

Two Workflows for the Digital Edition of the Unpublished Elisa Chimenti's Corpus (Naples 1883 - Tangier 1969)

Bianca Vallarano
Dept. of Literary, Linguistic and Comparative studies
University of Naples "L'Orientale"
Naples, Italy
bianca.vallarano@unior.it

Ada Desideri
Dept. of Italian studies
Sorbonne University
Paris, France
ada.desideri@sorbonne-universite.fr

Abstract — This paper presents two digitisation workflows in the digital processing of unpublished texts by Elisa Chimenti (1883-1969). The author's corpus is extremely varied and interesting both from the point of view of academic research and for a wider public. The experiments therefore move in the area of open access digital scholarly edition. The aim is to compare the different text processing chains, in order to reflect on the pros and cons of each and thus contribute to the ongoing project of organising, preserving and digitising this heritage.

Keywords — Elisa Chimenti, Digital Scholarly Edition, XML-TEI, Digitisation workflows, Text encoding.

I. INTRODUCTION

This paper presents the preliminary proceedings of a research which aims to build a model for the digital edition of Elisa Chimenti's work. This research is in the context of an International Associated Laboratory, established between the University of Rome La Sapienza and the University of Lille, about the writing of women's exile in the Mediterranean region (XIX-XX century), coordinated by Silvia Tatti and Camilla Cederna.

We therefore discuss two different digitisation workflows we developed, trying to point out the pros and cons of each experiment. Both experiments aim at a scholarly open access digital edition of unpublished contemporary texts. Analysing the two workflows will provide an opportunity to reflect on the broader issue of editing unpublished twentieth-century typewritten texts under copyright. As a result, we will address the issues of transcription and OCR softwares, the still unmodelled status of digital author philology, copyright, and software ownership. The aim of our intervention would therefore be to contribute to the first axis proposed in the call for papers, "Digitisation workflows, tools and infrastructures", with methodological reflections on some techniques in the field of digitisation. The topic would be tackled starting from experiments on text editions, in the context of the digitisation of the Elisa Chimenti Fund.

II. ELISA CHIMENTI'S WRITTEN HERITAGE: WHY A DIGITAL SCHOLARLY EDITION?

A. Elisa Chimenti

Elisa Chimenti (Naples 1883-Tangier 1969) was an Italian-born writer, journalist and teacher who grew up and lived in the Maghreb. Her work focuses on the dialogue between cultures, languages, and religions from one side of the Mediterranean to the other, with a particular attention on

the condition of women [4] [5] [6]. Her approach promotes understanding and respect for every human and not-human being; she criticises the behaviour of European people and institutions towards Maghrebian cultures [20]. These represent pressing issues in the present context of repression against people on the move from Africa to Europe. Within academic research, we furthermore can notice an increasing consideration for the cultural hybridization processes, whose Elisa Chimenti represents a peculiar example.

Although she published only five works during her lifetime,¹ she left an extremely large corpus of unpublished works, including novels, short stories, essays, and poems (more than 30 works). The corpus is preserved today in the Elisa Chimenti Fund [26] in the Palace of Italian Institutions in Tangier, Morocco. The preservation conditions of the documents are precarious and access to them is very difficult; most of them were photographed with personal instruments by two researchers, who shared the images with other colleagues to ensure their preservation and allow academic research.

B. A Digital edition

As the texts are still subject to copyright, digitisation initially focused on producing critical editions that would remain in an academic environment, solely to facilitate scholarly research on the texts. As new possibilities for dissemination have opened, thanks to the ongoing negotiations with the heirs for the acquisition of rights, workflows oriented towards open access digital editions, which can also be consulted by a non-academic public, are beginning to be explored. Open access publication is important to enable the circulation and study of this heritage, which is still mostly unpublished and difficult to access, both for security reasons (documents are very fragile) and material conditions of access to the fund (the facility is not accessible, and documents are not fully archived). Currently, a cataloguing and metadata project is being undertaken with the "Archives Chimenti" project, promoted by the UMR Thalim (CNRS-ENS-Sorbonne Nouvelle) [30] in the EMAN digital library [24].² In this way, it is possible to know the content of the corpus in broad terms, even if the texts are not available for the public. This can increase the interest of academics or other actors in publishing and culture regarding Elisa Chimenti. For those who are already working on this writer, access to metadata makes it possible to contextualise the examined object within the whole work. This is particularly

¹ *Èves marocaines* (Tangier 1935), *Chants de femmes arabes* (Paris 1942), *Au cœur du harem* (Paris 1958), *Légendes marocaines*, (Paris 1959), *Le sortilège et autres contes séphardites* (Tangier 1964). Three works were translated: *Tales and legends of Morocco* (New York 1965), *Al cuore dell'harem* (Rome 2000), *Cuentos del Marueccos español* (Madrid 2003).

The works were recently collected in [7]. Finally, two other texts were supposedly published by the author in her youth in Germany, but no evidence remains (probably lost during the First World War): *Meine Lieder* (Leipzig 1911), *Taitouma* (Leipzig 1913).

² The project started in January 2023, coordinated by Camilla Cederna.

interesting in this corpus, in which some texts (short stories above all) are present in several collections.

Elisa Chimenti's corpus is extremely diverse, both in terms of genre and content of works. While we can find clean, ready-to-publish texts, there are also several documents with minor to major rewrites, corrections, and variants. Therefore, digitisation processes also vary. Obviously, with regard to texts with a rich apparatus of variants, specific considerations will have to be made in the field of authorship philology. The issue of modelling heterogeneity in author philology is as open as is the discussion around the standardisation of an encoding model and set of scalable visualisation software.

Therefore, the choice of digital seems to have several advantages: the possibility of preserving and studying these texts in an academic context; the possibility of relating recurring themes by considering many texts, even apparently very different from each other, by the means of the creation of new paths and links; the possibility of carrying out linguistic research; the possibility of making the texts accessible on the Web. Moreover, from the digital-humanities point of view, the access to such a varied corpus and the great openness in the possibilities of experimentation, represent a very fruitful framework for testing and formalising old and new workflows for the creation of digital scholarly editions.

For this paper, we will only consider novels and short story collections, as no essay has yet been fully digitised, and the poetry collections are already the subject of an ongoing paper edition [8]. The texts that have been the subject of this work both have minor corrections and authorial variations. This is the case for a large part of the documents in the Chimenti corpus, and for this reason the two texts can be illustrative and serve as a model. However, they are clearly not exhaustive of the variety of textual situations that can be encountered.

III. TWO DIGITISATION WORKFLOWS FOR A SCHOLARLY DIGITAL EDITION OF CHIMENTI'S WORK³

In this section, we will detail two workflows that have been applied to three Elisa Chimenti's unpublished texts:

1. The first workflow is based on text editors and XSLT transformations. It was used for digitising a short story of the collection *Contes Berbères* (*Mennouch la Rifaine*) and an excerpt of about 50 pages of a science-fiction novel (*Une étrange aventure*) [3]. They are both unique copy;⁴
2. The second workflow is based on XML-TEI [28] and EVT [25] and was used to digitise a short story of the collection *Contes d'autrefois* (*Histoire du prince Moustapha et de la princesse Djamila*) [1]. The short story is also included in another collection, *La veillée du harem* [2].

A. First workflow

The first scholarly digital edition produced in the digitisation project of the Elisa Chimenti Fund⁵ were created in the context of two master's thesis at the University of Lille and the University of Pisa.⁶ The two texts analysed are very

different from each other, both in terms of content and in terms of the form and condition of the document; different choices were therefore made regarding the phenomena to be processed and certain aspects of the final graphic rendering. Indeed, *Mennouch la Rifaine* is a love short-story, set during the Rif War and features numerous interventions and handwritten notes by the author. *Une étrange aventure* (fig. 1), on the other hand, is in a ready-to-print version, almost without manuscript interventions, and presents a very peculiar hybridisation between supernatural aspects of Western and Moroccan cultures.

The workflow applied remains the same in both cases and the work was carried out in close collaboration with the student from University of Lille, Pauline Modolo: the text was initially transcribed, starting from the images of the documents, in Word format, applying paragraph and character styles; then the XML file that allows the graphical rendering in Word was extracted,⁷ and was transformed into a second XML file.⁸ In this file, only the tags associated with the character and paragraph styles added during transcription appear. This allowed us to generate an XML-TEI file and then the HTML pages that constitute the edition interface, with further XSLT transformations (a scheme in Table I).

For the novel *Une étrange aventure*, we first tried to transcribe the text using tesseract, an open OCR software [29]. The results were not satisfactory, considering the regular form of the text, and we therefore opted for manual transcription (already chosen for the *Contes berbères*). In fact, due to the low quality of the images (less than 300 dpi) each page needs special contrast and colour adjustment to maximise the OCR result. Spending this time on manual transcription seemed to be preferable, rather than testing page by page for the best result. This also allowed a closer approach to the text, which is suitable for a critical edition work. At present, due to the presence of about fifty manually transcribed pages, we can consider training OCR for the transcription of the nearly 200 remaining pages.

During transcription, errors that could be considered typing errors (inversion, addition or missing letters, etc.) or due to the typewriter (missing accents and other symbols) were corrected. Capital letters and punctuation were also normalised. The application of styles in Word (or in

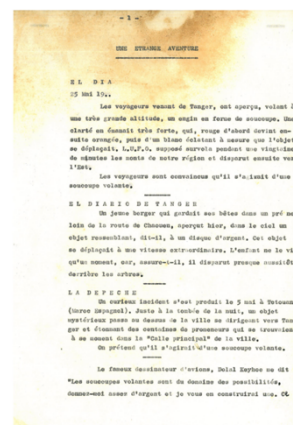


Fig. 1 Page 1 of the typescript of *Une étrange aventure*.

³ Gitlab hosts both editions. For research purposes, they can be shared privately upon request [27].

⁴ An Italian version of *Une étrange aventure* exists, translated by an unknown person ("L. Monsello"). We didn't take it into account for establishing this edition.

⁵ Coordinated by Camilla Cederna.

⁶ Pauline Modolo's dissertation on the short story *Mennouch la Rifaine*, and Ada Desideri's on the novel *Une étrange aventure*.

⁷ More precisely, we exported the Word file into an odt file. We then changed the extension to zip and extracted the content.xml file in this way.

⁸ Thanks to an XSLT stylesheet produced by Nathalie Gasiglia and Luc Audrain.

OpenOffice) is standardised in a common guide and concerns both structural phenomena (e.g. paragraphs, corrections or parts of text that cannot be read) and analytical phenomena (e.g. foreign words, passages relating to selected themes or to quotations). The conception of this guide constitutes a preliminary work of analysis of Elisa Chimenti's texts as a whole, in which several students participated, coordinated by Camilla Cederna and Nathalie Gasiglia. Actually, this guide features the encoding choices. The transcription in this format has the advantage of being able to be created and verified by people unfamiliar with XML-based languages, but it generates several irregularities that render some of the subsequent XSLT transformations partially ineffective. In fact, these transformations required some special arrangements to achieve the desired result. We also had to intervene manually in order to correct certain errors and insert data; therefore, any integration of changes at the transcription level would not be fully automated at present.

Consequently, in terms of workflow reusability, in theory transcribing a text into Word by applying the paragraph and character styles as indicated in the common guide is sufficient. In practice, we can expect specific intervention to be required to get the XML-TEI file and HTML page well-formed and functional.

TABLE I. (NOT EXHAUSTIVE) SCHEME OF THE TRANSFORMATIONS FOR THE FIRST WORKFLOW.

<i>text</i>	<i>TEI</i>
Page number	<pb>
Paragraph	<p>
Erasure/Restored text	<seg>
Adjustments	<choice>,<orig>,<reg>
Gloss	<gloss>
Foreign word	<foreign>
Location	<placeName>
Name or character	<persName>
Significant words	<term>
Word about significant topic	<distinct>
*Part of text about a significant topic	<milestone>
Bibliographic reference	<bibl>
*Reference symbol to external note	<ref>

Starting from the completed and annotated transcription, considering that the common annotation guide had already been created and that some transformation files had already

⁹ The files that enable the XLM file extracted from ODT to be transformed into an XML file in which only the tags corresponding to the styles applied during transcription are present.

¹⁰ For example, the editions produced with EVT.



Fig. 2 Screenshot of the digital edition of *Une étrange aventure*

been provided,⁹ the creation of the interface took approximately one month.

Finally, the interface (fig. 2) is functional but not optimal. It was created independently by adapting numerous templates provided by W3C. Below are some images of the functionalities developed for the digital edition of the novel *Une étrange aventure*. Being a unique, typewritten, ready-to-print copy, for the edition, only some images can be displayed, on a different page from the text. In general, priority was given to the development of functionalities for analysing the text, based on critical notes and annotations of words and passages relating to significant themes.

Since the images of the manuscript are not presented next to the transcribed text, this interface is more similar to commercial digital editions, or other types of texts available online, such as newspaper articles, blogs or fanfictions, which generally present the text in the centre, with links and information on the sides. However, the implemented functionalities are typical of academic productions, in particular the list of terms¹⁰ and the possibility of making information related to the critical apparatus appear or disappear from the body of the text.¹¹

This workflow allows to maximise the customisation of the interface; indeed some improvements have already been done from the first version in order to enhance user experience. However, modifications are not easy to handle, in particular for students and researchers who are not familiar with computer programming. So, even if we decided to start with an easy format for non-computer scientists, we finally obtain a complicated process that does not allow them to directly take charge of the modifications. Hence, this workflow is particularly suited for exploring interfaces and hypertext fitted on the corpus. However, it becomes efficient only if it succeeds in setting up a collaboration between developers and computer scientists on one side, and researchers and students in literature on the other side, within a transdisciplinary frame. Both humanists and computer scientists are deeply interdependent for the realisation of this kind of edition: for this reason, the presence of a specialist in digital humanities is highly desirable.

B. Second workflow

A second type of edition was developed in the context of a Master's degree at the University of Siena.¹² The work, which included preparing the text, the edition and writing the presentation paper, took a couple of months - not full time.

¹¹ For example, the Electronic Beowulf [23].

¹² Bianca Vallarano's dissertation for the Master *Infotext-Infomatica del testo ed edizione elettronica* (2022).

With the assistance of Roberto Rosselli del Turco (University of Turin), the edition is now in the process of being corrected and improved. A broader reflection on workflows and software is also being carried out within the framework of a doctoral thesis aiming to create a digital edition of Elisa Chimenti's collection "La veillée du harem".

The text chosen for the edition was the "Histoire du prince Moustapha et de la princesse Djamila" in its version annotated by Chimenti (fig. 3), contained in the collection "Contes d'autrefois" (n.d.) [1]. A number of variants - albeit not numerous - are present in the typescript of the selected text, which makes the work a matter of authorial philology as well [19]. Another corrected version of this text exists in fine copy, in the collection "La veillée du harem" (n.d.) [2]. This particular case allowed us to reflect on a text whose genetic stages are explicit. We were thus able to focus on how to markup and display the genesis of the text (the erasures, additions, and substitutions) in the visualisation.

In this workflow, XML format, TEI standard, and EVT 2js visualisation tool were used.¹³ In the first instance, the text was manually transcribed from the images of the typewritten pages in TXT format (plain text file). Human transcription was chosen because it was difficult to find open access OCR software with an acceptable text output due to the poor-quality images and damaged text in the typewritten documents. As far as the edition criteria are concerned, since these are preserved originals, there are two options: diplomatic edition and interpretative edition [18].¹⁴ In view of the fact that the chosen text has handwritten notes and corrections, the base text is the corrected text - i.e. taking into account the author's corrections - with notes that account for the genetic aspects of those corrections [14]. Typographical errors have been corrected, punctuation have been normalised and - where necessary - introduced, and capital letters have been added as needed.

The text was copied in an XML file and marked up according to TEI standards. Regarding the encoding choices, TEI was used to mark up the following (more details on Table II). The arborescence of the text: titles of the novels, paragraphs, direct discourses, verses (in the text, prose and

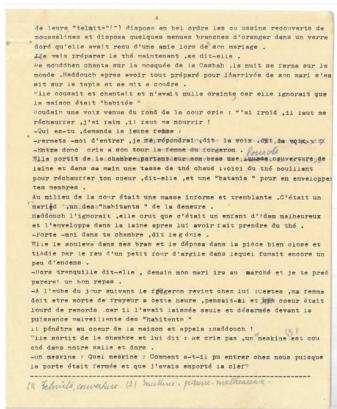


Fig. 3 Page 4 of the typescript of Histoire du prince Moustapha et de la princesse Djamila

¹³ Similar DSEs - built with the same schema, but richer - are the Codice Pelavicino [21] directed by Enca Salvatori and the Digital Vercelli Book [22] directed by Roberto Rosselli del Turco.

¹⁴ As Stussi (1988) writes, a critical edition in the sense of reconstruction of the original is excluded. The diplomatic edition is the faithful reproduction of the original chosen as the base text. The interpretative edition, on the other hand, implies an accounting of authorship variants and a standardisation of

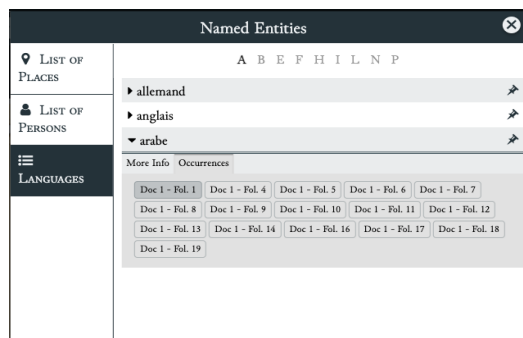


Fig. 4 Exemple of a list of entites of Histoire du prince Moustapha et de la princesse Djamila on EVT

poetry intersect and there are frequent inserts of poems and folk songs). The annotations for exploring the text and create lists of entities: person's names (to navigate within the occurrences of the characters), place's names (to navigate the places cited in the text; this is relevant because Chimenti's texts often focus on exile and travel experiences), foreign languages (as we can see in fig. 4, from the list of languages the user can see the occurrences of each language and click on the folios in which the language is used; this is relevant since Chimenti's texts are multilingual). The apparatuses, encoded using the TEI parallel segmentation method (according to EVT):¹⁵ basic apparatus (additions, deletions, substitutions, author's notes, editor's notes) and critical apparatus (witnesses, lemmas, readings, lacunas).

TABLE II. LIST OF CODING ELEMENTS FOR THE SECOND WORKFLOW.

Titles	<head>
Paragraphs	<p>
Periods	<seg>
Verses	<l>, <lg>
Direct discourse	<said who="">
Personal names	<persName>
Place names	<placeName>
Foreign languages	<foreign xml:lang="">
List of witnesses	<listWit>
Witnesses	<witness xml:id="">
Critical apparatus entry	<app>
Lemma	<lem>
Variant	<rdg>

formal variants (graphic, phonetic, morphological), e.g. the transliteration of Arabic terms into the Latin alphabet.

¹⁵ In addition to being the one supported by EVT, the parallel segmentation method is also the most functional: it is more precise than the location-referenced method and is more intuitive than the double-end-point-attached method. Although the latter is the most effective, it is to date difficult for visualization software to handle.

Lacuna	<lacunaStart wit=""/> <lacunaEnd wit=""/>
Critical notes	<note resp="author"> <note resp="editor">
Addition	<add>
Deletion	
Substitution	<subst>

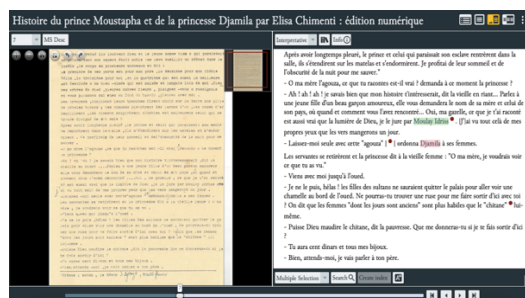


Fig. 5 Screenshot of the digital edition of *Histoire du prince Moustapha et de la princesse Djamilia*.

Concerning the visualisation, EVT needs the XML:TEI file, the configuration file in JSON format, the CSS file to customise the visualisation and the images (in JPG format) associated with the text, for the facsimile edition (if needed). We didn't edit the CSS file. Regarding the JSON file, we just chose the display modes (text-image, reading text and collation), the editions to show (critical and interpretative)¹⁶ and set the lists. An example of the visualisation interface in fig. 5.

This model has a high tax of reusability. The configuration file is set to work at a general level with texts that have author variants (even minor) and/or multiple witnesses (as is the case with most of Chimenti's corpus).¹⁷ The annotations also trace paths that are either fundamental to any basic analysis (people and places) or specific to the entire Chimenti corpus (languages). Once the configuration file is set, the schema for the visualisation is established and you will only need to insert your XML:TEI file.

IV. CRITICAL ASSESSMENTS

A. Critical assessments of the outputs

We can schematise the steps of the two workflows (text structuring, markup, and visualisation) in a comparison table (Table III).

TABLE III. COMPARISON SCHEME FOR THE TWO WORKFLOWS.

	I step	II step	III step
1° workflow	WORD	XML-TEI	HTML
2° workflow	TXT	XML-TEI	EVT

In the end, the two products are both html pages. In both cases, the main page contains the text, which is more or less interactive: in the first experiment, all terms linked to a specific thematic path can light up within the text: characters, places, references, animals, plants, food, religion/magic, foreign languages, glosses, exile/travel, women. Scholar and cultural notes may be read both in a dynamic and in a static visualisation. In the second experiment, genetic variants, names and places can light up within the text. In both experiments there is the possibility of displaying thematic lists, but in the first one they are obviously much richer. In the first experiment, one can switch from the main page via a menu to view the images of the original, the glossary and the transcription criteria. A diplomatic transcription is also

available, where one can visualise structural phenomena, such as corrections. In the second experiment, on the other hand, via the top menu one can switch to the different views of the same text (text-image, reading-text, collation), and to the "toc, front and back" page (where you have information on table of contents, named entities and the project); also via an "option" button you can download the xml, switch languages and learn more about EVT.

Among the two experiments of digital critical editions, the first one is more laborious and requires the availability of a large working group, including not only humanists, but also people with IT skills, especially for writing XSLT transformation sheets and creating the website. Most of all, this process is time-consuming, but also allows for the construction of critical digital editions that are rich in terms of hypertextual paths, thereby leaving the field open for experimentation. Although, in theory, this workflow could only be carried out by a digital humanist, it is actually a work that has a strong collective vocation. The importance of a digital humanist, therefore, lies above all in facilitating communication between experts from different fields.

The second experiment, on the other hand, involves a simpler workflow, which can be realised by digital humanists with transversal skills in humanistic computing and basic knowledge of markup languages, moving along a path already mapped out as far as the programming part is concerned, thanks to the EVT software (which is easy to understand and for which there is an agile online manual). On the other hand, EVT, by making the work simpler, also runs the risk of limiting it, and does not allow the experimentation that is possible on a newly created HTML page - unless you work directly on the code. The EVT edition is probably more appropriate in the case of texts in which there is an important critical apparatus and in which it may be interesting to show images of the originals (e.g. for the *Contes berbères*, rich in manuscript variants).¹⁸ On the other hand, the HTML edition is more suitable if you want to give priority to different hypertextual paths.

B. Overall critical assessments

A number of issues remain open. From the point of view of tools, the problem remains that some of the software used or that could have been used - Word, Oxygen, OCR software - are proprietary and require a licence, which is easy to obtain for Oxygen and Word in the academia, but it may be problematic in other cases. On the other hand, open tools are

¹⁶ As EVT is organised, the 'critical edition' mode shows the final corrected text, while the 'interpretative' mode highlights the genesis (deletions, additions).

¹⁷ Small changes are, however, extremely easy to make. Assuming, for example, that your text does not have multiple witnesses, simply change the 'availableViewModes'/'collation' option from true to false.

¹⁸ The EVT tool is optimal for apparatus visualisation since it was developed specifically for critical editions with multiple levels of apparatuses.

totally available but are often difficult to handle. The issue of transcription is illustrative: in the second workflow it was done manually - or at least, for some portions of text, heavily corrected by the proprietary OCR tools used. In the first workflow, while an open OCR software was used and the failure was mostly due to the image quality, a quite long time was spent on setting and learning how to use the open software. In both cases, the result was a great consumption of time. In addition, while using open software in itself is not enough to ensure the full sustainability of a project, it still increases it, thanks to available documentation about format and use.

A further aspect that remains open is the tag model for marking up aspects of author philology. In the second experiment, the proposed markup combines elements of the standard TEI with elements of the Critical apparatus module in an attempt to render the two levels of criticism and genetics concisely on the text [12]. One of the limitations of these models is that, in terms of graphical rendering, they are closely linked to the final visualisation program. As far as the philological aspect is concerned, however, the TEI markup model may be considered sufficiently detached from the visualisation software.¹⁹

With regard to the time required to create editions, the reflection must consider the three different activities that have to be performed: the conception of the edition, the transcription of the text and the generation of the files that constitute the edition. The time required for the conception of the edition is difficult to assess a priori, because it is experimental work, which also depends on the flexibility and customisation potential of the chosen medium. The time for transcription is also very variable, depending on the content and form of the text, its material conditions and the possible mark-up. In the two workflows analysed in this paper, the former requires more time for transcription and conception than the latter, but quantification is much more difficult since the work was carried out over a long period of time and not on a full-time basis. Instead, we can compare with more certainty the creation time of the files for the edition: once the editorial choices were made and the text was transcribed in plain text format, the creation of the files for EVT took about two weeks. While the time to create the XSLT transformation files and correct the output for the first workflow presented is almost twice as long: about a month. In both cases, this is considered part-time work.

In terms of reusability, the two projects are very different. EVT provides a tool that can be used relatively easily: the experience of creating an edition with this tool can therefore be reproduced by looking at how it is used. Conversely, the first experiment presented can, in theory, be reproduced simply by applying XSLT transformations to the XML files, provided the text is transcribed using the same rules; in practice, the intervention of a computer scientist to correct errors and make the final interface operational is necessary. Furthermore, in this case, the workflow is adapted to Elisa Chimenti's texts, whereas EVT is much more generic.

Concerning the target audience, the first experiment appears more suitable for a broader audience, which may be attracted by a dynamic interface with thematic paths. The second experiment, on the other hand, results in a product that is more useful to the scientific community, being more suited to the construction of complex critical editions with multiple levels of apparatus.

At present, the issue of copyright is still open, since Chimenti's heirs own the rights, and the reflection on how to address it is ongoing. The texts can not be published online yet, but the dialogue with the heirs is open for selected texts, on a case-by-case basis.²⁰ Either way, the texts should be shareable for research purposes. For this reason, two projects are ongoing. On the one hand, an agreement is being sought for State Archives to acquire the archive (in order for it to be well preserved and available for researchers, which at present is not). On the other hand, a website is being built where texts and editions - with annotations, links and thematic paths, result of the work of researchers and students in DH - can be collected, with restricted access, waiting for it to become public.²¹

V. CONCLUSIONS

The models presented are still experimental and should not be regarded as definitive. Nevertheless, it is important that we continue to work and reflect on this in order to achieve a common and shared model for the digital edition of this corpus. For this corpus, it is particularly important to achieve a shared digital publishing model, in several aspects. It facilitates consultation and linguistic research (in a corpus that is so difficult to access and with such a high level of linguistic hybridity), as well as emphasises and interweaves the richness of the thematic paths that the author proposes (not exhaustive): aetiological legends, women's everyday life, folk and oral literature and culture, supernatural, as well as the intersection east/west and north/south, between Maghrebians, Europeans and Berbers traditions, and the three monotheistic religions, each highly represented in the texts with its own heritage of legends and beliefs.

Many other workflows can still be experimented, depending on the context and objectives, for this very rich corpus. For example, we have considered the possibility of using DSLs (Domain-Specific Language) but have not yet made any attempts in this direction. So, we hope, in conclusion, that this paper can be a stimulus to reflect on the topic and trigger a dialogue within the community about the relation between the goal, the context and the technologies involved in a digitisation project.

VI. REFERENCES

A. Primary sources

- [1] E. Chimenti, "Histoire du prince Moustapha et de la princesse Djamila", in E. Chimenti, *Contes d'autrefois*, unpublished, n. d.
- [2] E. Chimenti, "Histoire du prince Moustapha et de la princesse Djamila", in E. Chimenti, *La veillée du harem*, unpublished, n. d.
- [3] E. Chimenti, *Une étrange aventure*, unpublished, n. d.

¹⁹ A software-bound markup was instead used, in the workflow with EVT, to create the lists of entities.

²⁰ We mentioned the edition of poems which is in press for ETS, Pise, as well as a short story, "Khadidja de l'île sarde", both edited by Camilla Cederna. The poems are translated by Camilla Cederna and Antonietta Sanna.

²¹ Unlike in D'Agata's case [9], the documents cannot be consulted separately by readers, so the text must necessarily appear.

B. Secondary sources

- [4] A. Benchebkroun, “Préface”, in E. Chimenti, *Anthologie. Mohammedia/Casablanca: Éditions du Sirocco/Senso Unico Éditions*, 2009, pp. 7-19.
- [5] C. Cederna, “Elisa Chimenti (Naples, 1883-Tanger, 1969): écrivaine et philosophe enragée, entre transgression et médiation” in E. Chaarani, L. Denooz, S. Thiéblemont-Dollet (ed.), *Plein feux sur les femmes (in)visibles*. Nancy: Presses universitaires de Lorraine, 2021, pp. 529-544.
- [6] C. Cederna, “Elisa Chimenti: écrivaine en exil, arabophile et antifasciste”, in *OLTREOCEANO*, n. 20, 2022, pp. 171-185.
- [7] E. Chimenti, *Anthologie, Sirocco/Senso Unico, Mohammedia/Casablanca*, 2009.
- [8] E. Chimenti, Marra, collection of poems translated into Italian by C. Cederna and A. Sanna (in press, ETS, Pisa).
- [9] C. D’Agata, “«Le varianti della rosa». Per un prototipo di edizione digitale del *Nome della rosa*: interpretazione, didattica, annotazione”, in F. Ciraci, G. Miglietta, C. Gatto (ed.), *AIUCD 2022 - Culture digitali. Intersezioni: filosofia, arti, media. Proceedings della 11a conferenza nazionale*. Lecce, 2022, pp. 105-110.
- [10] M. J. Driscoll, E. Pierazzo (ed.), *Digital Scholarly Editing: Theories and Practices*. Cambridge: Open Book Publisher, 2016.
- [11] P. Italia, C. Bonsi (ed.), *Edizioni critiche digitali/Digital Critical Editions. Edizioni a confronto/Comparing Editions*. Roma: Sapienza Università Editrice, 2016.
- [12] P. Italia, “Filologia d’autore digitale”, in *Ecdotica*, n. 1, 2019, pp. 203-216.
- [13] P. Italia, *Editing Duemila*. Roma: Salerno, 2020.
- [14] B. Nava, “Siamo tutti bédieriani? Prospettive per le edizioni genetiche digitali”, in *Umanistica digitale*, n. 14, 2022, pp. 19-40.
- [15] E. Pierazzo, *Digital Scholarly Editing: Theories, Models and Methods*. 2014. fhal-01182162.
- [16] R. Rosselli Del Turco, C. Di Pietro, C. Martignano, “Progettazione e implementazione di nuove funzionalità per EVT 2: lo stato attuale dello sviluppo”, in *Umanistica Digitale*, n. 7, 2019.
- [17] C. Schuwey, *Interfaces. L’apport des humanités numériques à la littérature*. Neuchâtel: Alphil, 2019.
- [18] A. Stussi, *Nuovo avviamento agli studi di filologia italiana*. Bologna: Il Mulino, 1988.
- [19] G. Tancredi, C. Fenu, “XML-TEI: Un modello per la filologia d’autore”, in F. Ciraci, G. Miglietta, C. Gatto (ed.), *AIUCD 2022 - Culture digitali. Intersezioni: filosofia, arti, media. Proceedings della 11a conferenza nazionale*. Lecce, 2022, pp. 218-222.
- [20] M. S. Zemmouri, “À la redécouverte d’Elisa Chimenti, écrivaine humaniste”, in A. Baida (ed.), *Langue française et contacts langagiers*. Rabat: Bouregreg, 2007, pp. 35-146.

C. Sitography

- [21] Codice Pelavicino: <http://pelavicino.labcd.unipi.it/evt/>
- [22] Digital Vercelli Book: <http://vbd.humnet.unipi.it/beta2/index.html>
- [23] Electronic Beowulf: <https://ebeowulf.uky.edu/ebeo4.0/>
- [24] EMAN Édition de Manuscrits et d’Archives Numériques: <https://eman-archives.org/EMAN/>.
- [25] EVT Edition Visualization Technology: <http://evt.labcd.unipi.it/>.
- [26] FMEC - <https://www.elisachimenti.org/accueildef.html>
- [27] GitLab: <https://about.gitlab.com/>.
- [28] TEI Text Encoding Initiative: <https://tei-c.org/>.
- [29] Tesseract OCR on Github: <https://github.com/tesseract-ocr/tesseract>.
- [30] THALIM Unité mixte de recherche 7172 “Théorie et histoire des arts et des littératures de la modernité XIX-XXI siècles”: <https://www.thalim.cnrs.fr/>.