En *Digital Humanities in Medieval and Early Modern Spanish Texts*. (España):
Routledge.

The Writing and the Territory: Argentina Revisited in Digital Scholarly Editions.

del Rio Riande, Gimena.

Roberto José González Zalacain, profesor titular de Historia Medieval de la Universidad de La Laguna, como investigador principal del *Proyecto Cordican: transcripción y edición digital de textos para el conocimiento de la aportación de Canarias a la historia Atlántica en la primera globalización*, financiado por la Agencia Canaria de Investigación, Innovación y Sociedad de la Información con cofinanciación del Programa Operativo FEDER Canarias 2014-2020, para el periodo 2020-2022,

Hace constar

como coordinador, junto con Gael Vaamonde dos Santos, profesor ayudante doctor de Filología Española de la Universidad de Granada, del libro *Digital Humanities in Medieval and Early Modern Spanish Texts*, resultado de los trabajos del citado proyecto y aceptado para su publicación por la editorial Routledge, que en el mencionado libro figura la contribución de la doctora Gimena del Río Riande titulada "The Writing and the Territory: Argentina Revisited in Digital Scholarly Editions".

Y para que conste, a los efectos oportunos, lo firma, en San Cristóbal de La Laguna, el día consignado en la firma electrónica

The Writing and the Territory: Argentina Revisited in Digital Scholarly Editions

Gimena del Rio Riande (CONICET)

1. Narratives of space, narrative space, digital space

Investigating close and distant reading approaches to the narratives of the Argentinian space between the 16th and the 18th century entails building on previous work in various areas of research: Literary Studies, History, Philology, Literary Theory Geography, and Digital Humanities, among others. For instance, from a Literary Studies perspective, Morillas Ventura (2008, 2009) studied the discursive strategies of the so-called foundational texts of the Southern territory, from the Viceroyalty of the Río de la Plata up to the Modern Argentina (1870-1914). According to this scholar, the turn from the genre "report" to the "story" implies the inclusion of more subjective content, which results in an abundance of particular tales, legends and myths over general histories, especially throughout the 17th century. Cicerchia (2000) explains that the texts describing travels in the Argentinian region in the course of the 17th and 18th century follow the discursive strategies of the Renaissance, more associated with adventure and humanistic mysticism, although detailed information, critical observation, and a referential system of sources are also emphasized. Funes (2010) asks himself about the type of relationship that exists between the medieval European and the *indiana* historiography. Funes reviews the ways in which the narrative voice of colonial chronicles is nurtured by a network of discursive strategies which had already been used for almost three centuries in European memoirs, chivalric biographies, personal chronicles, and travelogues. Ehrlicher (2015) analyzes the travelbook of Columbus, finding that the obvious difference between the expected (Japan) and the observed (Cuba) leads to an epistemological crisis resulting in the flexibilization of the interpretation of observed phenomena but still with the aim to integrate the new experiences into the fixed schemes of what was previously known. Añon and Battock (2013) understand the colonial historical narrative as a possible future or even a utopia: a textual meeting space for different traditions and a new *locus* of enunciation. Finally, El Jaber (2011, 2016) focuses on the conception of the colonial American space not as a simple scenario where stories take place, but as a part of social practices which generate a territory in print and in drawing.

The works of the aforementioned scholars can also be related to the approaches of Literary Geography, a multidisciplinary field of study which emerged in the 1980's as part of the impact of the *spatial turn* in the Arts and Humanities and the *cultural turn* of the Human Geography (Alexander 2015: 1-2; Piatti et al. 2008; Moretti 1999). Both of these disciplines also allowed the emergence of Literary Cartography (Piatti et al. 2009), which highlights the scenery of literary works, understanding landscape as something to be read even beyond genres like travelogues, chronicles, memoirs, etc. (Bishell 2012: 151; Barth et al. 2017). However, the most relevant change in Literary Cartography in the last years took place inside the Digital Humanities (DH) field, thanks to Geographical Information Systems (GIS). This approach resulted in the rise of the so-called Geohumanities (Richardson et al. 2011: 3), which contribute to the standardization of the geographical analysis of literary texts through digital tools like geoparsers and digital geo-historical dictionaries or *gazetteers*. Through the retrieval and visualization of the geographic information in the texts, different aspects of the connections between space and literature are made explicit, approximating, in this ways, to Moretti's concept of *distant reading* as the way in which machines "read" (Ramsay 2011) or help humans reading patterns and data inside big corpus (Moretti 2005), something that Jockers (2013) has named as macroanalysis.

However, the constitution of the narrative space is usually defined by more complex causes than explicit mentions of toponyms in databases. Fos instance, Murietta-Flores and Howell (2017a, 2017b) advocate for a holistic analysis of space and place for DH projects, arguing that the aforementioned approaches usually deal with geographical places and spaces that can be related to geographic coordinates or *mappable*, while Literary Geography also deals with places that are vague or imaginary. They observe that not only literary texts, but also historical documents often deal with a combination of vague and located places. For an analysis of French medieval romances, they design a methodology combining Corpus Linguistics and spatial technologies to identify, extract, and investigate all these types of places and spaces. In the analysis of the narratives of the Argentinian space in chronicles and travelogues, the issues about *mappability* versus vagueness also needs to be taken into account.

Linguistic annotations with Natural Language Processing techniques (NLP), especially through Named Entity Recognition (NER), have been extensively used in DH for the analysis of literary and historical texts, and have revealed themselves as a fundamental tool to automatically extract mentions of places. Another technique that has not been too much operationalized yet for this purpose, is Topic Modeling, a quantitative method based on assumptions made in distributional semantics and aimed at the detection of hidden semantic structures in big text collections. It is an unsupervised and undeterministic method, meaning that no previous definition, classification or knowledge of topics in the texts is needed and that that there is no unique solution to the problem of finding the semantic structures hidden in the texts (Blei 2003, 2012). Topic Modeling has achieved a high degree of popularity in DH projects and has been used to analyze thematic structures in a range of historical and literary texts (Blevins 2010; Rhody, 2012; Jockers, 2013; Schöch, 2017a).

When digital methods are used to investigate narratives of space and narrative space, they can be associated with different reading methods. Text encoding of texts in digital format, and semantic and geographic annotation can prepare the texts for complex computational practices. Still, these practices are in the first place related to a close reading of the texts, aimed at the thorough individual interpretation (Singer 2013, Chaturvedi et al. 2012). Despite this, when annotations are performed automatically and without comprehensive manual verification, and when the digital texts are interpreted on quantified automatized and unsupervised analysis results and their visualization, these digital techniques can be understood as methods of distant reading, where the textual structure is dissolved in order to generate abstract views (Moretti 2005, 2013; Jockers 2013; Jänicke et al. 2015).

This chapter explains the decisions, methods, and results put into practice while working with a corpus of Argentinian written texts that describe for the first time in Spanish the vast and complex geography of the Río de la Plata (Reitano 2003, Ternavasio 2009): a Southern viceroyalty, a never ending river, and an emerging and powerful port and city (Santa María del Buen Ayre or, later, Trinidad). As mentioned, previous investigations on narratives of the Argentinian space, as well as DH close reading methods like text encoding, semantic and geographic annotation work and quantitative methods stemming from Computational Linguistics and unsupervised methods were part of an unusual workflow that also entailed working with open-source and low-tech infrastructures.

2.

In an article devoted mostly to the study of travelogues that describe the 19th century Patagonian space through European eyes, Enriqueta Morillas Ventura (2008) defines as "inaugural texts" a heterogenous corpus of early texts from the 16th and the 17th centuries that move from the depiction of the "Indias" to the one of the "New World" (pp. 155-156). Among these, she includes a group of written records that aim at drawing for the first time in Spanish the geography of the Río de la Plata: Pero Hernández's *Relación de las cosas sucedidas en el Río de la Plata*, also known as *Relación del descubrimiento del Río de la Plata, por Pero Hernández, secretario de Alvar Nuñez Cabeza de Vaca* (1545); Ulrico Schmidl's *Viaje al Río de la Plata* (1567); Martín del Barco Centera's poem *La Argentina* (1602); Ruy Díaz de Guzmán's *La Argentina manuscrita* (1612); and Acarete Du Biscay's *Relación de un viaje al Río de la Plata y desde allí por tierra a Peru…* (1672).

The book of Acarete du Biscay, like other similar chronicles of the time (e.g. Pedro de

Oña, Cabeza de Vaca, Díaz del Castillo, Bartolomé de las Casas) records the impressions of

the European voyager who travels through this recently discovered land while taking note of

the customs of the society as well as the religious, ethnic and political groups that constitute

these nascent nations. These records, filled with impressions and detailed descriptions, allow

us to glimpse the everyday life of the colonial period. Through their eyes we discover how the

settlement of the European colonists developed and how the native communities resisted the

tyranny of the empires. They present us with (unilateral) perspectives showing how societies

were structured and the role of language or religion in the colonization process. These

documents also provide us with information about the peculiarities of the geography, the

climate, the flora, and fauna.

How could we build a digital library of inaugural rioplatense texts that we could study at a microscopic and and macroscopic level?

3.

Digital scholarly editions are one of the oldest and most prolific forms of output of digital DH projects (Pierazzo 2019). Even though the interest in Digital Scholarly Edition (DSE) and the use of TEI standards for scholarly publishing started really early in the Spanish-speaking world, especially in Spain (Marcos Marín 1986, 1994), scientific literature about it only found some continuity since 2008 (Lucía Megías 2008). However, still nowadays

there are few publications dealing with DSE and the use of the TEI in Latin American projects (Allés-Torrent 2015; del Rio Riande and Zubillaga 2015; Priani 2017; Allés Torrent and del Rio Riande 2019, among others).

On the one hand, it is important to mention that historically Philology has not been part of the main academic curriculum in Latin American. Approaches inherited from the French Academy, like Literary Criticisim, and later the American Cultural Studies, relegated Philology to the peripheries or even eliminated it from the Faculties of Arts and Humanities (Rodríguez Temperley). Consequently, interest in the Scholarly Edition, and later, on the DSE field, was somehow relegated to Medieval Hispanic or Colonial texts, taking the field not only to Textual Criticism approaches, but to others more related with Diplomatic or Genetic ones.

The landscape of projects and DSEs related to Hispanic texts shows some very specific features. Taking a look at the two existing catalogues (by Greta Franzini and Patrick Sahle) we can establish that:

- DSEs are common among scholars interested in Hispanic Medieval texts (such as *Cantar de Mío Cid*) and Early Modern Literature —the so-called Golden Age period , in which we find editions of classical plays such as Miguel de Cervantes' *La Entretenida* or Lope de Vega's *La Dama Boba*, and others regarding *Don Quixote* through largely funded projects (Allés-Torrent 2017). Yet, early Mexican manuscripts are also represented with *El sitio de Guaman Poma*, or the *Codice Mendoza*. Other initiatives deal with 19th and 20th century, such as *"En el ojo del huracán" Cartas de Yltramar a España, 1823 - Edición Digital*, *The Pérez Galdós Edition Project*, the *Manuscrito digital de Juan Goytisolo* or *Obras Completas de José Luis Romero*, among others. It is interesting to note that many of them were carried out in non-hispanic institutions (*Cantar de Mio Cid*, *El sitio de Guaman Poma*, *Manuscrito digital de Juan Goytisolo*), and that others, like the *Quijote Interactivo*, a great site funded by the Biblioteca Nacional de España, were built with technologies such as Adobe Flash Player, or not using have textual markup, as in the case of the *Códice Mendoza*, funded by the INAH in Mexico, which seems just to be focused on the fascinating images of the mesoamerican manuscript.

- If using it, few of these editions declare the use of TEI. Just to name a few of them, we may point to *La Dama Boba, La Entretenida,* the *Castillian Cancionero*, the digital edition of the *Cartas de Yltramar*, *Repertorio Métrico de la Poesía Medieval Española* (ReMetCa), *Poetriae*, *Obras Completas de José Luis Romero*, *Biblioteca Digital del Pensamiento Novo Hispano*, or *Diálogo Medieval* (Priani 2017).

- TEI suffered from discontinued use in the two big Spanish projects in Digital Libraries, such as Biblioteca Virtual Miguel de Cervantes and Corpus Diacrónico del Español (CORDE).

With regards to infrastructures and software for editing, annotating and publishing texts, it must be said that the progress done in digital textual scholarship has also lead to the creation of research environments, frameworks, and platforms in different countries: TAPAS, Boilerplate, Juxta, Version Machine 4.0 (English), TextGrid, Ediarium (Germany), eLaborate (Holland), Edictor (Brazil). For the Spanish community, to date the only example is Virtual Research Environment (EVI-LINHD), developed at the

Laboratorio de Innovación en Humanidades Digitales (LINHD-UNED) in Spain. EVI-LINHD creates digital projects, and uses an intuitive XML-TEI platform to encode texts (González-Blanco et al. 2016, 2017b).

As stated in this section, the history and actual landscape of the training in Humanidades Digitales in Spain and Latin America, and, specifically, in the use of the TEI as the standard for encoding and exchanging data, lacked continuity and has been mostly limited to Spanish projects in the last ten years. The general feeling perceived is a lack of awareness and mastery when compared to the rest of the Anglophone DH field, as well as little communication with other research groups. The case of the TEI is paradigmatic, marginal and not well known, especially if contrasted with other neighbour countries, such as France, where it was adopted early. Also, the use of open access tools and open access to data is not a very common practise.

In Latin America, the HD Lab, the digital humanities laboratory at the Argentinian CONICET (*Consejo Nacional de Investigaciones Científicas y Técnicas*) has been creating minimal editions via a workflow built around Recogito, an open source semantic annotation software developed by Pelagios Network (https://github.com/pelagios/recogito2), incorporating TEI markup and rendering the edited texts in static sites built with Jekyll and GitHub pages. This minimal, low-infrastructure approach was directly determined by the very limited funding and technological support granted to the lab.

These were supported by the international project Pelagios Commons, a community forum for using open data methods to link and explore historical places.[1] The projects *Pelagios al Sur* (2017)[2] and *LatAm: a Historical Gazetteer of Colonial Latin America and the Caribbean* (2018)[3] were funded through Pelagios' Resource Development Grants (RDG). Both projects focused on Argentinian and Latin American texts, approached from a Digital Humanities perspective and using tools for semantic tagging and georeferencing as well as Linked Open Data (LOD) technologies and Geographic Information Systems (GIS). It is worth noting that these were the first Digital Humanities projects focusing on Geohumanities or Spatial Humanities of Argentinian texts. The project *Pelagios al Sur* aimed at expanding the Pelagios Commons tools, originally designed for texts and geographies of the Antiquity, into the spatial and geographical frontiers of the New World through the use of their annotation platform Recogito[4]. The project focused on the interests of the Argentinian research community, defining the task to study texts and maps from the 16th and 17th centuries and the area of the Rio de la Plata. For the project, a group of historians from the *Instituto de las Culturas* (IDECU) CONICET and others collaborated using the Recogito platform to annotate and georeference a foundational text: *Historia argentina del descubrimiento, población y conquista de las provincias del Rio de la Plata*, also known as *La Argentina* or *La Argentina Manuscrita*, by Ruy Díaz de Guzmán (first edition 1602),

---

[1] http://commons.pelagios.org/.

[2] http://commons.pelagios.org/2017/12/pelagios-south-investigar-ensenar-y-aprender-informe-final/

[3] http://commons.pelagios.org/2019/02/final-report-on-latam-a-historical-gazetteer-of-colonial-latin-america-and-the-caribbean/

[4] https://recogito.pelagios.org/login, a tool for annotating and georeferencing texts and maps.

as well as related maps.[5] One year later, thanks to a second RDG, the Argentine project group joined an international team in order to develop a gazetteer based on a dataset from colonial Latin America, stemming from the *Diccionario geográfico histórico de las Indias Occidentales o América* (1786) by Antonio de Alcedo. The main goal of this second Pelagios project, *LatAm*, was to use Alcedo's dictionary to create a dataset linking Pelagios Commons and the World Historical Gazetteer (WHG). This project notably improved the georeferencing of Ruy Diaz de Guzman's text and the 17th century maps.

The first subcorpus will be annotated manually and semi-automatically with semantic (mentions of persons, events, and places) and geographic information (georeferencing) and will be used to improve the gazetteer previously developed in the *LatAm* project. This subcorpus will be primarily approached with close reading methods, investigating the development of the discourse about the region from medieval to early modern traditions. For semantic annotation, the tool Recogito will be used. The close reading approaches will be complemented by analyses on a large scale using maps and visualizations.

The second subcorpus will be primarily treated with distant reading methods, to answer questions about the differentiaton of narrations of space in the 19th century into various factual and fictional genres and about the national discourse about the region. For the distant reading approach, the following methods will be used: (a) linguistic annotation by means of Natural Language Processing (NLP), in particular Named Entity Recognition (NER) to automatically extract mentions of places, (b) Topic Modeling to computationally determine text passages containing topics and motifs related to descriptions and narrations of places and space,

to our knowledge only NER has been previously employed to analyze narratives of space in historical and literary texts (Losada Palenzuela 2017; Tilton et al. 2018). The three methods also have the initial limitation that they have been primarily developed for the analysis of texts in English, so that an evaluation with texts in Spanish is indispensable in order to use them, a task which means a qualitative leap for the study of texts in Argentinian Spanish.

The use of Topic Modeling for the analysis of literary texts in Spanish language are still scarce, though. For example, Navarro-Colorado (2018) applied the method to extract themes and motifs from a corpus of Spanish Golden Age poetry.

Representation (latin representatĭo) is the action and effect of representing (making something present with figures or words, referring, replacing someone, executing a work in public). A representation, therefore, can be the idea or image that replaces reality. Representation is also the way in which the world is socially constructed and represented at a textual and contextual level. The idea of representation is linked to the way in which the idea of subject is produced, that is, the descriptions and sets of values within which the different groups and / or individuals are identified. Moreover, following Davis, Shrobe & Szolovits (1993: 19) we can also claim that:

---

[5] https://recogito.pelagios.org/document/wzqxhk0h3vpikm/part/1/edit. Also, an index of Argentine toponyms was created: http://vocabularios.caicyt.gov.ar/geoar/vocab/index.php.

> [S]electing a representation means making a set of ontological commitments. The commitments are in effect a strong pair of glasses that determine what we can see, bringing some part of the world into sharp focus, at the expense of blurring other parts. These commitments and their focusing/blurring effect are not an incidental side effect of a representation choice; they are of the essence (…).

The respective combination of subcorpora and methods is on the one hand owed to the availability and properties of the sources: the earlier texts are fewer in number so that they lend themselves better to close reading approaches, while many more texts are available from the 19th century. Also, the language and orthography of the later texts is closer to modern Spanish so that less effort is needed to prepare them for the use with NLP and text mining tools.

Recogito is a web-based annotation tool developed by Pelagios Commons that enables annotation of geographic place references in text, images and data through a user-friendly online platform. Perhaps the most notable feature of Recogito is the ability to produce semantic data without the need to work with formal languages directly, while at the same time allowing the user to export the annotations produced as valid TEI-XML, RDF and GeoJSON formats.

Recogito supports import of plaintext (.txt extension) as well as TEI/XML encoded text (.xml extension). Users can identify concepts and entities —such as places, people and events—, and disambiguate and identify them. Currently, however, this LOD is transformed into TEI only partially, since not all annotation elements are supported in the export (e.g. commentary elements). Also, at present, the TEI export doesn't contain a schema declaration and if the original document had one, it is stripped out, meaning the output cannot be validated without some post-processing. For plain text documents, the TEI export is based on a fixed XML template, and for files originally uploaded as TEI, the original structure is not modified, nor XSLT transformations applied. In the case of plaintexts, only a tei_all schema actually validates the texts that are annotated in Recogito.

Many users of Recogito and Pelagios Commons tools have found the transition between the annotated text in Recogito and the final TEI result very useful for developing Digital Scholarly Editions (DSE), though this transition is not seamless at the moment.

plaintexts, only a tei_all schema actually validates the texts that are annotated in Recogito.

Likewise, Moretti (1999, 2005) understands maps as an instrument for textual analysis that illuminates hidden links.

The leading role of the anglophone community, as well as the obvious importance of English for global scientific communication, have outlined a landscape where resources in other languages are still difficult to find.

Una de las primeras cuestiones que surgen al afrontar la edición digital es la falta de modelos y prácticas consolidadas. A diferencia de los modelos impresos, no existe todavía una idea clara de cómo debe ser una edición digital, y eso dificulta tanto la concepción, como la ejecución y la implementación de cualquier proyecto editorial (Price 2009; Gabler 2010, Allés).

Following Goggin, a critical approach to new technologies relies in not only observing, documenting, analyzing, interpreting and constructing an account of what technology is but how it operates, what discourses and meanings are associated with it, who and what support it and how are the relations among different groups of people altered or believed to be altered through the introduction of the technology (122, 2012). Coined by Franco Moretti in 2005, distant reading refers to the analysis of large sets of data through computer based methods that aims to reveal correlations between the texts and deeper structural relations among them.

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 will enable us to

investigate these maps in new and innovative ways. The annotation of both elements, this is

to say, the logographic toponyms and names alongside alphabetic descriptions and place

names, 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.

Fig. 2 Close-up view of some of the Cempoala map annotations, including alphabetic place-names

(e.g. Tlacopa) and names of people (e.g. Don Francisco). Created using Recogito:

https://recogito.pelagios.org/.

The Project

Looking to introduce new and established scholars in Latin America to the field of Spatial

Humanities and train them in the use of Recogito, the aim of the project is to continue and

extend annotations of these maps, which we hope will offer new insight into perceptions of

space and place across sixteenth-century Mexico from both, an indigenous and European

perspective. By using Recogito to annotate and extract the information contained within these

maps, we will demonstrate the use of annotation software in the analysis of historical source

material which is not always entirely textual. This Recogito-LatAm project focusing on the

maps, will contribute new resources consisting of the vast array of geographic, historic, and

toponymical information featured on the maps. In addition, the information extracted from

these maps will expand upon the existing gazetteers available within Recogito, furthering the

scope of these gazetteers and showcasing sixteenth century geographies, particularly in

relation to indigenous
knowledge.

¿De qué hablamos cuando hablamos de ediciones digitales?

- Edición digital de textos (general)

- Edición digital académica (revistas, libros)

- Edición digital filológica (crítica, anotada, enriquecida)

Una de las primeras cuestiones que surgen al afrontar la edición digital es la falta de modelos y prácticas consolidadas. A diferencia de los modelos impresos, no existe todavía una idea clara de cómo debe ser una edición digital, y eso dificulta tanto la concepción, como la ejecución y la implementación de cualquier proyecto editorial (Price 2009; Gabler 2010).

DSEs are at the core of digital humanities [Earhart 2012].

While not all textual scholars might rely on the same definition of DSEs, they recognize their features and uses [Sahle 2016]. Free, open standards such as the ones developed by the Text Encoding Initiative (TEI), along with eXtensible Markup Language (XML) technologies, such as eXtensible Stylesheet Language Transformations (XSLT) and XQuery, and dedicated software have characterized the digital editing field.

We aim at using a set of methods stemming from Computational Linguistics and Information Retrieval for the distant reading of narratives of the Argentinian space and to compare the results achieved with them to findings obtained from close reading texts based on the methods of text encoding, semantic, and geographic annotation.

Summary

Investigating close and distant reading approaches to the narratives of the Argentinian space from the 16th to the 18th century entails building on previous work in various areas of research. The chapter delves in the formalization of models of narrative space. It is also concerned with text encoding and geographic annotation as methods for the close study of narrative space and narratives of space in the digital philological edition of a corpus of colonial texts that describe for the first time the Río de la Plata region. These close reading techniques will be then compared to quantitative methods stemming from Computational Linguistics and Information Retrieval to examine the same phenomena from the distance.

From the point of view of Literary Studies, general narratological approaches to the topic, as well as previous investigations on narratives of the Argentinian space are relevant points of reference for this chapter (Morillas Ventura 2008). Also, the approaches of Literary Geography, Literary Cartography, and GeoHumanities usually rely on explicit mentions of toponyms to build their data basis. Both of these disciplines allowed the Literary Cartography to emerge (Piatti et al. 2009), which highlights the scenery of literary works, understanding landscape as something to be read even beyond genres like travelogues, chronicles, memoirs, etc. (Bushell 2012: 151). However, the constitution of narrative space is usually defined by more complex factors, as well, which has already been investigated in narratological works on space (Dennerlein 2009), but has rarely been taken up in the Digital Humanities. An event region designates a space which is relevant for the plot while mentioned spatial facts do not directly relate to situations in the narration. The textual encoding of texts in digital format, enriching them for example with metadata, structural, semantic and geographic information, opens the texts up to a computational treatment. Still, depending on how the information is added and used, these practices are in the first place related to a close reading of the texts, aimed at the thorough interpretation of the individual texts. When annotations are realized automatically without the possibility for comprehensive manual verification and when the digital texts are interpreted on the basis of quantified analysis results and their visualization, the digital techniques can be understood as modes of distant reading, where the textual structure is dissolved in order to generate abstract views (Moretti 2005, 2013; Jockers 2013; Jänicke et al. 2015).

We will explain the digital close and distant reading methods used Digital Humanities Lab (HD Lab, CONICET) in the digital edition of La Argentina Manuscrita (Rui Diaz de Guzmán), Un viaje al Río de la Plata (Acarette Du Biscay) and Relación de las cosas

sucedidas en el Río de la Plata (Pero Hernández)[6].These editions approached the texts from a Digital Humanities perspective and using tools for semantic tagging and georeferencing as well as Linked Open Data (LOD) technologies and Geographic Information Systems (GIS). All the texts were encoded in the XML-based standard of the Text Encoding Initiative (TEI), adding metadata and structural information. Digital annotation techniques, NLP and text mining methods chosen (semantic and geographic tagging, named entity recognition) will be explained in order to describe what aspects of models of narrative space can be covered with these methods and how. As a result, a roadmap for the digital analysis of narrative Argentinian space will be formulated.

Bibliographic references

Bushell, S. (2012). The Slipperiness of Literary Maps: Critical Cartography and Literary Cartography. In: *Cartographica: The International Journal for Geographic Information and Geovisualization* 47 (3), 149-160.

Dennerlein, K. (2009). *Narratologie des Raumes*. Berlin: de Gruyter.

Jänicke, S., Franzini, G., Cheema, M. F., Scheuermann, G. (2015). On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In: *Proceedings of the Eurographics Conference on Visualization (EuroVis) 2015*. Edited by Borgo, R., Ganovelli, F., Viola, I. Italy, Cagliari: The Eurographics Association. http://dx.doi.org/10.2312/eurovisstar.20151113

Jockers, M. L. (2013). *Macroanalysis: Digital Methods & Literary History*. Urbana, Illinois: University of Illinois Press.

Moretti, F. (2005). *Distant Reading*. London: Verso.

Morillas Ventura, E. (2008). Textos inaugurales: Los relatos de los viajeros patagónicos. In: *Anclajes* 12. http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-46692008000200003

Piatti, B., Reuschel, A. K., y Hurni, L. (2009). Literary Geography –or How Cartographers Open up a New Dimension for Literary Studies. In: *Procedimientos de la 24ª International Cartographic Conference* (ICC). Santiago de Chile.

Ternavasio (2009). «1. Ser parte de un gran imperio». Historia de la Argentina 1806-1852. Siglo veintiuno. ISBN 9789876290937.

1.

What is text? I am not so naïve as to imagine that

question could ever be finally settled. Asking such a

question is like asking 'How long is the coast of

England?'."(J. McGann)

---

[6] http://hdlab.space/biblioteca-digital/.

.

.

"Text is what you look at. And how you look at it." (P.

Sahle)

Como cada uno de nosotros codifica "la realidad", lo que Alfred Korzybski denomina el **territorio**, de una manera particular, todos tenemos **mapas** diferentes de la misma. Esto es el **significado de** la frase "El **mapa no es el territorio**"

La esencia de la obra de Korzybsky es la declaración de que los seres humanos están limitados en su conocimiento por la estructura de su sistema nervioso y la estructura de sus lenguas. Los seres humanos no pueden experimentar el mundo directamente, sólo a través de sus abstracciones (impresiones no verbales que provienen del sistema nervioso e indicadores verbales que provienen de la lengua). A veces las percepciones y la lengua confunden al hombre que cree que son los *hechos* con los que debe tratar. El entendimiento humano de lo que está pasando carece en ocasiones de *similitud de estructura* con lo que está pasando realmente.

"If words are not things, or maps are not the actual territory, then, obviously, the only possible link between the objective world and the linguistic world is found in structure, and structure alone."

1. Philology, Textual Criticism, and Digital Scholarly Editions in Argentina

https://www.cervantesvirtual.com/obra-visor/amado-alonso/html/e586d6e9-ca1e-401c-8aa0-b1b361471eb5_4.html

El Instituto de Filología de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires se creó en 1923 gracias a la actividad conjunta de Ricardo Rojas, decano de dicha Facultad, y de Ramón Menéndez Pidal. Con su creación se pretendía inaugurar un centro donde se investigara la lengua castellana y, más concretamente, las peculiaridades de su variante argentina.

El complejo conjunto de procesos históricos entre los que se desarrolla la emancipación, formación y afianzamiento de los Estados en América del Sur comprende en distintos momentos intervenciones destinadas al conocimiento y la gestión de la(s) lengua(s) y la tradición (literaria, aunque no solamente) en el territorio, que apuntan a la concepción y organización de la comunidad política. Desde las preguntas por la gramática a enseñar, la potestad para intervenir la ortografía, la posibilidad de delimitar una tradición textual propia, la organización de las voces en su representación letrada, la pregunta por qué hacer y dónde ubicar a las lenguas indígenas, las diversas formas de vindicación de la unidad de la lengua o de la gestión soberana de una variedad propia, hasta los intentos más o menos exitosos de la formación de un campo específico para el estudio científico de la lengua, cobra forma un corpus que se proyecta en numerosas direcciones pero que está atravesado por líneas de consistencia que han permitido, en época reciente, la apertura de un campo de estudio transitado por disciplinas diversas. En ese marco, y con las herramientas que proveen la historiografía lingüística, la glotopolítica, la filología, la crítica genética y

la historia material de la prensa y el libro, es que se propone un trabajo de relevamiento, sistematización, expansión, profundización, difusión y examen crítico de ese corpus.

If editing is "without doubt one of the oldest scholarly activities within the Humanities" [Pierazzo 2016], DSEs are at the core of digital humanities [Earhart 2012]. While not all textual scholars might rely on the same definition of DSEs, they recognize their features and uses [Sahle 2016]. Free, open standards such as the ones developed by the Text Encoding Initiative (TEI), along with eXtensible Markup Language (XML) technologies, such as eXtensible Stylesheet Language Transformations (XSLT) and XQuery, and dedicated software have characterized the digital editing field. The scholarly editions themselves, however, haven't always been successful as open products of research. Bodard and Garcés posed this issue when they claimed that, analogous to the Open Source Software movement, DSEs—which they call "Open Source Critical Editions" or "OSCE"—should be licensed for reuse, including all sources, data, methods and software [Bodard and Garcés 2009]. While it is common practice to make TEI data publicly available, the debate over how DSEs need to be structured to be truly "open" is still ongoing and best practices have yet to be established. Hannesschläger, for example, has recently surveyed licenses appropriate for DSEs powered by TEI and singled out Creative Commons ones as appropriate for an international context and for enabling an open culture of reuse with a global impact; however, there has yet to be widespread agreement with this proposal [Hannesschläger 2019].

From a Global South perspective, the DSE field is perceived as being dominated by standards and technologies that are still unfamiliar to scholars; not surprisingly, these methods are typically described in the context of Anglophone projects [Allés-Torrent and del Rio Riande 2020]. Indeed, beyond some very specific projects and initiatives, multilingual resources related to DSEs, such as tutorials, software, books, and articles, are generally difficult to find in languages other than English. Moreover, the use of proprietary software for most editorial work has become a barrier for extending the DSE practice beyond Northern academies. Indeed, "[d]igital scholarly editions are expensive to make and to maintain" [Pierazzo 2019]: as such, long-term web hosting, preservation, and access to servers pose a significant obstacle for not-so-well-funded scholars who lack access to grant funding or other institutional resources. Overall, DSEs require substantial infrastructure and advanced technical skills, while diverse needs, capacities, priorities, languages, and academic traditions may require different features from DSEs at a global scale.

With that in mind, how can DSEs, one of the crown jewels of digital humanities [Pierazzo 2016], become global? From our perspective, this can be achieved by establishing a "digital commons." When the GO::DH Minimal Computing Working Group started a debate on power and inequality in digital humanities from a technological perspective, its intention was not to simply criticize or mourn a lack of diversity, but to establish an alternative discourse and create a new set of commons, namely technology of disobedience, architecture of necessity, and the moral modulor [Gil 2016]. The principles of minimal computing turned into new ways of undertaking digital humanities work and collaborating to building an alternative digital epistemology that has found a practical outlet in minimal DSEs with Ed., a tool for building minimal DSEs without text encoding.

Thus, minimal computing informed the design of our joint course that teaches digital publishing and textual scholarship with minimal computing and text encoding. Specifically, we highlight minimal computing as a shared set of values such as the use of open technologies, ownership of data and code, and reduction in computing infrastructure. Even though the interest in the use of TEI for Digital Scholarly Edition (DSE) started really early in Spain (Marcos Marín 1986, 1994), scientific literature about

it only finds some continuity in this country since 2008 (Lucía Megías 2008). Gradually, other articles came to light regarding general overviews that concentrate just on Spain, however, they do not reflect any research about the TEI in Latin American (Spence 2014; Spence and González-Blanco 2014; Allés-Torrent 2017) or they only cover study cases of projects (Revenga 2014; Rojas Castro 2017). There are few publications dealing with TEI and its use in Spanish and Latin American projects (Faulhaber 1996; Fradejas 2009; Allés-Torrent 2015; González-Blanco et al. 2014, 2015; del Rio Riande and Zubillaga 2015; Priani 2017, among others).

The landscape of projects and DSEs related to Hispanic texts shows some very specific features. Taking a look at the two existing catalogues (by Greta Franzini and Patrick Sahle) we can establish that:

- DSEs are common among scholars interested in Hispanic Medieval texts (such as *Cantar de Mío Cid*) and Early Modern Literature —the so-called Golden Age period , in which we find editions of classical plays such as Miguel de Cervantes' *La Entretenida* or Lope de Vega's *La Dama Boba*, and others regarding *Don Quixote* through largely funded projects (Allés-Torrent 2017). Yet, early Mexican manuscripts are also represented with *El sitio de Guaman Poma*, or the *Codice Mendoza*. Other initiatives deal with 19th and 20th century, such as *"En el ojo del huracán" Cartas de Yltramar a España, 1823 - Edición Digital*, *The Pérez Galdós Edition Project*, the *Manuscrito digital de Juan Goytisolo* or *Obras Completas de José Luis Romero*, among others. It is interesting to note that many of them were carried out in non-hispanic institutions (*Cantar de Mio Cid*, *El sitio de Guaman Poma*, *Manuscrito digital de Juan Goytisolo*), and that others, like the *Quijote Interactivo*, a great site funded by the Biblioteca Nacional de España, were built with technologies such as Adobe Flash Player, or not using have textual markup, as in the case of the *Códice Mendoza*, funded by the INAH in Mexico, which seems just to be focused on the fascinating images of the mesoamerican manuscript.

- If using it, few of these editions declare the use of TEI. Just to name a few of them, we may point to *La Dama Boba, La Entretenida,* the *Castillian Cancionero*, the digital edition of the *Cartas de Yltramar*, *Repertorio Métrico de la Poesía Medieval Española* (ReMetCa), *Poetriae*, *Obras Completas de José Luis Romero*, *Biblioteca Digital del Pensamiento Novo Hispano*, or *Diálogo Medieval* (Priani 2017).

- TEI suffered from discontinued use in the two big Spanish projects in Digital Libraries, such as Biblioteca Virtual Miguel de Cervantes and Corpus Diacrónico del Español (CORDE).

With regards to infrastructures and software for editing, annotating and publishing texts, it must be said that the progress done in digital textual scholarship has also lead to the creation of research environments, frameworks, and platforms in different countries: TAPAS, Boilerplate, Juxta, Version Machine 4.0 (English), TextGrid, Ediarium (Germany), eLaborate (Holland), Edictor (Brazil). For the Spanish community, to date the only example is Virtual Research Environment (EVI-LINHD), developed at the Laboratorio de Innovación en Humanidades Digitales (LINHD-UNED) in Spain. EVI-LINHD creates digital projects, and uses an intuitive XML-TEI platform to encode texts (González-Blanco et al. 2016, 2017b).

The use of the TEI as the standard for encoding and exchanging data, lacked continuity and has been mostly limited to Spanish projects in the last ten years. The general feeling perceived is a lack of awareness and mastery when compared to the rest of the Anglophone DH field, as well as little communication with other research groups. The case of the TEI is paradigmatic, marginal and not well known, especially if contrasted with other neighbour countries, such as France, where it was adopted early. Also, the use of open access tools and open access to data is not a very common practice.

https://pueblosoriginarios.com/recursos/mapas/coronelli/coronelli.html

https://www.diarioalfil.com.ar/2021/02/08/indicaciones-para-nunca-llegar-a-trapalanda-primera-parte/

https://heraldicaargentina.blogspot.com/2018/

https://cadenaser.com/nacional/2022/06/09/walt-whitman-el-poeta-de-america-cadena-ser/