpus digitised, we were able to begin the annotation process for the extraction and mining of concepts of our interest. Geographical Text Analysis (GTA) is a methodology originally developed by some members of our team. GTA is in essence a combination of theory and methods from Natural Language Processing, Corpus Linguistics, and GIS that facilitates the identification of unsuspected patterns of information by focusing on the geographic components of the historical narrative in large corpora [28-34]. In six stages, this analysis consists of:

I. Pre-processing the corpus to convert it into a machine-readable format. This involves digitising and/or

Citation: Liceras-Garrido, Raquel, Favila-Vázquez, Mariana, Bellamy, Katherine, Murrieta-Flores, Patricia, Jiménez-Badillo, Diego, et al. Digital Approaches to Historical Archaeology: Exploring the Geographies of 16th Century New Spain. Open Access J Arch & Anthropol. 2(1): 2019. OAJAA.MS.ID.000526. DOI: 10.33552/OAJAA.2019.02.000526.

transcribing the data, and cleaning any noise caused by the digitisation process.

II. Semi or automatic identification of toponyms and other important information in the corpus, including all named entities (i.e. people, institutions, locations, etc.). To achieve this, a Natural Language Processing (NLP) task called Named Entity Recognition (NER) needs to be executed to extract proper names (i.e. toponyms, dates, institutions and people's names). NER approaches can range in complexity, from methods based on rules and/or dictionaries, to methods based on statistical models inferred through Machine Learning, or hybrid approaches combining these different methodologies.

III. Disambiguation of place names extracted in the previous stage. This involves resolving any linguistic and geographic uncertainties surrounding the place names in order to assign coordinates to the majority of sites mentioned in the corpus. The use of existing directories of place names, or the construction of a gazetteer connecting place names to geographic footprints, is essential to both this and the previous stage.

IV. Further annotation of the corpus, depending on predicted research questions and analyses. The annotations may refer to "part of speech" syntactic categories (i.e. noun, pronoun, adjective, determiner, verb, adverb, preposition, conjunction, and interjection), or they may refer to ad-hoc categories of terms (e.g. tagging information that refers to "deities"), created according to the research questions and managed through a tailored schema. This schema is later used to extract and cross-link information in each document or across the whole corpus. Following this stage, it is possible to carry out several data mining exercises and analyses. Two quick examples are:

a. extracting subsets of toponyms contained in the corpus.

b. creating thematic maps based on information from the corpus, linked to place names and geographic footprints.

V. Analysis of the data through an adapted form of Collocation Analysis (CA), a useful technique from Corpus Linguistics. CA allows extracting the most statistically significant words that are found next to a term or keyword of interest (the "collocates", in CL jargon). The Geographical Text Analysis variant of the technique corresponds to the analysis of "Geographic Collocations". This facilitates the recognition and retrieval, across the corpus, of all toponyms that collocate, or are close to, a keyword of interest according to a predefined proximity threshold (usually defined as the maximum number of words between the search-term and the collocate). In this manner, if a researcher is interested in investigating economic aspects related to agricultural products, s/he could locate, for example, all place names associated with the word "maize".

VI. The datasets obtained in previous stages can be imported into a geodatabase for analysing the information in a GIS environment.

VII. This process facilitates the aggregation and analysis of information at a macro scale, which, in computer-assisted literary studies, is called "distant reading". At the same time, the use of Geographic Collocation and the concept of "keywords-in-context" permit a more detailed granular comprehension of the corpus, by identifying those parts that deserve a "close reading". This flexibility has been demonstrated in numerous research papers employing GTA [4,30,35,36] and opens an unprecedented range of possibilities in terms of identifying information patterns and hypothesis-testing across large corpora.

Although GTA offers considerable advantages in terms of answering historical questions in a way that would be impossible otherwise, it is quite limited by the capabilities of standard NER methodologies, as these only deal with recognising proper names and dates. Therefore, in order to expand the entities recognised by the computer, DECM is using a hybrid NLP approach that combines Machine Learning (ML) with dictionaries and rules. This work is being developed in collaboration with our industry partner tagtog, a company that has created a collaborative online text annotation platform.

When working with ML, challenges begin with collecting training data because, first, labeled datasets in this domain (i.e. documents written in the 16th century "classical" Spanish, mixed with phrases in different indigenous languages of the Americas) are scarce or non-existent; and, second, the increasing complexity and changing nature of linguistic nuances require knowledge and consistent verification from subject-matter experts. In the context of modern NLP approaches, this knowledge comes in the form of text annotations. Consequently, the first step was to created a tailored ontology that encompassed the complexity of RGs, resulting in twenty-one entities or categories. These entities were defined and linked to DBpedia, a semantic standard, to ensure the interoperability of our dataset. Additionally, most entities contain labels, also defined and linked to DBpedia to enrich the semantic extraction process (see Table 2).

Once the ontology is defined, the main steps are to:

I. annotate a data sample through tagtog (Figure 3), augmenting the historical documents with information on the spans of text corresponding to entities of interest;

II. carry out an Inter-Annotator Agreement assessment for data sample quality, which measures the level of discrepancies between different annotators, enabling us to spot biases and ambiguous cases, and to evaluate the frequency of different class occurrences (to check for unbalanced class usage or oversampled data);

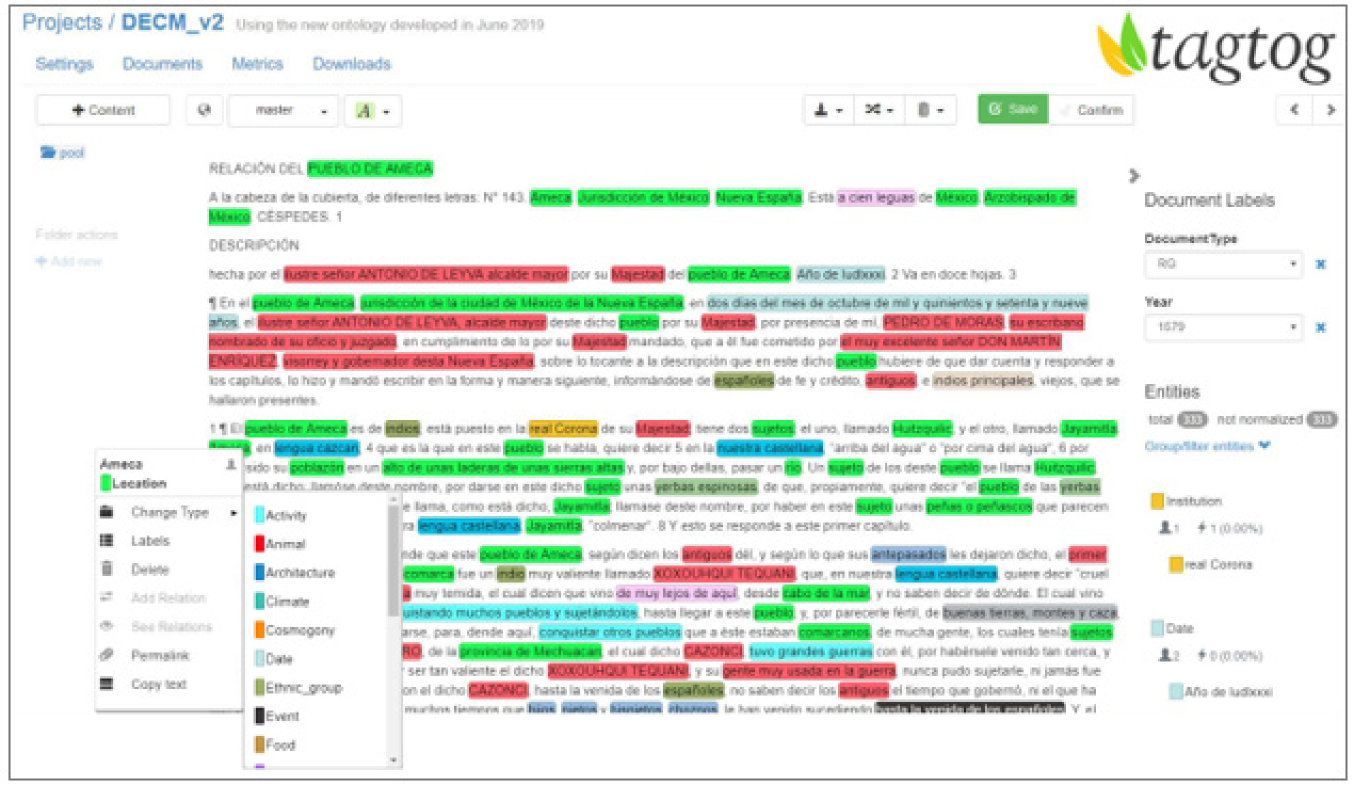
III. annotate a second data sample, facilitating the subsequent training of NER models to automatically recognise entities, and also the creation of additional resources (e.g., dictionaries and rules) for the fine-tuned labels;

IV. train the machine; and finally

V. extract and mine the text.

Table 2: Annotation taxonomy: entities, DBpedia definitions and labels per entity.

Entity	Ontological definition	Labels
Person	http://dbpedia.org/page/Person	female/male, title, profession
Date	http://dbpedia.org/ontology/date	
Institution	http://dbpedia.org/page/Institution	civil, ecclesiastical
Location	http://dbpedia.org/ontology/location	settlement type, generic location, geographic feature type, toponym, address, imaginary, ecclesiastical jurisdiction, civil jurisdiction
Activity	http://dbpedia.org/ontology/activity	agriculture, warfare, economy, mining, domestic, female/male
Animal	http://dbpedia.org/ontology/animal	insect, mammal, reptile, bird, amphibian, aquatic, domesticated
Plant	http://dbpedia.org/page/Plant	
Food	http://dbpedia.org/page/Food	
Natural Resource	http://dbpedia.org/page/Natural_resource	
Cultural artefact	http://dbpedia.org/page/Cultural_artifact	house goods, commodities, clothing, weapon, tool
Architecture	http://dbpedia.org/page/Architecture	religious, civil, domestic
Cosmogony	http://dbpedia.org/page/Cosmogony	ritual, festivity, activity, deities, saints, object
Health	http://dbpedia.org/page/Health	disease, remedy
Route of Transportation	http://dbpedia.org/ontology/RouteOfTransportation	terrestrial, aquatic, route direction, distance
Kinship	http://dbpedia.org/page/Kinship	
Climate	http://dbpedia.org/page/Climate	
Ethnic Group	http://dbpedia.org/page/Ethnic_group	
Social Class	http://dbpedia.org/page/Social_class	
Language	http://dbpedia.org/page/Language	
Event	http://dbpedia.org/page/Event	historical, disasters
Measurement	http://dbpedia.org/page/Measurement	value, tribute, weight, population



DocumentType: RG

Year: 1579

total 333 not normalized 333

Institution 1 1 (0.00%)

Date 2 0 (0.00%)

RELACIÓN DEL PUEBLO DE AMECA

A la cabeza de la cubierta, de diferentes letras: N° 143. Ameca. Jurisdicción de México. Nueva España. Está a cien leguas de México. Arzobispado de México. CÉSPEDES. 1

DESCRIPCIÓN

hecha por el ilustre señor ANTONIO DE LEYVA alcalde mayor por su Majestad del pueblo de Ameca. Año de ludxxxi. 2 Va en doce hojas. 3

¶ En el pueblo de Ameca, jurisdicción de la ciudad de México de la Nueva España, en dos días del mes de octubre de mil y quinientos y setenta y nueve años, el ilustre señor ANTONIO DE LEYVA, alcalde mayor deste dicho pueblo por su Majestad, por presencia de mí, PEDRO DE MORAS, su escribano nombrado de su oficio y juzgado, en cumplimiento de lo por su Majestad mandado, que a él fue cometido por el muy excelente señor DON MARTÍN ENRÍQUEZ, visorrey y gobernador desta Nueva España, sobre lo tocante a la descripción que en este dicho pueblo hubiere de que dar cuenta y responder a los capítulos, lo hizo y mandó escribir en la forma y manera siguiente, informándose de españoles de fe y crédito, antiguos, e indios principales, viejos, que se hallaron presentes.

1 ¶ El pueblo de Ameca es de indios, está puesto en la real Corona de su Majestad, tiene dos sujetos, el uno, llamado Huitzquilic, y el otro, llamado Jayamitla.
Ameca, en lengua cazcan, 4 que es la que en este pueblo se habla, quiere decir 5 en la nuestra castellana, "arriba del agua" o "por cima del agua", 6 por
sido su población en un alto de unas laderas de unas sierras altas y, por bajo dellas, pasar un río. Un sujeto de los deste pueblo se llama Huitzquilic,
está dicho: llamóse deste nombre, por darse en este dicho sujeto unas yerbas espinosas, de que, propiamente, quiere decir "el pueblo de las yerbas
e llama, como está dicho, Jayamitla: llamase deste nombre, por haber en este sujeto unas peñas o peñascos que parecen
ra lengua castellana, Jayamitla, "colmenar". 8 Y esto se responde a este primer capítulo.

nde que este pueblo de Ameca, según dicen los antiguos dél, y según lo que sus antepasados les dejaron dicho, el primer
comarca fue un indio muy valiente llamado XOXOUHQUI TEQUANI, que, en nuestra lengua castellana, quiere decir "cruel
a muy temida, el cual dicen que vino de muy lejos de aquí, desde cabo de la mar, y no saben decir de dónde. El cual vino
uistando muchos pueblos y sujetándolos, hasta llegar a este pueblo, y, por parecerle fértil, de buenas tierras, montes y caza,
arse, para, dende aquí, conquistar otros pueblos que a éste estaban comarcanos, de mucha gente, los cuales tenía sujetos
RO, de la provincia de Mechuacan, el cual dicho CAZONCI, tuvo grandes guerras con él, por habérsele venido tan cerca, y
ser tan valiente el dicho XOXOUHQUI TEQUANI, y su gente muy usada en la guerra, nunca pudo sujetarle, ni jamás fue
on el dicho CAZONCI, hasta la venida de los españoles, no saben decir los antiguos el tiempo que gobernó, ni el que ha
muchos tiempos que hijos, nietos y bisnietos, chozmos le han venido sucediendo hasta la venida de los españoles. Y el

Figure 3: Annotation process. Created using tagtog: https://www.tagtog.net/

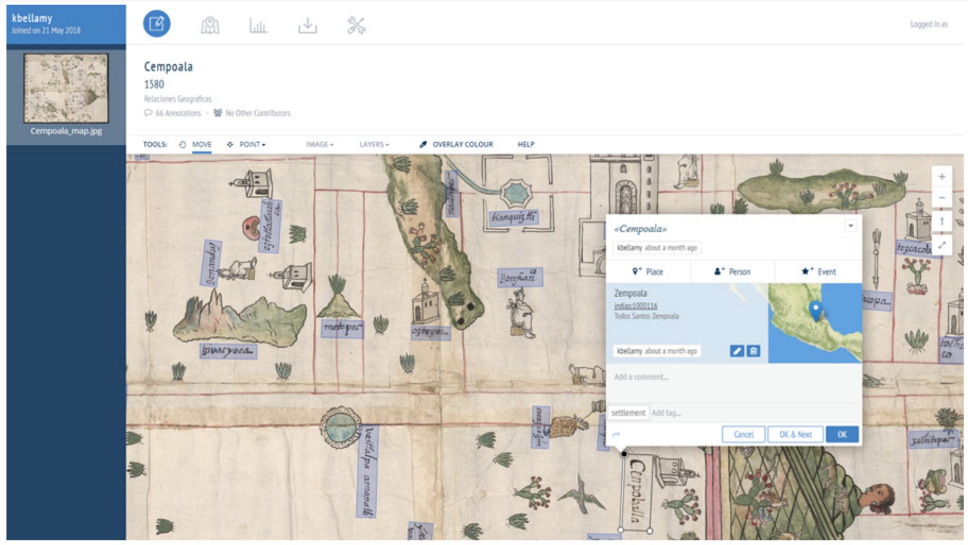
Currently, we are at the third stage, producing the annotation sample that will train the model. The tagtog annotation platform offers the possibility to upload dictionaries (pre-existing lists of toponyms, flora, political positions, institutions, etc.) to facilitate the classification of text according to the labels and entities in the ontology, facilitating the annotation process and enabling the human annotators to better focus on the problematic nouns. Simultaneous to advancing on this third stage, the data resulting from initial annotation efforts has also been used to support the training of a NER model based on neural networks, focusing on person and location names (i.e. we used the initial data to validate our ideas regarding the development of Stage 4, combining the in-domain data with other pre-existing Spanish corpora containing annotations for named entities and parts-of-speech tags). Initial cross-validation tests showed that this NER model can produce high quality results (i.e. an F1 score of approximately 85% in terms of correctly identifying the spans corresponding to persons or locations), although further improvements are still planned (e.g. we plan to use parts-of-speech tags to latter help in the recognition of entities corresponding to common nouns). Once we have mined and extracted the information relevant to us from the text, we will proceed to connect this dataset with the gazetteer, which will enable us to work with this data in a GIS environment and to carry out spatial and statistical analyses.

Working with digital methods to analyse images: the maps (*pinturas*) of the RGs

One of the questions in the RGs requested that the recipient create "a plan in colour of the streets, plazas, and other significant features such as monasteries, as well as can be sketched easily on paper, indicating which part of the town faces south or north" [12]. Whilst some recipients chose to ignore this request, a considerable number ensured the creation of these maps and, as a result, a corpus of 78 maps has survived for the Mexico area. These maps are a unique reflection on 16th century settlements in Mexico, drawn using a combination of indigenous and European techniques and ideas. This interplay of indigenous and European voices is a key part of these maps' significance, offering a unique insight into multiple perceptions of space and place during this crucial period in Mexico's history. The importance of these maps is clear, and they have been the subject of numerous studies, which have provided invaluable contributions to our understandings of indigenous and European perceptions of space and place in 16th century Mexico [15].

The maps of the RGs contain a great variety of information, both textual and pictographic, which offer invaluable insight into the historical and geographical contexts in which these maps were produced. This information includes proper names in the form of both traditions, logographic Mesoamerican toponyms and people's names, and European alphabetic glosses. Recent advances in digitisation techniques have allowed new ways of interacting with, and understanding, these valuable sources.

Digital annotation of maps is one method which offers a promising way of analysing a corpus, which is not heavily text-based, but may feature text alongside pictographic depictions of space and place. We have recently begun the process of annotating the maps of the Geographic Reports using Recogito, which has enabled us to investigate these maps in new and innovative ways. The tool developed by Pelagios Commons [38] allows us to identify, record and export places and other information depicted in the maps through a Linked Open Data model. With the support of a Pelagios Commons Resource Development Grant, we were able to develop the project *Subaltern Recogito: Annotating the sixteenth-century maps of the Geographic Reports of New Spain.* In collaboration with our colleagues in the LLILAS Benson Latin American Studies and Collections at The University of Texas at Austin, the National School of Anthropology and History (ENAH), The National Autonomous University of Mexico (UNAM), the National Institute of Anthropology and History (INAH), and the University of Lisbon, we delivered an online workshop and trained 27 scholars from UNAM and ENAH to complete the annotation of the full corpus of the maps.



Cempoala
1580
Relaciones Geograficas
66 Annotations · No Other Contributors

«Cempoala»
kbellamy about a month ago
Zempoala
indias:1000116
Todos Santos Zempoala
kbellamy about a month ago
settlement

Figure 4: Close-up view of some of the Cempoala map annotations, including alphabetic place-names (e.g. Metepec) and names of people (e.g. Don Francisco). Created using Recogito: https://recogito.pelagios.org/.

Citation: Liceras-Garrido, Raquel, Favila-Vázquez, Mariana, Bellamy, Katherine, Murrieta-Flores, Patricia, Jiménez-Badillo, Diego, et al. Digital Approaches to Historical Archaeology: Exploring the Geographies of 16th Century New Spain. Open Access J Arch & Anthropol. 2(1): 2019. OAJAA.MS.ID.000526. DOI: 10.33552/OAJAA.2019.02.000526.

The annotation of both elements, this is to say, the logographic toponyms, and names alongside alphabetic descriptions and place names (Figure 4), is enabling us to better understand the different ways in which Mesoamerican indigenous spatial knowledge and portrayals changed over time, and the processes through which these became 'subaltern' to European thinking. When possible, each toponym was assigned to its corresponding modern geospatial coordinates in the Recogito platform (Figure 5). The collectively annotated information produced from this workshop will be available at the end of the project to any scholar with a CC-BY licence, and it will be also deposited at the Nettie Lee Benson Latin American Collection at the University of Texas as an enhanced dataset and published in their data repository dedicated to the RGs: https://dataverse.tdl.org/dataverse/relaciones_geograficas. It is our aim that the resulting dataset can also support the development of automated approaches (e.g., computer vision methods leveraging advances in machine learning) for annotating the considered information elements on other similar maps.

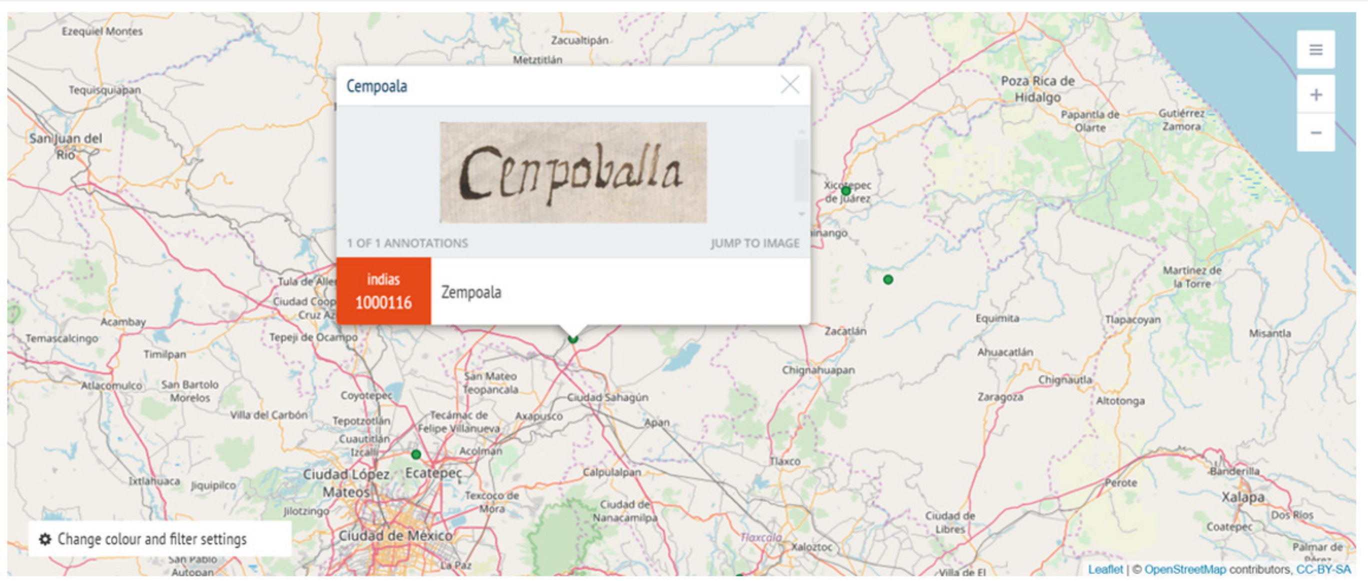
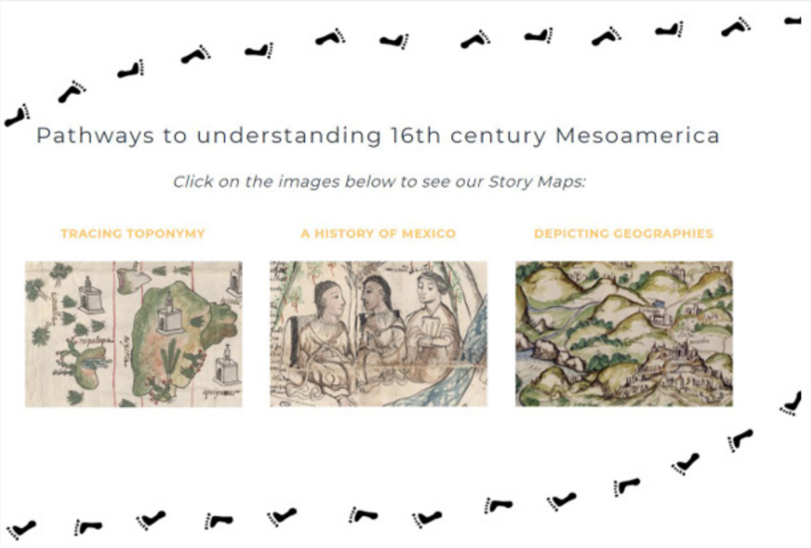


Figure 5: Close-up view of the annotated location of the alphabetic place-name of Cempoala on a modern map. Created using Recogito: https://recogito.pelagios.org/.

Mesoamerican Pathways: a public engagement online resource

DECM is also committed to the idea that digital public engagement is an effective way to communicate research results directly to non-specialists, and to reach a wide range of public. With this in mind, and using the resources already created for the gazetteer, we have also developed the impact project 'Pathways to understanding 16th century Mesoamerican geographies'. This resulted in the creation and publication of three ArcGIS StoryMaps with two primary objectives: first, to disseminate the content of our research targeting non-academic public and, secondly, to create an online, interactive tool for teaching.


Pathways to understanding 16th century Mesoamerica

Click on the images below to see our Story Maps:

TRACING TOPONYMY

A HISTORY OF MEXICO

DEPICTING GEOGRAPHIES

Figure 6: Mesoamerican Pathways landing page (https://www.lancaster.ac.uk/digging-ecm/portfolio-items/pathways-to-understanding-16th-century-mesoamerica/)

Story Maps are teaching/dissemination devices that enable the narration of any kind of story based around geographic information, particularly maps. The easy identification of locations, main characters and episodes, alongside its interactive, visual nature make this tool the perfect medium to present historical narratives. ESRI's ArcGIS StoryMaps offers a dynamic storytelling environment to combine textual narratives, current and historical maps, and images. Focused on the Mesoamerican Postclassical and Colonial period of Central Mexico, covering a period from the 14th to the mid-16th century, the StoryMaps are divided into three themes: History, Toponymy and Geography (Figure 6).

Mesoamerican pathways | A history of Mexico: The first and largest of the StoryMaps explores the history of the Mexica people, beginning with their journey to the foundation of Tenochtitlan in 1325, which would become (alongside its neighbour city to the north, Tlatelolco), the heart of the Triple Alliance. Following this, it shows how the Mexica began to expand, featuring the lists of conquered settlements as recorded in the Codex Mendoza. This leads up to the arrival of the Spanish, and the ultimate meeting of Moctezuma II and Hernán Cortés in 1519. It then proceeds to describe how Cortés, with considerable assistance from his indigenous allies, conquered Tenochtitlan. This Story Map concludes with a look at the beginning of the colonial era, exploring how the Spanish began to impose their own institutions across 'New Spain', with varying success due to the continuing influence of indigenous institutions across Mesoamerica.

Mesoamerican pathways | Tracing toponymy: This Story Map explores the nature of historic place-names across what is currently Mexico, introducing the importance of place-names and language as a tool of colonisation and empire. It explores how this tool was used not only by the Spanish, but also by the Mexica and the Triple Alliance (not to mention other indigenous groups), as part of their systematic colonisation of conquered settlements and people. The Story Map goes on to explore how indigenous place-names continued to be used, despite the processes of colonisation at the hands of both the Triple Alliance and the Spanish. In addition, it explores the meaning of Nahua toponymy in particular – demonstrating the use of suffixes such as -tepec (which means 'inhabited place') and showing the distribution of some of these examples. Following this are some case studies of individual place-names, explaining their meaning and how they have been depicted in the historical record. The Story Map concludes by giving a brief overview of colonial naming, and how indigenous influences have continued.

Mesoamerican Pathways | Depicting Geographies: The third and final StoryMap discusses depictions of geographic space and place. This starts with an explanation of why this is an important discussion, with particular reference to, and problematisation of, the use of Geographic Information Systems for representing historical geographies. Following this, it introduces the idea of representations of space that may be unfamiliar to the modern reader and explores the various types of pre-Hispanic Nahuatl documents, including those which represented geographies. The StoryMap gives an introduction to the state of Spanish cartography in the 16th century, before going on to discuss how the Spanish and Nahua traditions of depicting geography began to merge during the conquest of Mexico. There is considerable evidence of this merging of traditions throughout the historical record, which is explained in the StoryMap alongside two specific examples.

Concluding Remarks

Given the complexity of historical data and the technical development in Digital Humanities thus far, DECM is addressing several challenges in all the disciplines involved. Starting with the analysis of text, the fact that the corpus is multilingual, written in the 16th century Spanish, and peppered with words in 69 indigenous languages, is pushing Corpus Linguistics (CL), Natural Language Processing (NLP) and Machine Learning (ML) beyond their current limits, which are based upon, and restricted by, tests which use predominantly modern texts written in English.

Additionally, since the RGs are responses to a questionnaire, the topics discussed are multiple and diverse, and consequently the annotation ontology contains a large number of entities and labels. In the same vein, the majority of our annotations are nouns. While the identification of proper nouns is well tested in NER, the automated classification of words beyond these is more problematic. Therefore, to improve the results when the computer automatically assigns entities and labels, we are using a hybrid approach that combines the use of rules and dictionaries, with statistical models with parameters inferred through Machine Learning (ML). Research in Mexico has produced a large number of dictionaries focused on a variety of topics, including toponyms, flora, political positions, institutions, etc. The tagtog annotation platform offers the possibility to upload dictionaries to labels and entities to facilitate the classification of text, enabling the sorting of diverse topics and problematic nouns. At present, ML is being considered to support the recognition of entities corresponding to proper nouns (e.g. persons and locations) but, beyond this, we also plan to experiment with Machine Learning approaches for more ambiguous entities and labels where classification is difficult using dictionaries, such as cosmogony, activities, cultural artefacts, or events.

Working with 16th century sources means constant spelling variations and ambiguity. There is a lack of standardisation in the spelling of people's names and places, in addition to the forms of naming ethnic groups, social classes, regional titles, plants, animals or local artifacts, amongst others. This has the potential to hamper the expandability of the gazetteer (as capacity for these alternative or additional words must be built-in to the database), as well as complicating the text mining and Machine Learning processes. The challenges posed by semantic and geographic ambiguity are present in all project stages and research areas. In the case of a place name corresponding to several locations, the geographic disambiguation process necessarily includes a stage for close, dedicated research. When training the computer for the automatic identification of information using Machine Learning, one of the challenges that we are experiencing is that the same word can have several meanings. For instance, the word '*vino*' could be a conjugation of the verb 'to come' (s/he/it comes) or the noun 'wine', depending on the context. The creation of a wide and diverse set of training data, or the use of NLP techniques to recognise the 'parts of speech', can both contribute to the differentiation and classification of particular word types.

Regarding the spatial depictions of the RGs maps, we have been able to assign geographical coordinates to many of the logographic toponyms and place names written in European glosses that appear in these *pinturas*, using tools including Recogito. The geolocation of these places is, however, a minute aspect of the information contained within these maps. There are other elements such as orographic features, pathways, religious symbols and persons that together construct a semantic riddle that contributes to the interpretation of the spatial logic of the map [39]. The possibility to include this data may reveal new patterns in the construction of space, offering insight into indigenous conceptions of the surrounding world. We also aim to develop new kinds of research questions that complement and build upon the perspectives of established theories and interpretations of indigenous history.

With the advancements of the Digital Humanities, we hope that initiatives such as the DECM project encourage historical archaeologists to consider the different types of spatial data that can enrich research, developing its potential in the social sciences and humanities by establishing new research methodologies. In our commitment to FAIR data, at the end of the project in January 2021, all datasets will be published openly in Lancaster University, Lisbon University and INAH repositories.

Acknowledgement

This research is supported by a grant from the ESRC-UK, CONACyT-Mexico, and FCT-Portugal under the Transatlantic Platform of Social Sciences and the Humanities (Digging into Data), grant number ES/R003890/1: 'Digging into Early Colonial Mexico: A large scale computational analysis of 16th century historical sources' (DECM). The Portuguese researchers working in the project have also received additional support from FCT-Portugal, through the INESC-ID multi-annual funding from the PIDDAC programme (UID/CEC/50021/2019). Pelagios Commons funded the 'Subaltern Recogito: Annotating the sixteenth-century maps of the Geographic Reports of New Spain' with a small grant. The creation of the public engagement resource was funded by the Lancaster University's History Department under the project 'Pathways to understanding 16th century geographies'. We would also like to thank our DECM team for their support.

Conflict of Interest

There is no conflict of interests.

References

1. Jeremy Huggett (2012) Core or Periphery? Digital Humanities from an Archaeological Perspective. Historical Social Research 37(3): 86-105.
2. David M Berry (2012) Understanding Digital Humanities. Palgrave Macmillan, London.
3. David Cooper, Christopher Donaldson, Patricia Murrieta-Flores(Eds.) (2016) Literary Mapping in the Digital Age. Digital Research in the Arts and Humanities. Ashgate, Farnham, Surrey, England Burlington, VT.
4. Patricia Murrieta-Flores, Christopher Donaldson, Ian Gregory (2017) GIS and Literary History: Advancing Digital Humanities Research Through the Spatial Analysis of Eighteen-Century Travel Writing. Digital Humanities Quarterly 11(1).
5. Raúl Gómez (2004) Los Virreinatos Americanos. Dastin Export, Madrid.
6. Howard F Cline (1972) Ethnohistorical Regions of Middle America. In: Howard F Cline (Edt.), Handbook of Middle American Indians 12: Guide to Ethnohistorical Sources, Part One. University of Texas Press, Austin, pp: 166-182.
7. Patricia Murrieta-Flores, Diego Jiménez Badillo, Bruno Martins, Ian Gregory, Mariana Favila Vázquez, Raquel Liceras-Garrido, Katherine Bellamy (2019 In press) Exploring Early Colonial Mexico. An analysis of historical texts through the application of computational linguistics and Geographic Information Systems. In: Diego Jiménez Badillo (Ed.), Computational Methods and Digital Techniques to Analyze and Disseminate Cultural Heritage. National Institute of Anthropology and History, Mexico City.
8. Howard F Cline (1964) The Relaciones Geográficas of the Spanish Indies, 1577-1586. The Hispanic American Historical Review 44 (3): 341-374.
9. Howard F Cline (1972) Introductory Notes on Territorial Divisions of Middle America. In: Howard F Cline (Ed.), Handbook of Middle American Indians, Volume 12: Guide to Ethnohistorical Sources, Part One. University of Texas Press, Austin, pp. 17-62.
10. Howard F Cline (1972) The Relaciones Geográficas of the Spanish Indies, 1577-1648. In: Howard F. Cline (Edt.), Handbook of Middle American Indians, Volume 12: Guide to Ethnohistorical Sources, Part One. University of Texas Press, Austin, pp. 183-242.
11. Mercedes de la Garza (1983) Relaciones HistóricoGeográficas de La Gobernación de Yucatán: (Mérida, Valladolid y Tabasco) 1,2 (1St edn), Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas, Centro de Estudios Mayas, México, North America.
12. René Acuña (1982-1988) Relaciones Geográficas Del Siglo XVI, 1-10(1St edn), Universidad Nacional Autónoma de México, Instituto de Investigaciones Antropológicas, México, North America.
13. Francisco del Paso y Troncoso (1905) Papeles de La Nueva España Publicados de Orden y Con Fondos Del Gobierno Mexicano, Tipográfico Sucesores de Rivadeneyra, Madrid.
14. Carmen Manso Porto (2012) Los mapas de las Relaciones Geográficas de Indias de la Real Academia de la Historia. Revista de Estudios Colombinos 8: 23-52.
15. Barbara Mundy (1996) The Mapping of New Spain: Indigenous Cartography and the Maps of the Relaciones Geográficas. University of Chicago Press, Chicago, Illinois.
16. Linda L Hill, James Frew, Qi Zheng (1999) Geographic Names: The Implementation of a Gazetteer in a Georeferenced Digital Library. D Lib Magazine 5 (1).
17. Linda L. Hill (2000) Core Elements of Digital Gazetteers: Place names, Categories, and Footprints. In: Borbinha J, Baker T (Eds.), Research and Advanced Technology for Digital Libraries. ECDL 2000. Lecture Notes in Computer Science, vol. 1923. Springer, Berlin, Heidelberg, pp. 280-290.
18. https://www.inegi.org.mx/
19. https://www.geonames.org/
20. http://www.getty.edu/research/tools/vocabularies/tgn/
21. http://geonames.nga.mil/namesgaz/
22. Peter Gerhard (1972) Colonial New Spain, 1519-1786: Historical Notes on the Evolution of Minor Political Jurisdictions. In: Howard F. Cline (Editor) Guide to Ethnohistorical Sources, Part One. Handbook of Middle American Indians 12. University of Texas Press, Austin, pp. 63-137.

23. Peter Gerhard (1972) A Guide to the Historical Geography of New Spain. Cambridge University Press, Cambridge.

24. Peter Gerhard (1991) La Frontera Sureste de La Nueva España. IIH-UNAM, México, North America.

25. Alejandra Moreno Toscano (1968) Geografía económica de México (siglo XVI). Colegio de México, México, North America.

26. Dorothy Tanck de Estrada, José Antonio Álvarez Lobato, Jorge Luis Miranda (2005) Atlas Ilustrado de Pueblos de Indios de La Nueva España, 1800(4) El Colegio de México, México, North America.

27. https://www.hgis-indias.net/cmv-app-master/viewer/

28. David Cooper, Ian N Gregory (2011) Mapping the English Lake District: A Literary GIS: Mapping the English Lake District. Transactions of the Institute of British Geographers 36 (1): 89-108.

29. Patricia Murrieta-Flores, Alistair Baron, Ian Gregory, Andrew Hardie, Paul Rayson (2015) Automatically Analyzing Large Texts in a GIS Environment: The Registrar General's Reports and Cholera in the 19th Century. Transactions in GIS 19(2): 296-320.

30. Ian N Gregory, Christopher Donaldson, Patricia Murrieta-Flores, Paul Rayson (2015) Geoparsing GIS, and Textual Analysis: Current Developments in Spatial Humanities Research. International Journal of Humanities and Arts Computing 9(1): 1-14.

31. Catherine Porter, Paul Atkinson, Ian Gregory (2015) Geographical Text Analysis: A New Approach to Understanding 19th Century Mortality. Health & Place 36: 25-34.

32. Bruno Martins, Patricia Murrieta-Flore (2017) Geo Humanities 2017 Workshop Report. Newsletter SIGSPATIAL Special 9 (3): 22-23.

33. Patricia Murrieta-Flores, Ian N Gregory (2017) Cruzando Fronteras En Humanidades Digitales: Análisis Geográfico de Textos de Interés Histórico y Arqueológico Con Sistemas de Información Geográfica. In: Diego Jiménez Badillo (Edt.), Arqueología Computacional. Nuevos Enfoques Para El Análisis y La Difusión Del Patrimonio Cultural. INAH, México. pp. 199-212.

34. Ian N Gregory, Christopher E. Donaldson, Andrew Hardie, Paul E. Rayson (2018) Modelling Space in Historical Texts. In: Julia Flanders and Fotis Jannidis (Edt.), The Shape of Data in Digital Humanities: Modeling Texts and Text-Based Resources. Routledge, London, pp. 133-149.

35. Joanna E, Taylor, Ian N Gregory, Christopher E, Donaldson (2017) Combining Close and Distant Reading: A Multiscalar Analysis of the English Lake District's Historical Soundscape. International Journal of Humanities and Arts Computing 12(2): 163-182.

36. Laura L Paterson, Ian N Gregory (2018) Representations of Poverty and Place: Using Geographical Text Analysis *to* Understand Discourse. Palgrave, London.

37. http://www.tagtog.net

38. http://commons.pelagios.org/

39. Patricia Murrieta-Flores, Mariana Favila Vázquez, Aban Flores Morán (2019) Spatial Humanities 3.0: Qualitative Spatial Representation and Semantic Triples as New Means of Exploration of Complex Indigenous Spatial Representations in Sixteenth Century Early Colonial Mexican Maps. International Journal of Humanities and Arts Computing 13(1-2): 53-68.