DH as an Ideal Educational Environment:
the Ethnographic Museum of La Spezia

Letizia Ricci
University of Pisa
l.ricci29@studenti.unipi.it

Francesco Melighetti
University of Pisa
f.melighetti@studenti.unipi.it

Federico Boschetti
CNR-ILC, Pisa &
VeDPH, Ca' Foscari Venezia
federico.boschetti@ilc.cnr.it

Angelo Mario Del Grosso
CNR-ILC, Pisa & University
of Pisa
angelo.delgrosso@ilc.cnr.it

Enrica Salvatori
LabCD,
University of Pisa
enrica.salvatori@unipi.it

Abstract

English. The authors present the outcomes of an educational experimentation that took place in the academic year 2018-2019 at the degree course in Informatica Umanistica at the University of Pisa. The first objective of the project concerned the digitization of a corpus of postcards from the period of the First World War owned by the ethnographic Museum of La Spezia "G. Podenzana". The aims of the work are not only the historical study of the corpus, but also the organization of a public history project with the Museum.

Italiano. Gli autori presentano i risultati di una sperimentazione didattica svolta durante l'anno accademico 2018-2019 presso il Corso di Laurea in Informatica Umanistica dell'Università di Pisa. Il primo obiettivo del progetto riguarda la digitalizzazione di un corpus di cartoline del periodo della Prima Guerra Mondiale, di proprietà del Museo Etnografico di La Spezia "G. Podenzana". Gli obiettivi del lavoro non sono solo lo studio storico del corpus, ma anche l'organizzazione di un progetto di Public History con il Museo.

1 Introduction

The authors present the outcomes of an educational experimentation that took place in the academic year 2018-2019 at the degree course in Informatica Umanistica at the University of Pisa. The experimentation involved the courses of Digital Public History, Digital Text Encoding as well as Digital Philology, and at the beginning concerned the digitization of a corpus of postcards from the period of the First World War owned by the ethnographic Museum of La Spezia "G. Podenzana".

The postcards have been historically contextualized, digitized, placed on a collaborative web platform and distributed to the students in order to be recorded, transcribed and encoded in XML-TEI. Students have been involved in the development of the web platform by tracking the usability issues as beta-testers. Students contributed also to the requirement analysis and the definition of the specifications necessary to extend the platform to a broader audience of users without specific skills in Digital Humanities.

Indeed, the project aims were not only the historical study of the corpus, but also the organization of a public history project with the Museum, its targeted audience and the High School students of La Spezia. Arriving almost at the end of this educational experiment, we propose now to discuss the current achievements, based on the common educational statement of "learning by doing" and announced months ago at AIUCD 2019.

2 Background

Within the previous annual conference of the Italian Digital Humanities Association held in Udine (AIUCD2019), a preliminary work towards a profitable collaboration between students and teachers of different DH classes at the University of Pisa was presented. The context of the project (Booth, 1996; Cati, 2006; Cole, 2016; Delle Cave, 2013) gave the actors the opportunity to collaborate with each other and with the "G. Podenzana" museum of La Spezia outside the formal classroom constraints, involving also some activities carried out by non academic communities (Salvatori, 2017). In Salvatori et al. (2019), the authors discussed the objectives and the outcomes that have been achieved during the bootstrap phases of the project. In particular, within that paper, they pointed out the main problems and the added values of the initiative introduced above.

To date, the collaboration has been getting wider and a few internships have been activated to improve the design and development skills of the interested students in order to enhance the tools already developed within the *Euporia* platform (Mugelli et al., 2016). These new activities facilitate students and general users in digital encoding historical documents, which have an inherently complex nature. This objective has been possible by adopting a formal but "common" rules for annotating and for processing textual data (Fowler, 2010). Moreover, as far as the actual TEI-XML encoding work (Burnard, 2014; Pierazzo, 2015) which concerns the digitization, the recording and transcription of the postcards provided by the involved cultural institutions, the main problems have been overtaken by putting in place a chain of document processing tasks. This process has been developed by using XSLT technology and by implementing a Web environment to publish the encoded documents (Del Turco and Di Pietro, 2016).

3 Methods

The project involves students, teachers and representatives of the Ethnographic Museum of La Spezia, which collaborate by sharing information, resources and tools. The Museum has made available two large corpora of postcards dating back to the Great War period, which have been digitized, uploaded onto the *Euporia* platform and encoded in XML-TEI by the students under the supervision of the teachers, according to a custom subset of tags declared in the *CartolineXML* schema. Furthermore, the Museum played the role of an interdisciplinary meeting center among High School students of La Spezia, students and teachers of the University of Pisa and the Museum managers, in order to share ideas about the project from different perspectives.

Two students of the aforementioned courses and co-authors of this contribution, made an internship at the CNR-ILC to improve their skills in text encoding. They focused on the simplification of the annotation process, in order to involve High School students and volunteers that could actively collaborate in transcribing and annotating postcards, even if they have not specific skills in XML-TEI encoding.

Therefore, they have defined a Domain Specific Language (DSL), *CartolineDSL*, with the same expressivity of its counterpart in XML-TEI but much less verbose. A DSL is a formal language with a simple, understandable and suitable syntax for the domain of interest we are dealing with and based on a limited and controlled vocabulary. A DSL must be defined by a Context-Free Grammar (CFG), that is a set of recursive rewriting rules used to generate string patterns. Therefore *CartolineDSL* is a language suited to the domain of postcards characterized by a series of "*attribute: value" fields that users can easily fill in.

```
 3    doc: title description? body note;
 4    title: TITLE_HEADER text;
 5    description: DESCRIPTION_HEADER figure? notes?;
 6    body: bodyText recipient;
 7    note: NOTE_HEADER text?;
 8    bodyText: TEXT_HEADER opener letterbody closer notes?;
 9    recipient: RECIPIENT_HEADER address notes?;
10
11    opener: date placeName? initialSalute?;
12    closer: finalSalute? signed;
13    notes: NOTES_TAG text?;
14
15    figure: FIGURE_TAG text;
16    date: DATE_TAG dateValue;
17    placeName: PLACE_NAME_TAG text;
18    initialSalute: INITIAL_SALUTE_TAG text?;
19    finalSalute: FINAL_SALUTE_TAG text?;
20    letterbody: LETTERBODY_TAG (text|linebreak)+;
21    signed: SIGNED_TAG text;
22
23    address: persName street city province?;
24    persName: NAME_TAG text;
25    street: STREET_TAG text;
26    city: CITY_TAG text;
27    province: PROVINCE_TAG text?;
28    dateValue: DATE;
29
30    linebreak: '//';
```

Figure 1: Formal grammar code snippet for the postcard corpus within the *Euporia* digital environment

Fig. 1 illustrates some rewriting rules in the CFG of *CartolineDSL*, whereas Fig. 2 shows an example of data and metadata encoded in *CartolineDSL*.

===TITOLO===
Sommeil Interrompu

===DESCRIZIONE===
*immagine: Una donna interrompe il sonno di un uomo che si riposa vicino ad un albero.
*note: Sono presenti tre timbri e un francobollo.

===TESTO===
*data: 24/07/1913
*luogo: Ravenna
*saluto:
*corpo: A Ravenna piove sempre: è una gioia per chi non è andato ai bagni. Per ora niente di nuovo, tutti bene. Stasera qui c'è musica, come ieri che fu S.Apollinare: ad Arona, niente
*commiato: Saluti affettuosi,
*firma: Giuseppe
*note:

===DESTINATARIO===
*nome: Signorina Oliva Turtura
indirizzo: Via Cavour 12
*località: Arona (Lago Maggiore)
*provincia:
*note:

===NOTE===
La cartolina si presenta in un buono stato di conservazione.

Figure 2: *CartolineDSL* snippet

The conversion of *CartolineDSL* to XML-TEI is performed in two steps. In the first phase, the annotations encoded in *CartolineDSL* are parsed by the ANTLR compiler compiler (Parr, 2013) and converted in XML. The proprietary *CartolineML* schema allows the serialization in XML of the Abstract Syntactic Tree (AST) parsed by ANTLR. In the second phase, the proprietary XML document is converted to XML-TEI by an XSLT transformation. The XSLT style-sheet has been created by the students on the latest part of their internship at the ILC-CNR.

As usual, other XSLT style-sheets are necessary to transform XML-TEI in HTML for visualization purposes. Fig. 3 shows the designed interface.

Figure 3: Mockup sketch of the ongoing web-app aimed at publishing the archive

4   Results

The main achievements of this didactic experimentation are listed below: 1. coordination of three courses, in order to work on the same materials from different perspectives (Public History for the historical contextualization of the project, Text Encoding for XML-TEI models and technologies, and Digital Philology for the treatment of uncertain readings and for the creation of an optimized human readable DSL); 2. engagement of students in the annotation process of a large sample of the postcards corpus (learning by doing); 3. transfer of knowledge and experience from students of the University of Pisa and High School students of La Spezia during meetings at the Museum; 4. involvement of students in the creation of Domain-Specific Languages meant to bridge the gap between the best practices of Digital Humanists and the simple practices of unskilled citizens that desire to participate in projects of Public History.

5   Conclusion and Future Work

We guess that the educational model that we experimented can be easily exported in other contexts, with a broader involvement of multidisciplinary communities of practice and applied to different textual and/ or iconographic (or multimedia) digital resources.

In the next academic year we will release an updated version of *Euporia*, which currently is just a prototype, in order to allow High School students and volunteers to annotate further postcards in *CartolineDSL*.

References

A. Booth. 1996. *Postcards from the Trenches: Negotiating the Space Between Modernism and the First World War*. Oxford University Press.

L. Burnard. 2014. *What is the Text Encoding Initiative? Encyclopédie numérique 3*. OpenEdition Press.

I. Cati. 2006. *Cara mamma ti scrivo: le cartoline dei soldati della grande guerra*. Gasparri.

K.J. Cole. 2016. *Postcards from the Front 1914-1919*. Amberley Publishing Limited.

R. Del Turco and C. Di Pietro. 2016. *Between innovation and conservation: the narrow path of UI design for the DSE*, University of Graz, Graz.

L. Delle Cave. 2013. *Orme di guerra : lettere e cartoline dal fronte (1912-1919)*. Samus.

M. Fowler. 2010. *Domain-Specific Languages*. Addison-Wesley Signature Series (Fowler). Pearson Education.

G. Mugelli, F. Boschetti, R. Del Gratta, A.M. Del Grosso, F. Khan, and A. Taddei. 2016. A user-centred design to annotate ritual facts in ancient greek tragedies. *BICS* 59(2):103–120.

T. Parr. 2013. *The Definitive ANTLR 4 Reference*. Pragmatic Bookshelf, 2nd edition.

E. Pierazzo. 2015. *Digital Scholarly Editing : Theories, Models and Methods. Digital Research in the Arts and Humanities*. Ashgate, Farnham Surrey.

E. Salvatori. 2017. Digital (public) history: la nuova strada di una antica disciplina. *RiMe* 1-I:57–94.

E. Salvatori, F. Boschetti, and A.M. Del Grosso. 2019. From collaborative transcription to interdisciplinary education: the postcards of the great war case. In *AIUCD2019 Book of Abstracts*. Udine.