Versione PROVVISORIA del contributo presentato al Convegno Annuale

Sharing knowledge digitally; the Muruca case study

Chiara Aiola[1], Giulio Andreini[2], Francesca Di Donato[3], Tiziana Lombardo[4]

[1] Net7, Italy - aiola(«»)netseven.it
[2] Net7, Italy - andreini(«»)netseven.it
[3] ILC-CNR, Italy - francesca.didonato(«»)ilc.cnr.it
[4] Net7, Italy - lombardo(«»)netseven.it

ABSTRACT

In this poster we present the evolution of *Muruca*[1], a platform that has been conceived as a framework to allow Digital Humanities researchers and research teams to create, curate and share their own Digital Editions. The poster describes the main features of the framework, its evolution over the years in synergy with the researchers' needs and with the progress of technologies and policies, and the open challenges and opportunities of the framework in a European context.

KEYWORDS

Open Culture, Digital Humanities, Digital Libraries, Digital Editions, Electronic Objects.

1. INTRODUCTION

In the Digital Humanities, there is a growing demand for software suites that allow users to manage the Digital scholarly Edition[2] of a corpus of digital documents (texts, images, metadata), from the management phase (entering texts and metadata, semantic enrichment, management of editorial workflows), to the publication on the Web.

This demand is indicative of an expanding market, characterized by a high number of small individual projects, normally with very high scientific content but often with an insufficient level of optimization and pooling of software resources and technical and editorial procedures [3]. In this context, the collaboration between academic and technical teams is at the centre of the Digital Edition publication process.

2. MURUCA'S WORLD

Muruca is a framework composed of different components. In its macroscopic structure we can divide it into 2 parts:

- The backend. It's the platform used to insert and manage data. It allows to create "records" for each work with its metadata, like in physical libraries, but also to associate transcriptions and images or other multimedia. The type of data relay on each project's object of study.
- The frontend. It's the interface used by the users for consultations. It gives tools to consult the materials, such as textual search, search of metadata, image viewer and text reader. It should be a tool for researchers, but also a way to disseminate the results of the research itself.

Its evolution took place in line with its research ecosystem and rests upon these three main pillars:

- Easiness of use. Technology shouldn't be an obstacle for users, but a useful work tool
- Standardization of the framework. It is a choice of maintainability and sustainability
- Split between content and its design. The contents of a publication are the real value of the project. The results of the researches had to be maintained regardless of the evolution of the digital platform.

Data model and data preservation

After an analysis of the whole scenario of available content management systems, it has been decided to build the backend on the popular CMS *Wordpress*[3] by developing a group of plugins, one for each functionality, which allows the creation of as many content types as we need with all their related metadata. This choice responds to the need of a flexible data model definition coupled with the need of having a usable interface to manage all these entities. Furthermore, attention to preservation issues has been given by adding a functionality that allows to export data into files that can be stored everywhere, such as in a Zenodo[4] record or into thematic databases' aggragators .

---

[1] http://www.muruca.org/

[2] While providing a complete definition of the expression Digital scholarly Edition is out of the scope of this presentation, we define it here as a critical representation of texts and historic documents, created, shared, and published on the Web. See [1], [2]

[3] https://wordpress.com/ was born in 2003 as a simple blogging platform, today in 2020 (with version 5.5) it covers the 44% of all the Internet sites using a CMS.

[4] "ZENODO is a CERN Data Centre-backed research data repository for the long-tail of science, enabling researchers to preserve and share their research output from any science, regardless of the size and format." (from

Data visualization

The Muruca's frontend is the result of many years of experience and iterations with Digital Humanities research teams; its design and development took into account also case studies belonging to different domains like data visualisation and digital archiving to offer the possibility of organizing all the elements to obtain a scientifically valid content visualization and a usable interface for selected typologies of users

Sustainability and reusability in the long term

Muruca is designed to be a modular platform, in which there are a number of components available to build each project, but each component is the same (with a different configuration) for all Muruca installations.This solution makes it easy to create a new instance of Muruca or update all its installations.

The frontend uses external tools and libraries to implement some functionalities; the choice of which of them have to be included in the framework is based on those that are widely used and accepted as a standard or developed by a large community. The choice to adopt TEI Publisher[5] as the text reader in the new version of Muruca is a step forward in this direction.

This way the platform is conceived as a modular object where the sustainability of each single object is meticulously planned.

3. PERSPECTIVES AND FUTURE EVOLUTIONS

Despite we present Muruca as a virtuous example, the fast changing scenario in the Digital Humanities opens to new challenges and issues to deal with.

The first tension is between the so-called *pret-a-porter* and tailor made solutions [3]. Digital scholarly Editions are offered from many different platforms, which are very often designed and developed for one specific project and are mostly produced for it by research teams and specialised developers. The consequences of this situation are manyfold: firstly, the costs of the development are very high and are not scalable; secondly, despite the fact that most editions use the TEI format for their data model, software standardization is far from being reached. The fragmentation and (unpredictable) discontinuity of funding was the main limitation we had to face when building a full "pret-a-porter" solution which, on the contrary, needs a planned long-term strategy and financial support.

Moreover, there is a lack of harmonization in the specifications of users' needs - which depends on the fact that there's not a full agreement on common standards within the community.

In order to address these challenges, we are planning a new design phase in Muruca, centered around the development of open science methodologies and standards. Muruca’s evolution will be focused on FAIR data, reproducible processes, open access and long-term sustainability, in order to conceive a way to measure the scholarly edition impact and outreach, taking into account multiple metrics and indicators and, at the same time, considering standardization and preservation. All these aspects will be investigated thanks to a collaboration between the authors of this paper and other potential European partners experts on this domain.

BIBLIOGRAPHY

[1] Gabler, Hans Walter. «Theorizing the Digital Scholarly Edition». Literature Compass 7/2 (2010): 43–56. 10.1111/j.1741-4113.2009.00675.x

[2] Sahle, Peter. «What is a Scholarly Digital Edition?». (https://books.openedition.org/obp/3397?lang=it)

[3] Pierazzo, Elena. «What future for digital scholarly editions? From Haute Couture to Prêt-à-Porter». International Journal of Digital Humanities (2019) 1:209–220. https://doi.org/10.1007/s42803-019-00019-3.

[4] Griffin Gabriele, Hayler Matt Steven «Collaboration in Digital Humanities Research – Persisting Silences». Digital Humanities Q 2018 volume 12 nr 1

[5] Mancinelli, Pierazzo «Che cos’è un’edizione scientifica digitale». Carocci editore, 2020.

[6] Gomes Daniel, Mario J. Silva «Modelling information persistence on the web». Proceedings of the 6th international conference on Web engineering, July 2006

https://zenodo.org/record/51902#.X1czg3kzZEY). The platform also gives a Rest API Service to publish content used by Muruca. Documentation on https://developers.zenodo.org

[5] https://teipublisher.com/index.html is a text viewer for XML-TEI encoded text developed by the community and licensed under the GPLv3.