*e*Viterbo global report

Outlines on the creation of a research open platform using MediaWiki and Wikibase.

Alice Santiago Faria

Rute Correia

This document intends to give an overview of the organization of the *e*Viterbo platform system. It describes the choices and challenges behind its creation, the outline of the infrastructure and the data model, reflecting on the main challenges of creating and maintaining such a platform in a social sciences and humanities academic environment. Some of its content contains lessons learned in practice, doubts, and open questions; in this way tries to achieve to be a best practices reference for future projects.

*"e*Viterbo global report" was produced under the scope of the research project TechNetEMPIRE - *Technoscientific Networks in the construction of the built environment in the Portuguese Empire (1647-1871)*, supported by FCT - Fundação para Ciência e Tecnologia (PTDC/ART-DAQ/31959/2017) and Wikimedia in Residence at FCSH - Faculdade de Ciências Sociais e Humanas, Universidade NOVA de Lisboa - contract with the support of Wikimedia Portugal, a Wikimedia Foundation national chapter. Because the technical infrastructure and content of the platform will evolve, so will this report, and future versions of this report are to be expected.

DOI:10.5281/zenodo.7380115

Alice Santiago Faria

Auxiliary researcher at CHAM - Centre for the Humanities, FCSH, Universidade NOVA de Lisboa

Rute Correia

Wikimedia in Residence at FCSH

First Version: November 2022

Index

*e*Viterbo?

*e*Viterbo is the main research output of TechNetEMPIRE, an FCT-funded research project looking into the technical and scientific networks of the built environment (engineering, architecture, cartographers, carpenters, masons, etc.) in the Portuguese Colonial Empire. The project is led by researchers Alice Santiago Faria and Renata Malcher de Araújo, both affiliated to CHAM - Centro de Humanidades (a FCSH, Universidade NOVA de Lisboa and Universidade dos Açores joint research unit), and also had the support of lab2pt at Universidade do Minho and IEETA of Universidade de Aveiro.

The platform was called eViterbo, as a tribute to Francisco Viterbo, the author of one of the main 19th century sources in which the project is rooted. Like in the life work of Francisco Viterbo, which on top of its very known *Dictionary of Architects and Engineers* studied other arts & crafts portuguese agents and aspects, eViterbo contains content from other art history-research projects from the "Art, History and Memory" research group within CHAM. Among them C.1892, the individual research project developed by the PI, that enlarges *e*Viterbo chronology, and content created under an exploratory project called "As palavras com que nos cosemos...", financed by CHAM, and partially developed under the DRESS|VESTE research projects, led by Carla Alferes Pinto. The inclusion of several areas of knowledge from within the "Art, History and Memory" research group, makes the future of *e*Viterbo possible in several areas and a prime tool for Art History and related fields.

eViterbo is a multifaceted platform that combines a MediaWiki-based encyclopedia with an open linked data structured database, built on a Wikibase suite. It is a pioneer platform combining these two parts into a single platform in a research setting.

Why Wikimedia free and open source software tools in academia?

The principle behind TechNetEMPIRE was to gather scattered knowledge, bringing together researchers working in different geographical spaces and chronologies of the former Portuguese empire. This means gathering data fragmented across academic frameworks, sharing it among peers, identifying information gaps, trying to complete data, co-edit it and re-share it with the general public. With the broad aim of updating two 19th-century sources (both of which had the form of an encyclopedia/dictionary) a digital encyclopedia was created, to allow crossed and networked readings, built over the data organization provided by such sources.

In the past twenty years, Wikipedia stands out as a collaborative encyclopedia. Launched in 2001, it is the largest and most widely used encyclopedia in the world. Available in more than 280 languages, the project aims to be the “sum of all human knowledge”. Wikipedia may be the most well-known Wikimedia platform, but there are many others, each one focused on a specific type of knowledge or approach: for instance, Wikimedia Commons serves as a multimedia archive, Wikisource compiles transcriptions of texts and books, Wikispecies is a species directory, and many others. All these are built collaboratively by thousands of volunteers around the world. Thus, they embody the definition of a collaborative environment, where it is simple to create, share and co-edit content. Every change is tracked and can be reversed, ensuring that any edit can be easily checked and reinstated - creating a true version control system for content. The software used to build these platforms is called MediaWiki. Its familiar interface makes it a great option for encyclopedia-style solutions, such as *e*Viterbo, as most people will recognize the layout, facilitating navigation.

MediaWiki is a software developed by the Wikimedia Foundation (WMF); it is made available as free and open source software. It is thought of as an environment for the co-creation of content: it allows easy co-editing, discussion and following up of each stage of the process. Its open codebase grants users with programming knowledge the possibility to create new features, extensions or code libraries to interact with the software in an improved manner, making it particularly flexible. This customization is often shared: many of these extras are redistributed, keeping the code open to ensure that the software will meet the users' needs.

The financial costs for running it are marginal compared to software as a service-solutions (SaaS) or other types of proprietary software. Unlike these types of software, it is free to install, run and modify by any number of users without the requirement of any additional licenses or subscription fees.

It is also worth noting that the hardware requirements to host a MediaWiki installation are very low: only 256MB of RAM and 85 MB of storage are required to start an instance. For a public university in Portugal, meeting these demands should be straightforward. MediaWiki also allows private environments with several levels of access, from "anyone can read or edit" to a closed system with limited access. This setting can be changed at any time (however this may cause some other problems, see: Main Challenges). This means, for example, that it is possible to work on a private project with a closed team for a given amount of time and, once the content is ready, the same wiki can be made publicly available.

Last but not least: by using open and interoperable file formats, it also guarantees the longevity and portability of any outputs, in this case research ones, beyond their native software habitat.

As academia slowly moves forward with open science (open outputs, open data, etc.) and with the evidence that many tailored software solutions in academia are going obsolete, sometimes passing a date of no return, the longevity and portability of any digital outputs were problems seriously considered by TechNetEMPIRE coordinators. MediaWiki, along with other Wikimedia software tools, seemed to ensure those conditions in the project afterlife.

TechNetEMPIRE project application to FCT funding stated that

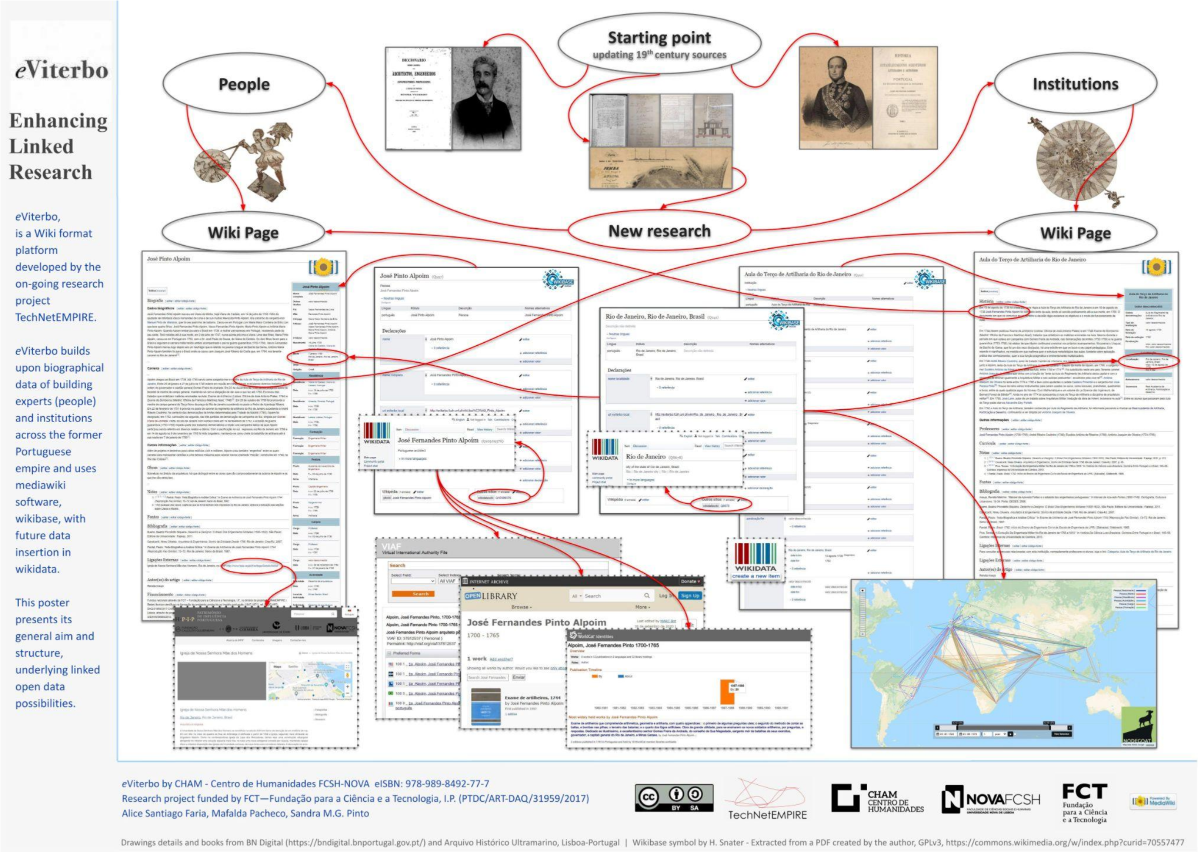
> “a set of structured data will be created - curated database (T3) - whose future free use will be guaranteed through the connection of the platform to DBpedia (http://pt.dpedia.org). One of the outputs of the project will be the open access platform, which if fruit of multidisciplinary work, bringing together historians and computer engineers. The platform will have the external form of a dictionary of people and institutions that will allow others to use and explore the methodological approach, the data, the sources used, the visualizations and results in the analysis, thus establishing a replicable analysis model for other contexts and chronologies.”

Findable, Accessible, Interoperable, Reusable (FAIR) data was always a central point of the project. MediaWiki through the hypothesis of structuring data via the infoboxes and linking possibilities with DBpedia seemed as a safe, and full of opportunities, place to go.

The infrastructure. Past & Present

Project applications, and the bureaucracy and time needed to get a research project running, sometimes seem incompatible with the timings of information, computer and software technology.

TechNetEMPIRE application was submitted in May 2017, and the results were known in July 2018. The project officially started in October 2018 and work started in January 2019, but due to general bureaucratic difficulties, the team was only completed in April 2019. Not long after its beginning, it became clear that the software options proposed on the application had to be reviewed. By the beginning of October 2019 decisions were made to install Wikibase as a way to deal with structure data and linking to Wikidata.



eViterbo

Enhancing Linked Research

*e*Viterbo, is a Wiki format platform developed by the on-going research project TechNetEMPIRE.

*e*Viterbo builds upon biographical data of building experts (people) and institutions across the former Portuguese empire and uses mediawiki software, wikibase, with future data insertion in wikidata.

This poster presents its general aim and structure, underlying linked open data possibilities.

*e*Viterbo by CHAM - Centro de Humanidades FCSH-NOVA eISBN: 978-989-8492-77-7
Research project funded by FCT—Fundação para a Ciência e a Tecnologia, I.P. (PTDC/ART-DAQ/31959/2017)
Alice Santiago Faria, Mafalda Pacheco, Sandra M.G. Pinto

TechNetEMPIRE
CHAM CENTRO DE HUMANIDADES
NOVAFCSH
FCT Fundação para a Ciência e a Tecnologia

Drawings details and books from BN Digital (https://bndigital.bnportugal.gov.pt/) and Arquivo Histórico Ultramarino, Lisboa-Portugal | Wikibase symbol by H. Snater - Extracted from a PDF created by the author, GPLv3, https://commons.wikimedia.org/w/index.php?curid=70557477

Image 1 - *e*Viterbo. *Enhancing Linked Research*. Poster presented at Linked Pasts VIII Symposium, 2021

Linked data possibilities went beyond Wikidata and the possibility of linking the encyclopedia type of pages from *e*Viterbo to other related webpages using hyperlinks was also considered. Image 1 underlines the platform general aim and structure. From the starting point of updating, in terms of methods and content, fundamental 19th century and its possibilities in terms of Linked Open Data (LOD). Moreover, the combination of MediaWiki-based encyclopedia style platform with a Wikibase-powered database, and the possibility of connection to external software for visualization and analysis of data are also represented.

MediaWiki, Wikibase and the Semantic Web

Released as free and open source software (FOSS), MediaWiki is the backbone of thousands of web-encyclopedias worldwide, helping organizations and people compile knowledge in our digital-oriented society. The software was first released in 2002, followed by multiple versions and upgrades since then. It is still actively developed by the WMF engineering teams, with significant input from multiple stakeholders within the wider Wikimedia community, from MediaWiki volunteer developers to Wikipedia editors, among others. With such longevity and frequent update cycles, the software is considered stable, but specific functionalities often require installing additional extensions. Such extensions generally have their own requirements and compatibility issues, increasing the complexity of the project, as different pieces of software interact with each other creating an ever-growing jigsaw puzzle of small intricacies. This will be further discussed later in this report.

There probably are thousands of small Mediawiki encyclopedias around the world but a close example to *e*Viterbo in the use of Mediawiki would be Bestor - Belgian Science and Technology Online Resources, which was implemented in 2008, as an "encyclopedia for Belgian science".[1] Nevertheless Bestor does not use structured data and its "database" actually runs by aggregation, using categories.

[1] https://www.bestor.be/wiki/index.php/Bestor:Pr%C3%A9sentation

Wikibase is the engine behind Wikidata, Wikimedia's linked open data platform. It was launched in 2012, developed by Wikimedia Deutschland (WMDE), the Germany-based subsidiary of WMF. Wikibase is actually a collection of MediaWiki extensions that allow working “with versioned semi-structured data in a central repository (...) instead of the unstructured data of MediaWiki wikitext”.[2] This means that, instead of having pages with text-based content, such as those in an encyclopedia-type MediaWiki, Wikibase enables users to create databases of structured linked data that can then be used to create graphs of knowledge, connecting various objects, and facilitating an integration with the “semantic web*”.*

Coined in 1999 by the World Wide Web’s father, Tim Berners-Lee, the term “semantic web” points towards the idea of a web of data, an Internet that can be read by machines.[3] As technology evolved, semantic web became possible through a variety of file formats and protocols, such as the Resource Description Framework (RDF) and Web Ontology Language (OWL), that encode semantics into data. Thus, these technologies empower the alignment of ontologies, taxonomies and data interchange (see Image 2) across databases with different purposes and organization systems.

[2] https://en.wikipedia.org/wiki/Wikibase

[3] Berners-Lee e Fischetti, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*.

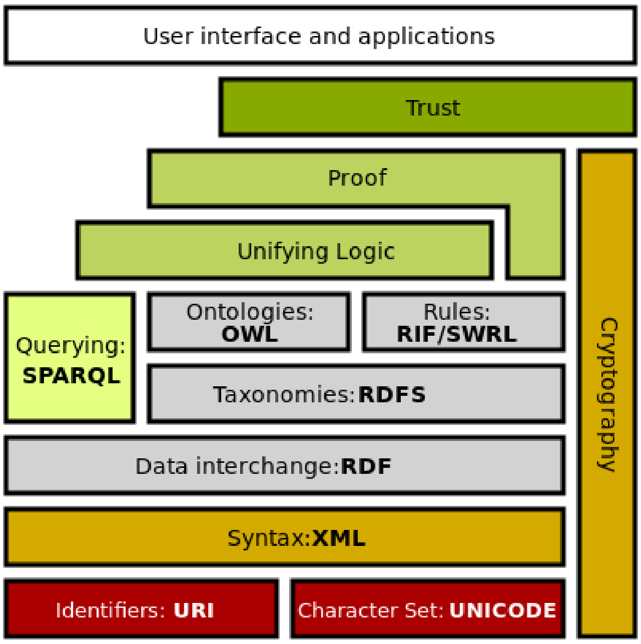


Image 2 - *Semantic Web Stack*, by Marobi1 (CC0)

Without diving into the engineering behind it, Wikibase uses the afore-mentioned semantic web stack to store and access data. It is worth noting that Wikibase is not the only way of addressing the connection between MediaWiki and the semantic web. Semantic MediaWiki and Cargo are the two other possibilities; like Wikibase they are MediaWiki extensions. Each one of these offers its own set of advantages and potential setbacks: for instance, Cargo is easy to install, but has stability issues; Semantic MediaWiki (SMW) has a flexible data model, but its learning curve for system administrators is quite high;[4] Wikibase provides a visual editor for structured data but, while adoption is widely increasing, documentation is fairly scarce and maintenance can be hard.

Cargo vs. Wikibase - Managing data on eViterbo

Cargo was the initial choice to manage data on *e*Viterbo. This extension launched in 2015 and was designed as a lightweight version of SMW. As such, Cargo tends to be

[4] See among others: Daniel K. Schneider & Julien Da Costa (2013). *Adding power to educational and research wikis with Semantic MediaWiki*, Paper presented at the 8th conference on Semantic MediaWiki (SWMCon 2013), available at: https://edutechwiki.unige.ch/en/Adding_power_to_educational_and_research_wikis_with_Semantic_MediaWiki

simpler to use than its older sibling SMW. It stores data in templates, which are presets that facilitate the transclusion of text and/or data across MediaWiki pages. By calling these templates within a page, they will display the requested data. For experienced Wikipedia editors, this is a very familiar way of accessing information since templates are included on Wikipedia, but it is arguable whether this is easier to use than SMW or Wikibase without this prior knowledge. Cargo also supports inline queries, which can be particularly useful to visualize information.

As the main goal behind building *e*Viterbo under TechNetEMPIRE was to build a database to explore the connections between multiple agents - persons, institutions and locations -, Cargo's simplistic data model proved limited, since it provided no real way to curate data, meaning to "fool-proof" the data insertion process (typos, capitalization inconsistencies, etc.). The solution for this was to install Wikibase. However, this also has not yet lived up to expectations as it will be discussed in section "Constraining properties". It should be stressed that both solutions seemed adequate to the incompleteness of historical data.

DBpedia vs. Wikidata

DBpedia and Wikidata are both Linked Open Data resources created around Wikipedia and Wikimedia universe. The main difference between both is that DBpedia extracts data from Wikipedia while Wikidata can provide data to Wikipedia and its use for example by Google's Knowledge Graph. In fact, they are seen by many as complementary tools.[5] Another difference is that Dbpedia, like Wikipedia, runs on "notability factor" similarly to Wikipedia and Wikidata has very low notability rules. For TechNetEmpire research this was an important point since its technical networks can include for example "carpenters" or other technical staff which might not be considered "notable" enough by the Wikimedia community of volunteers.

[5] See among others: Kingsley Uyi Idehen, «On the Mutually Beneficial Nature of DBpedia and Wikidata», OpenLink Virtuoso Weblog, 11 de fevereiro de 2017, https://medium.com/virtuoso-blog/on-the-mutually-beneficial-nature-of-dbpedia-and-wikidata-5fb2b9f22ada; D. Abián et al., «Wikidata and DBpedia: A Comparative Study», em Semantic Keyword-Based Search on Structured Data Sources, ed. Julian Szymański e Yannis Velegrakis, Lecture Notes in Computer Science (Cham: Springer International Publishing, 2018), 142–54, https://doi.org/10.1007/978-3-319-74497-1_14.

Wikidata is inclusive to all data and at the moment it is used as a main tool to mediate different ontologies.

Images 3 and 4 represent the “path of data" from the archive to the database and the difference between ”Cargo vs. Wikibase” and“DBpedia vs. Wikidata”.   In both images, presented by the PI in team work meetings and later at WikiDataCon2021, is clear an uncomfortable mystery of what happened to data once in “the data transformer” and how it could be extracted for data analysis. Wikibase presented itself as a smooth and clear way of data curation, modeling and extraction.

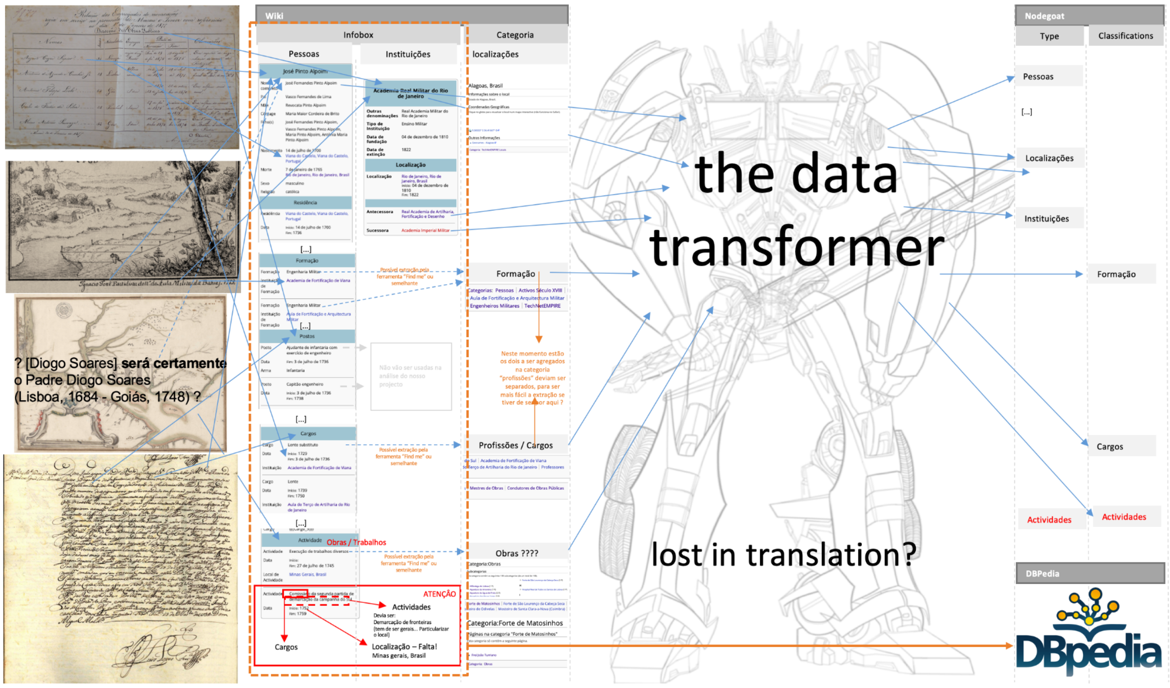

? [Diogo Soares] será certamente
o Padre Diogo Soares
(Lisboa, 1684 - Goiás, 1748) ?
Wiki
Infobox
Pessoas
José Pinto Alpoim
Nome completo José Fernandes Pinto Alpoim
Pai Vasco Fernandes de Lima
Mãe Revocata Pinto Alpoim
Cônjuge Maria Maior Cordeira de Brito
Filho(s) José Fernandes Pinto Alpoim, Vasco Fernandes Pinto Alpoim, Maria Pinto Alpoim, Antónia Maria Pinto Alpoim
Nascimento 14 de julho de 1700 Viana do Castelo, Viana do Castelo, Portugal
Morte 7 de janeiro de 1765 Rio de Janeiro, Rio de Janeiro, Brasil
Sexo masculino
Religião católica
Residência
Residência Viana do Castelo, Viana do Castelo, Portugal
Data Início: 14 de julho de 1700 Fim: 1736
[...]
Formação
Formação Engenharia Militar
Instituição de Formação Academia de Fortificação de Viana
Formação Engenharia Militar
Instituição de Formação Aula de Fortificação e Arquitectura Militar
[...]
Postos
Posto Ajudante de infantaria com exercício de engenheiro
Data Fim: 3 de julho de 1736
Arma Infantaria
Posto Capitão engenheiro
Data Início: 3 de julho de 1736 Fim: 1738
[...]
Cargos
Cargo Lente substituto
Data Início: 1729 Fim: 3 de julho de 1736
Instituição Academia de Fortificação de Viana
Cargo Lente
Data Início: 1739 Fim: 1750
Instituição Aula do Terço de Artilharia do Rio de Janeiro
[...]
Actividade
Obras / Trabalhos
Actividade Execução de trabalhos diversos
Data Início: Fim: 27 de julho de 1745
Local de Actividade Minas Gerais, Brasil
Actividade Comissário da segunda partida de demarcação da campanha do Sul
Data Início: 1752 Fim: 1759
ATENÇÃO
Actividades
Devia ser:
Demarcação de fronteiras
(tem de ser gerais... Particularizar o local)
Cargos
Localização – Falta!
Minas gerais, Brasil
Instituições
Academia Real Militar do Rio de Janeiro
Outras denominações Real Academia Militar do Rio de Janeiro
Tipo de Instituição Ensino Militar
Data de fundação 04 de dezembro de 1810
Data de extinção 1822
Localização
Localização Rio de Janeiro, Rio de Janeiro, Brasil Início: 04 de dezembro de 1810 Fim: 1822
Antecessora Real Academia de Artilharia, Fortificação e Desenho
Sucessora Academia Imperial Militar
Possível extração pela ferramenta “Find me” ou semelhante
Não vão ser usadas na análise do nosso projecto
Categoria
localizações
Alagoas, Brasil
Informações sobre o local
Coordenadas Geográficas
Outras Informações
Formação
Categorias: Pessoas | Activos Século XVIII | Aula de Fortificação e Arquitectura Militar | Engenheiros Militares | TechNetEMPIRE
Neste momento estão os dois a ser agregados na categoria “profissões” deviam ser separados, para ser mais fácil a extração se tiver de ser?por aqui ?
Profissões / Cargos
do Sul | Academia de Fortificação de Viana
do Terço de Artilharia do Rio de Janeiro | Professores
Mestres de Obras | Condutores de Obras Públicas
Obras ????
Categoria:Obras
Subcategorias
Forte de Matosinhos | Forte de São Lourenço da Cabeça Seca
Mosteiro de Odivelas | Mosteiro de Santa Clara-a-Nova (Coimbra)
Categoria:Forte de Matosinhos
Páginas na categoria "Forte de Matosinhos"
Esta categoria só contém a seguinte página.
Frei João Turriano
Categoria: Obras
the data transformer
lost in translation?
Nodegoat
Type
Classifications
Pessoas
[...]
Localizações
Instituições
Formação
Cargos
Actividades
Actividades
DBPedia
DBpedia

Image 3 - *Text from wiki to Nodegoat and Dbpedia. eViterbo Plan 2018*
in *e*Viterbo: Implementing an open platform in academia. Wikidatacon2021

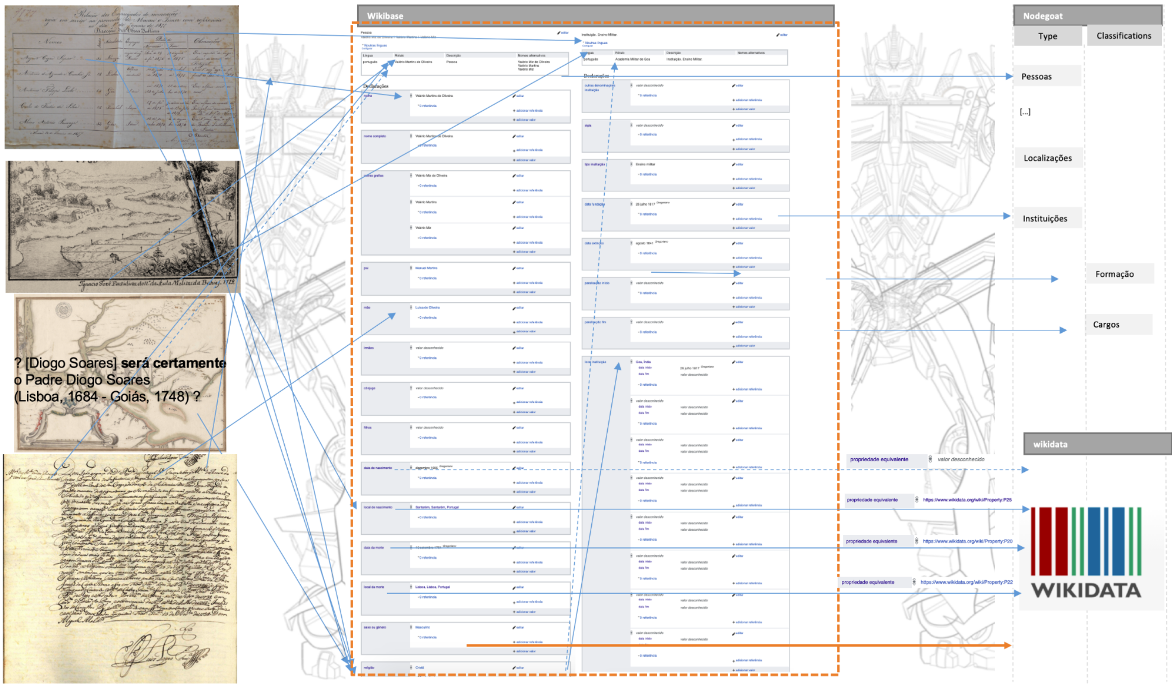


Image 4 - *Text from wiki to Wikidata (to Nodegoat and more. eViterbo Plan end of 2019* in eViterbo: Implementing an open platform in academia. Wikidatacon2021

Wikibase State of the Art

Wikibase has received increasing interest from the wider public. Powering the world's largest open knowledge database is a business card like no other: Wikidata is itself a Wikibase instance. Since its inception, hundreds of organizations have incorporated Wikibase in their infrastructure or stacks, be it for single projects or wider integrations. Wikidata has been used extensively by multiple national libraries around the world, such as the Library of Congress (USA), Bibliothèque Nationale de France (France), Biblioteca Nacional de Portugal (Portugal) or Deutsche Nationalbibliothek (Germany), among others. Similarly, the Online Computer Library Center (OCLC), a "global library organization that provides shared technology services", has made significant advances into testing the usage of Wikibase to create linked data in a library context on their Project Passage[6].

[6] Jean Godby et al., Creating Library Linked Data with Wikibase: Lessons Learned from Project Passage, 2019, https://doi.org/10.25333/faq3-ax08.

Libraries are not the only type of institutions to embrace these solutions. Europeana, a pan-European virtual library and digitalized cultural artefacts aggregator, has been using both Wikidata and self-hosted Wikibase environments on their workflows. Wikidata has been used to both reconcile ontologies across languages and institutions, as well as enrich the metadata associated with their collections[7]. Wikibase, on the other hand, is the engine behind EAGLE, the Europeana network of Ancient Greek and Latin Epigraphy. Wikibase was used to catalogue over 1.5 million items scattered between 25 countries, in a partnership that involved Wikimedia Italia.

Beyond cultural heritage projects, the European Union has been using Wikibase as the infrastructure powering the EU Knowledge Graph. As explained by D. Diefenbach et al, "this graph contains heterogeneous data items such as countries, buildings, and projects funded by the European Union", which encompasses thousands of data points.[8]

As far as academic research projects in History go, Wikibase powers significant achievements in data collection and exchange, such as FactGrid and Enslaved. Labeled as a "database for historians", FactGrid is operated by the University of Erfurt, but maintained with the primordial help of WMDE and the Deutsche Nationalbibliothek, also gathering support from other European universities and research institutes, such as Universitat de Barcelona (Spain) and the Institut d'histoire moderne et contemporaine (IHMC) in Paris, among others. In November 2022 FactGrid just became part of the Nationale Forschungsdateninfrastruktur (NFDI), the German National Research Data infrastructure[9]. Enslaved uses Wikibase to reconcile more than two dozen datasets in order to map the lives of people related to historical slave trade. It is maintained by the Matrix, a research unit at Michigan State University (USA), with the support of the University of Maryland, the University of California Riverside and other local institutions.

[7] https://pro.europeana.eu/page/get-your-vocabularies-in-wikidata

[8] See: Diefenbach, Dennis, Max De Wilde, e Samantha Alipio. «Wikibase as an Infrastructure for Knowledge Graphs: The EU Knowledge Graph». Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-88361-4_37.

[9] See: https://blog.factgrid.de/archives/3104

To the best of the authors' knowledge, eViterbo is a pioneer project in terms of using a structured database with an encyclopedia, simultaneously, in a production environment. Typically, Wikibase installations tend to exist as stand-alone instances, not tied to a MediaWiki encyclopedia counterpart. The only other project that we could find that proposed a similar approach was DeJonge Wiki. This Arenberg Castle-focused MediaWiki plus Wikibase instance was set up as a prototype to explore the viability of using such a solution for building scientific databases. As of November 2022, the project is still in development. Since the subject of research is so granular, arguably the complexity of the project is far smaller than eViterbo's. And while eViterbo explores the same idea - whether a combination of MediaWiki and Wikibase can be used for research purposes - the initial goal was simply to use these software tools as a way to build a database that could enhance the connections between the multiple agents. Thus, the challenges and concerns of such a setup may differ greatly from those experienced by the DeJonge Wiki, as eViterbo was not designed nor developed with the purpose of testing the software. Instead, it was a de facto application of the software's capabilities in a research setting.

eViterbo's system description

As previously mentioned, eViterbo is an encyclopedia that contains text-based entries with a structured data counterpart about people, professions, places and institutions.

Most of the information on eViterbo can be divided into one of the following groups of pages: content, data, navigation and system. All these have their own meta-information: "Categories" are present on all types of pages except "data pages", aggregating pages by topic but also system-wise, as there some of them highlight technical specificities, for example. eViterbo's ecosystem is deeply connected within itself, through hyperlinks (or “interwiki links”, as they mostly point to different pages on eViterbo).

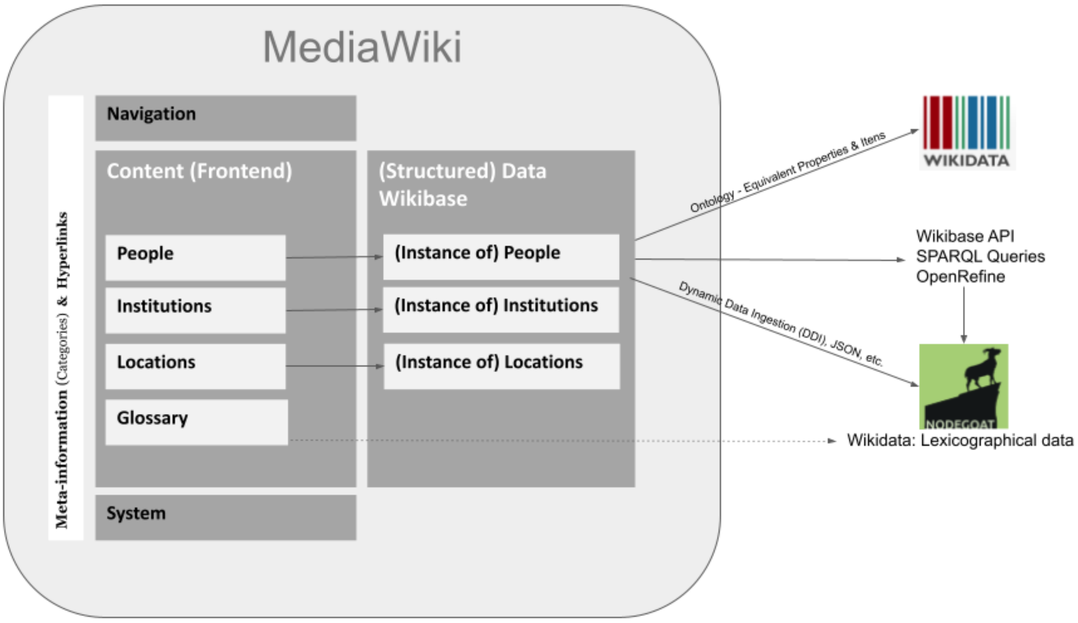
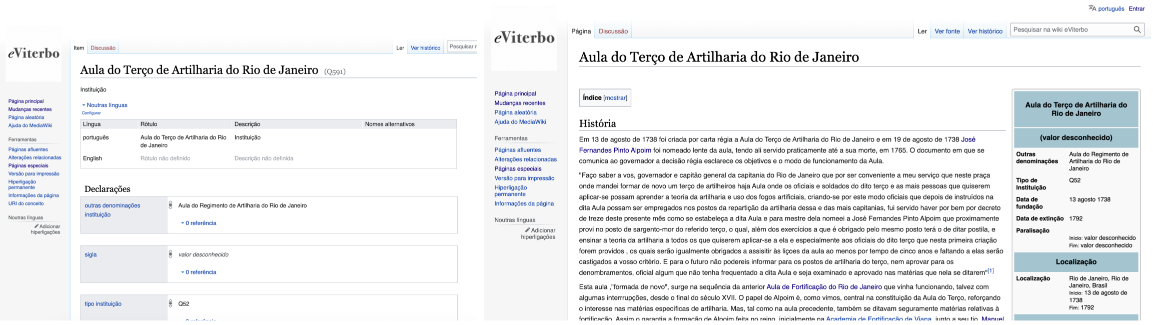

MediaWiki
Meta-information (Categories) & Hyperlinks
Navigation
Content (Frontend)
People
Institutions
Locations
Glossary
(Structured) Data Wikibase
(Instance of) People
(Instance of) Institutions
(Instance of) Locations
System
Ontology - Equivalent Properties & Itens
WIKIDATA
Wikibase API
SPARQL Queries
OpenRefine
Dynamic Data Ingestion (DDI), JSON, etc.
NODEGOAT
Wikidata: Lexicographical data

Image 5 - Wiki eViterbo general model

By ‘content’, we mean the encyclopedia-like entries; by ‘data’ we mean the structured data items (Wikibase items) associated with each one of the content entries or to its related data. Thus, most entries in eViterbo contains two pages: the content one and the data one, both linked bilaterally. There is also a Glossary, which will be addressed later in this report (see: eViterbo content pages).



Aula do Terço de Artilharia do Rio de Janeiro (Q591)

Instituição

Língua	Rótulo	Descrição	Nomes alternativos
português	Aula do Terço de Artilharia do Rio de Janeiro	Instituição	
English	Rótulo não definido	Descrição não definida	

Declarações

outras denominações instituição	Aula do Regimento de Artilharia do Rio de Janeiro 0 referência
sigla	*valor desconhecido* 0 referência
tipo instituição	Q52

Aula do Terço de Artilharia do Rio de Janeiro

História

Em 13 de agosto de 1738 foi criada por carta régia a Aula do Terço de Artilharia do Rio de Janeiro e em 19 de agosto de 1738 José Fernandes Pinto Alpoim foi nomeado lente da aula, tendo ali servido praticamente até a sua morte, em 1765. O documento em que se comunica ao governador a decisão régia esclarece os objetivos e o modo de funcionamento da Aula.

"Faço saber a vos, governador e capitão general da capitania do Rio de Janeiro que por ser conveniente a meu serviço que neste praça onde mandei formar de novo um terço de artilheiros haja Aula onde os oficiais e soldados do dito terço e as mais pessoas que quiserem aplicar-se possam aprender a teoria da artilharia e uso dos fogos artificiais, criando-se por este modo oficiais que depois de instruídos na dita Aula possam ser empregados nos postos da repartição da artilharia dessa e das mais capitanias, fui servido haver por bem por decreto de treze deste presente mês como se estabeleça a dita Aula e para mestre dela nomeei a José Fernandes Pinto Alpoim que proximamente provi no posto de sargento-mor do referido terço, o qual, além dos exercícios a que é obrigado pelo mesmo posto terá o de ditar postila, e ensinar a teoria da artilharia a todos os que quiserem aplicar-se a ela e especialmente aos oficiais do dito terço que nesta primeira criação forem providos , os quais serão igualmente obrigados a assisitir às liçoes da aula ao menos por tempo de cinco anos e faltando a elas serão castigados a vosso critério. E para o futuro não podereis informar para os postos de artilharia do terço, nem aprovar para os denombramentos, oficial algum que não tenha frequentado a dita Aula e seja examinado e aprovado nas matérias que nela se ditarem"[1]

Esta aula ,"formada de novo", surge na sequência da anterior Aula de Fortificação do Rio de Janeiro que vinha funcionando, talvez com algumas interrrupções, desde o final do século XVII. O papel de Alpoim é, como vimos, central na constituição da Aula do Terço, reforçando o interesse nas matérias específicas de artilharia. Mas, tal como na aula precedente, também se ditavam seguramente matérias relativas à fortificação. Assim o garantia a formação de Alpoim feita no reino, inicialmente na Academia de Fortificação de Viana, junto a seu tio, Manuel

Aula do Terço de Artilharia do Rio de Janeiro	
(valor desconhecido)	
Outras denominações	Aula do Regimento de Artilharia do Rio de Janeiro
Tipo de Instituição	Q52
Data de fundação	13 agosto 1738
Data de extinção	1792
Paralisação	Início: valor desconhecido Fim: valor desconhecido
Localização	
Localização	Rio de Janeiro, Rio de Janeiro, Brasil Início: 13 de agosto de 1738 Fim: 1792

Image 6 - Example Aula do Terço de Artilharia do Rio de Janeiro (Q591) (content + data pages)

Meta-information acts as an aggregator of content. In this case, through categories, which are displayed at the bottom of the content pages (see: eViterbo content pages). Select Categories have their own pages, aggregating information, and creating Navigation pages. For example, a category related to the a profession aggregates all the entries of all people which were military engineers or using a geographical category can join all the people working in Bahia, Brazil.

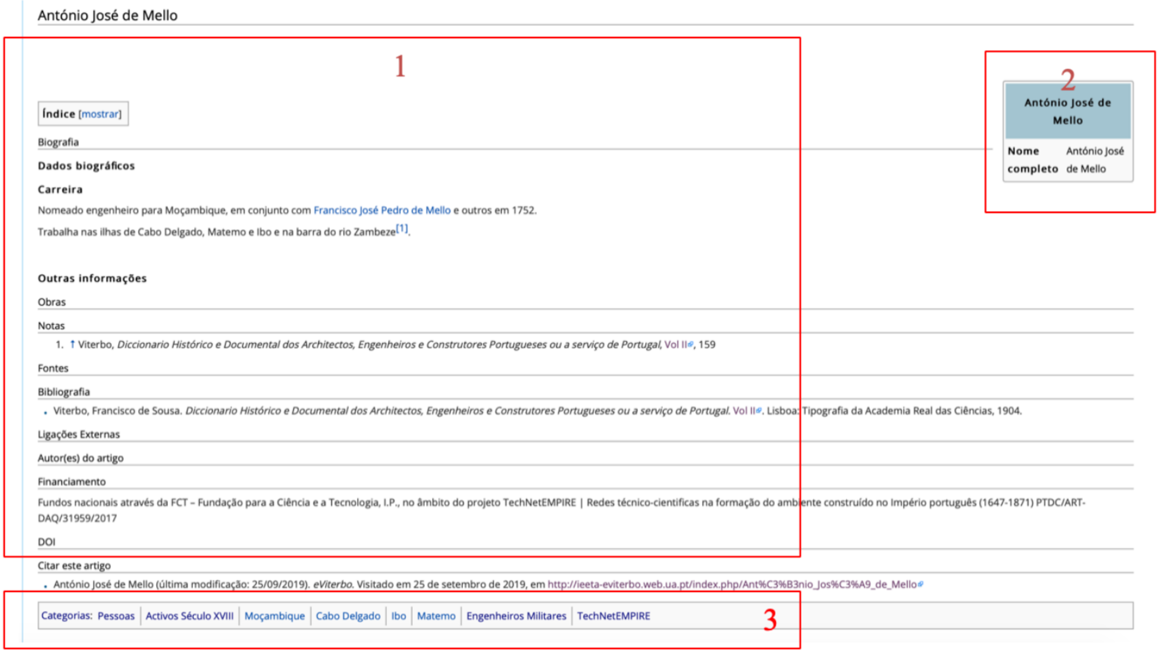
On top of these categories' pages, there is one main navigation hub: the Home page, which allows users to quickly access content segmented into thematic areas and by research project.

System pages are anything that relates to the system itself, encompassing a wide variety of pages: from the creation of new users, information about the software, some extension settings, maintenance pages, templates, and many others.

eViterbo content pages

MediaWiki allows, similarly to Wikipedia, the existence of several templates (predefinições) for content pages. eViterbo contains two types of content pages: type 1 corresponds to "instance of" pages in Wikibase, that is case of the entries on "people" and on "institutions"; type 2, pages not linked to structured data, i.e., the locations and the glossary (glossary pages do not have structured data for the time being).

The pages for Locations were created to solve a practical problem of curation of data in the beginning of the project and should probably not exist in a second upgraded version of eViterbo (see: Ontology v2 . To be continued...). They have structured data associated, but are not connected through an infobox.



António José de Mello

Índice [mostrar]

Biografia

Dados biográficos

Carreira

Nomeado engenheiro para Moçambique, em conjunto com Francisco José Pedro de Mello e outros em 1752.

Trabalha nas ilhas de Cabo Delgado, Matemo e Ibo e na barra do rio Zambeze[1].

Outras informações

Obras

Notas

1. ↑ Viterbo, *Diccionario Histórico e Documental dos Architectos, Engenheiros e Construtores Portugueses ou a serviço de Portugal*, Vol II, 159

Fontes

Bibliografia

- Viterbo, Francisco de Sousa. *Diccionario Histórico e Documental dos Architectos, Engenheiros e Construtores Portugueses ou a serviço de Portugal.* Vol II. Lisboa: Tipografia da Academia Real das Ciências, 1904.

Ligações Externas

Autor(es) do artigo

Financiamento

Fundos nacionais através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto TechNetEMPIRE | Redes técnico-cientificas na formação do ambiente construído no Império português (1647-1871) PTDC/ART-DAQ/31959/2017

DOI

Citar este artigo

- António José de Mello (última modificação: 25/09/2019). *eViterbo*. Visitado em 25 de setembro de 2019, em http://ieeta-eviterbo.web.ua.pt/index.php/Ant%C3%B3nio_Jos%C3%A9_de_Mello

Categorias: Pessoas | Activos Século XVIII | Moçambique | Cabo Delgado | Ibo | Matemo | Engenheiros Militares | TechNetEMPIRE

António José de Mello	
Nome completo	António José de Mello

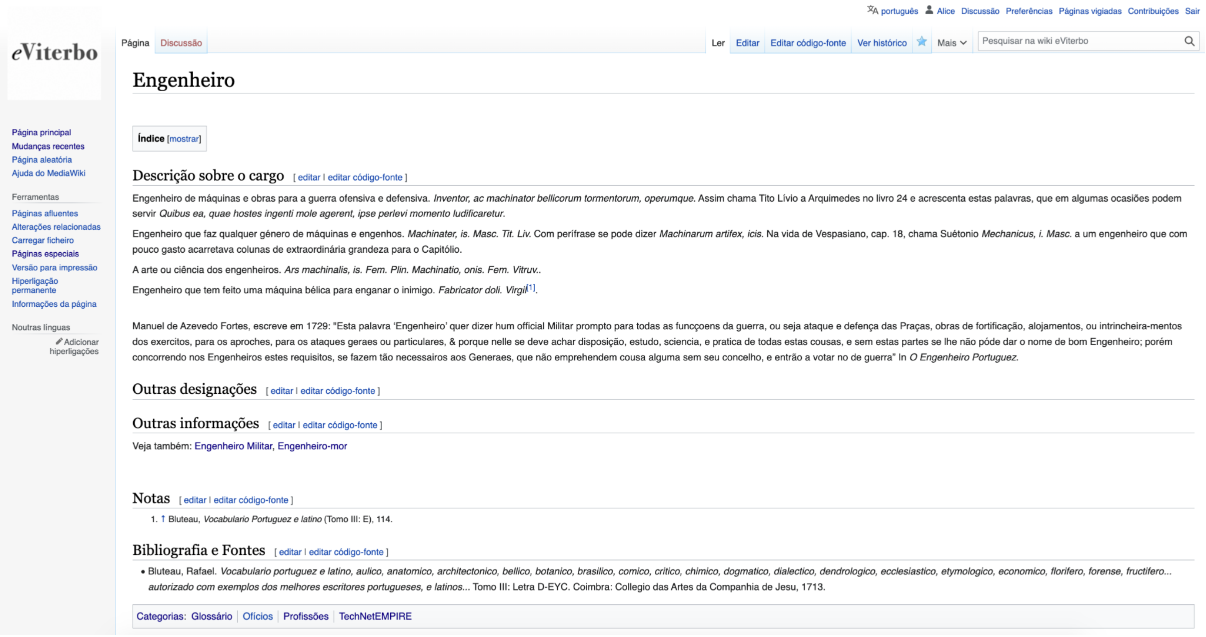
Image 7 - *e*Viterbo's people content pages general scheme (type 1)

A type 1 content page has three distinct parts: the text body (1), the infobox(2) and the categories(3), as shown in image 7. The first part is a descriptive text, allowing for details that may not be part of the structured information found in the infobox. The second one, the infobox, displays the information from the Wikibase entry, through a code snippet inserted in the page. The third part is the categories at the bottom of each article, which is meta-information under which the page is labeled.

In eViterbo, categories were divided into main themes: first, the nature of the the article (whether is about a location, a person or an institution), second the research project under which the entry is developed, followed by several categories which are variable according to the nature of the page or in same case, like geography, that cross and link pages of different natures. For people, the professions and timeframes of when they were active are also typified. Lastly, there are categories for the research projects involved in the making of each article (for further details see: Ontology (v1 - 11.2022)). This categorization allows for easier navigation to other pages under the same topics, even without the use of the linked open data on Wikibase.

To assist researchers with writing the entries in a consistent manner, a series of guides was created, including one for entries about people and another for entries about institutions (*Guião de Redacção - pessoas*; *Guião de redação - Instituições*). Additionally, the team also released a guide defining the rules for citations (*Guião de citação*).

It is also worth mentioning that each eViterbo's content page has a Unique Resource Identifiers (URIs) at the Uniform Resource Locator (URL of the website), which is generated automatically when a page is created and is basically a wiki ID. Select entries also have a Digital Object Identifier (DOI), facilitating adequate citations and referencing of the content, with appropriate authorship included. Authorship is detailed at the bottom of each article and, whenever possible, authors are identified along with their ORCID.



Engenheiro

Descrição sobre o cargo

Engenheiro de máquinas e obras para a guerra ofensiva e defensiva. *Inventor, ac machinator bellicorum tormentorum, operumque.* Assim chama Tito Lívio a Arquimedes no livro 24 e acrescenta estas palavras, que em algumas ocasiões podem servir *Quibus ea, quae hostes ingenti mole agerent, ipse perlevi momento ludificaretur.*

Engenheiro que faz qualquer género de máquinas e engenhos. *Machinater, is. Masc. Tit. Liv.* Com perífrase se pode dizer *Machinarum artifex, icis.* Na vida de Vespasiano, cap. 18, chama Suétonio *Mechanicus, i. Masc.* a um engenheiro que com pouco gasto acarretava colunas de extraordinária grandeza para o Capitólio.

A arte ou ciência dos engenheiros. *Ars machinalis, is. Fem. Plin. Machinatio, onis. Fem. Vitruv.*.

Engenheiro que tem feito uma máquina bélica para enganar o inimigo. *Fabricator doli. Virgil*[1].

Manuel de Azevedo Fortes, escreve em 1729: "Esta palavra 'Engenheiro' quer dizer hum official Militar prompto para todas as funcçoens da guerra, ou seja ataque e defença das Praças, obras de fortificação, alojamentos, ou intrincheira-mentos dos exercitos, para os aproches, para os ataques geraes ou particulares, & porque nelle se deve achar disposição, estudo, sciencia, e pratica de todas estas cousas, e sem estas partes se lhe não póde dar o nome de bom Engenheiro; porém concorrendo nos Engenheiros estes requisitos, se fazem tão necessairos aos Generaes, que não emprehendem cousa alguma sem seu concelho, e entrão a votar no de guerra" In *O Engenheiro Portuguez.*

Outras designações

Outras informações

Veja também: Engenheiro Militar, Engenheiro-mor

Notas

1. ↑ Bluteau, *Vocabulario Portuguez e latino* (Tomo III: E), 114.

Bibliografia e Fontes

- Bluteau, Rafael. *Vocabulario portuguez e latino, aulico, anatomico, architectonico, bellico, botanico, brasilico, comico, critico, chimico, dogmatico, dialectico, dendrologico, ecclesiastico, etymologico, economico, florifero, forense, fructifero... autorizado com exemplos dos melhores escritores portugueses, e latinos...* Tomo III: Letra D-EYC. Coimbra: Collegio das Artes da Companhia de Jesu, 1713.

Categorias: Glossário | Ofícios | Profissões | TechNetEMPIRE

Image 8 - *e*Viterbo glossary content pages general scheme (type 2)

eViterbo also has glossary pages (Image 8). As mentioned, these were initially created under an exploratory project called " As palavras com que nos cosemos...", and were partially developed under the DRESS research project. For TechNetEMPIRE, glossary pages aim to define and explain some terms, such as professions or positions held, that are used on the database and/or on the content pages. These glossary pages do not contain infoboxes, but are important for disambiguation and clarification of terms, which places these as meta-information pages as well.

eViterbo Wikibase

Wikibase is a suite of MediaWiki extensions that enables the possibility of creating collaborative knowledge bases, through linked open data. Comparing with other Wikibase architectures, such as the one used by the EU for its Knowledge Graph (image 9), eViterbo encompasses the following parts: MediaWiki (1), Wikibase Client and Repository (2), Maria DB (3) and Elastic Search (4). Blazegraph (5), the Query Service Updater (5) e Query Service UI (6) were installed but failed to work adequately together - this will be addressed later in this report.

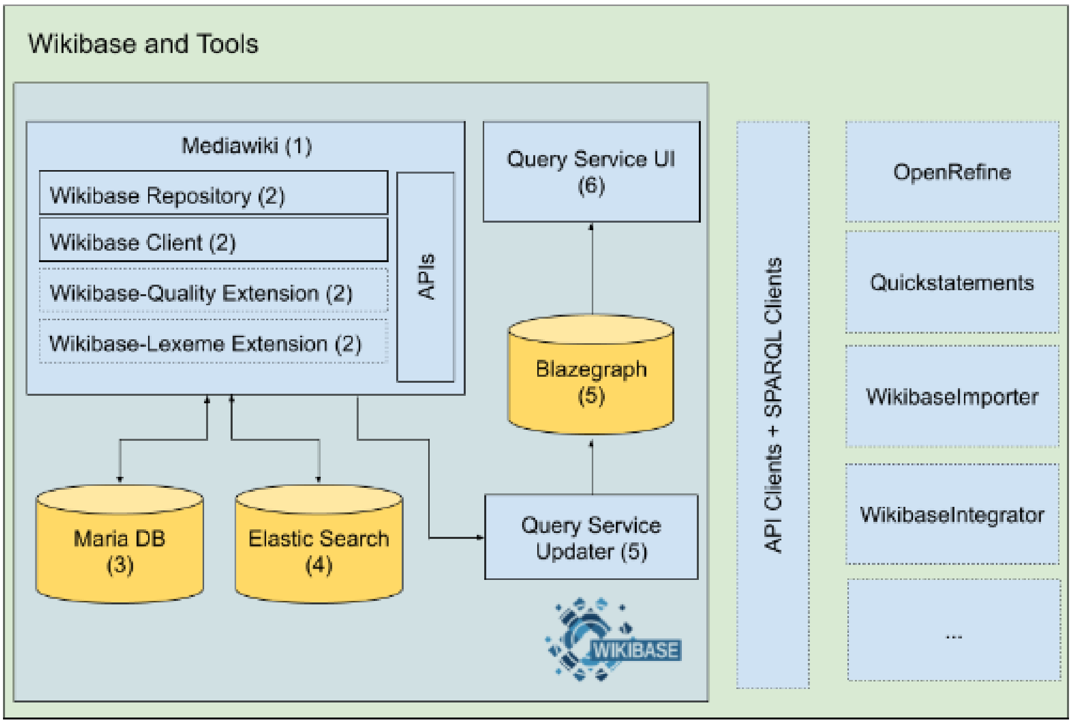
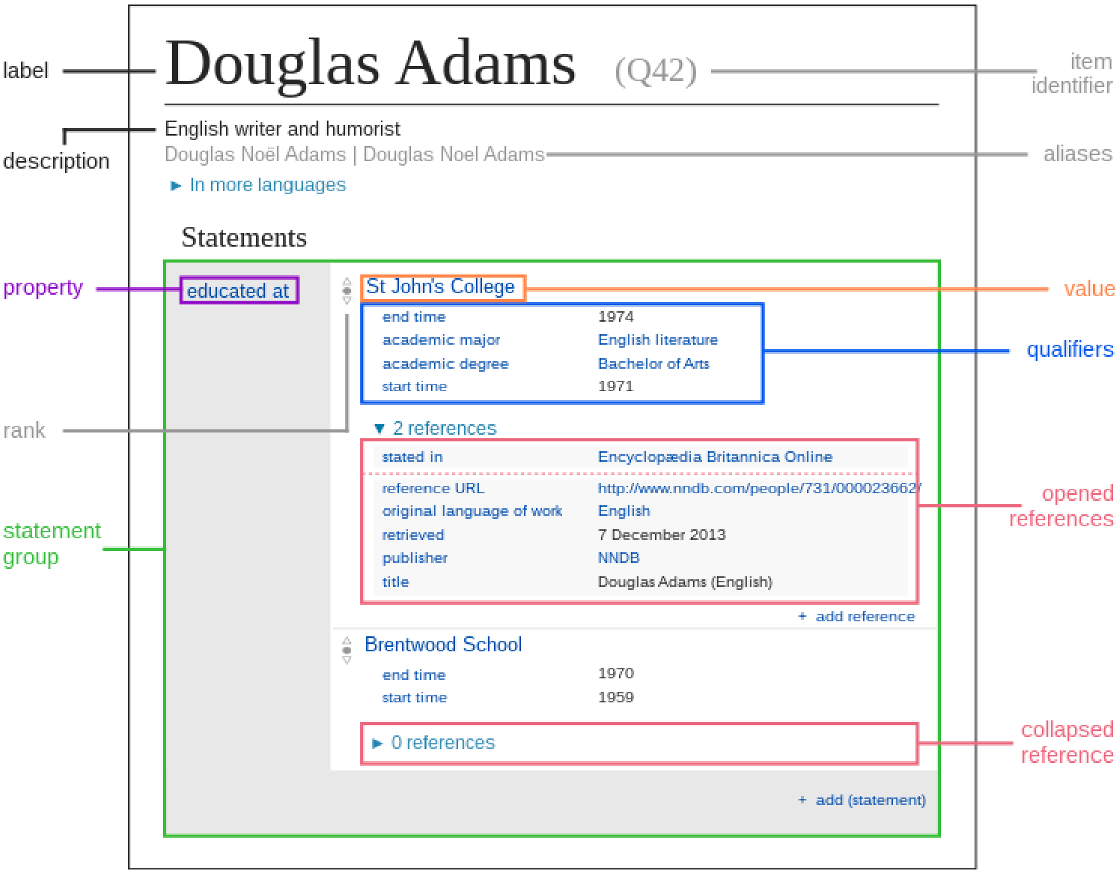


Image 9 - Wikibase Architecture of the EU Knowledge Graph (from D. Diefenbach et al, 2021)

A Wikibase page is called an entity or an item. The starting point of any item is its label and description. When it is created, Wikibase generates a URI which allows for the disambiguation of said item against another with a similar name. Each item is then described in a structured manner by using statements assigning values to a given property, along with adequate qualifiers whenever needed. This structure is similar to a human-language sentence: the item is the subject, the property acts as the verb and is followed by an object, the assigned value. This value is restrained to each property's specific data type (string, wikibase object, date/time, etc.); this is defined when a property is created.



Douglas Adams (Q42)

English writer and humorist

Douglas Noël Adams | Douglas Noel Adams

► In more languages

Statements

educated at

St John's College

end time	1974
academic major	English literature
academic degree	Bachelor of Arts
start time	1971

▼ 2 references

stated in	Encyclopædia Britannica Online
reference URL	http://www.nndb.com/people/731/000023662/
original language of work	English
retrieved	7 December 2013
publisher	NNDB
title	Douglas Adams (English)

+ add reference

Brentwood School

end time	1970
start time	1959

► 0 references

+ add (statement)

Image 10 - Data model in Wikidata/Wikibase, by Charlie Kritschmar (CC 0)

The data added to each entity must be stated and appropriately referenced on the corresponding content page. On eViterbo, values are not constrained to specific topics (see: Constraining properties). To ensure the integrity of the data, items for specific properties were created only by the core team of the project. To help with the process of filling in the appropriate data of each item, a guide was created, detailing what data is allowed in each property - *Guia de Itens*.

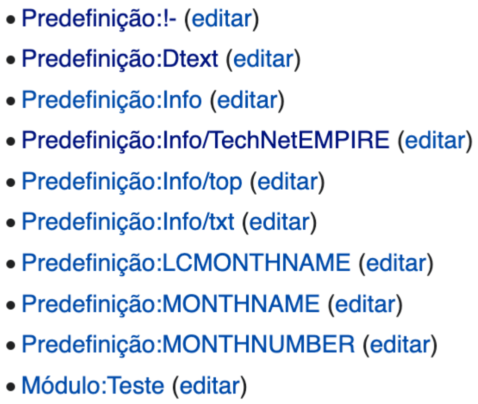
The Wikibase item page connects to the corresponding content page by using the local eViterbo URL as a property. In the opposite direction (content to data), this connection is made through the infobox templates, as previously mentioned.

eViterbo's Templates and Infoboxes

The main way in which the data connects to the content is through infoboxes; these are created using Templates. A Wiki template can work via substitution or transclusion. On eViterbo, Templates are mainly used for the latter, meaning that a specific page can be inserted into other pages. On Wikipedia, for instance, these are commonly used to easily display informational or navigational boxes, that can be anything from "citation needed" (a warning that states that a given article needs more references) or an "infobox", a summarized (typically itemized) version of the most important information about the article's subject. Changing a template that uses transclusion means that only the template needs to be changed, even if it is being used on multiple pages, something rather useful to display repeated information.

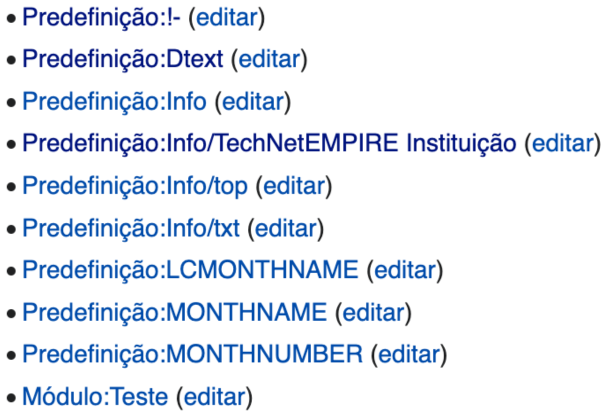
As detailed on image 5, most content pages on eViterbo have infoboxes, namely the pages about People or Institutions. These infoboxes work based on transclusion, but they are dynamic: their content is not inserted manually, but through a code snippet that is added to the source code of each page. The snippet is essentially the same across each category, with one minor adjustment on each page: by changing the identifier of the item associated with the box, different information will be displayed accordingly, as it is fetched from the corresponding Wikibase item.

A people content page has the following templates:



- Predefinição:!- (editar)
- Predefinição:Dtext (editar)
- Predefinição:Info (editar)
- Predefinição:Info/TechNetEMPIRE (editar)
- Predefinição:Info/top (editar)
- Predefinição:Info/txt (editar)
- Predefinição:LCMONTHNAME (editar)
- Predefinição:MONTHNAME (editar)
- Predefinição:MONTHNUMBER (editar)
- Módulo:Teste (editar)

A institution content page has the following templates



- Predefinição:!- (editar)
- Predefinição:Dtext (editar)
- Predefinição:Info (editar)
- Predefinição:Info/TechNetEMPIRE Instituição (editar)
- Predefinição:Info/top (editar)
- Predefinição:Info/txt (editar)
- Predefinição:LCMONTHNAME (editar)
- Predefinição:MONTHNAME (editar)
- Predefinição:MONTHNUMBER (editar)
- Módulo:Teste (editar)

eViterbo's connection with Wikidata

eViterbo is linked to Wikidata, firstly, through Wikibase's properties, as most of them have a designated "Equivalent property on Wikidata".

As new items are created on eViterbo's Wikibase, researchers also look for their counterpart on Wikidata, adding it to the item accordingly. This process is particularly challenging for locations, as eViterbo maps various types of locations, namely administrative units and geographical coordinates, among others, which are associated with different properties on Wikidata.

In the future, gathering these connections could be made in a more automated manner, by using a bot to interact with both eViterbo and Wikidata, or via OpenRefine for data set reconciliation. Nevertheless, the project aim is not to insert into Wikidata tons of new data, but only to input a relatively small amount of data (for now around 3000 people, 150 institutions and an unknown number of related items like professions, locations, activities, etc...), inserting new items, improving general quality of existing items, descriptions, etc. with which we were working with.

Main challenges. Past & Present

Some of the issues raised may have been solved if there was a dedicated engineering team throughout the project. From the project proposal, MediaWiki was suggested as a software solution for building an encyclopedia and database under the reasoning that it would be easy to work with, demanding only the initial technical stretch of the setup. Most of the technical feedback gathered before the set up of eViterbo was that a MediaWiki solution would only require six months of technical work. A masters student was hired for this task, under a masters fellowship, for six months. Despite his best efforts, there were significant development issues regarding eViterbo: from the lack of documentation about the system, to the trial and error approach which led to a significantly unstable platform. It would be rather unfair to imply that most of the issues surrounding eViterbo were his fault. His supervisor had some MediaWiki knowledge, but his commitment to the research project was only 10% of his work time, which is likely to have compromised appropriate technical guidance or supervision.

As the work with eViterbo's evolved, and technical problems became more evident, the researchers sought out help. Over and over again, they were told by technical staff that the technical efforts would be low, but after a few hours of interaction with the platform, technicians would often change their minds and say that anything would be too complicated to handle and that they would not touch the platform. This was persistent throughout almost the entirety of the project, creating a serious mismatch between researchers and technicians, as well as the expectation surrounding the behavior of the platform they were working on. It should be stressed that IT staff at NOVA FCSH is residual and often overloaded. Moreover, basic requirements for debugging software, such as having separate machines for development environments were not accounted for. In fact, the setup of a dev environment where technical improvements and tests could be made in a safe way - without potentially compromising the researchers' work (content and data) - was only possible in the last trimester of the project. During most of the project, there was only one IT technician at NOVA FCSH; later a second one was hired by the faculty, but their focus is on giving general support to the whole institution, without much leeway to provide dedicated support to researchers and their projects. This has serious implications on how researchers design and readjust their projects. While it was not expected from the faculty to provide technical support to the project, moving eViterbo from the servers of Universidade de Aveiro to NOVA FCSH was delayed for a whole year, due to the understaffed situation at the latter and all the extra-technicalities which occurred during the pandemic.

In March 2022, a Wikimedian in Residence (WiR) was hired by Wikimedia Portugal to work at the faculty - supporting research projects already using Wikimedia tools and doing community building and outreach work within the faculty. This was achieved after PI's reach for help to Wikimedia Portugal and Wikimedia Deutschland, which resulted in a proposal for a WiR to Wikimedia Foundation. One of the WiR's main missions was to help fix the technical issues that populated eViterbo, which included, among others: lack of backups and a backup system, database errors, Lua errors, errors and mistakes on infoboxes, ontology issues, CSS styling issues on the footer and duplicate items from previous system updates. While some of these issues were fairly straightforward to handle, many of them required more advanced technical knowledge, beyond her skills. In July 2022, a system administrator engineer was hired on a part-time basis to help with some of the more advanced tasks.

The list below is mostly based on the experience of the Wikimedian in Residence and the engineer hired at the final stage of the project, and their technical research. There may be other solutions to some of the issues faced during eViterbo's lifecycle, that somehow we did not come across or think of, despite our best efforts. Nonetheless, the difficulties detailed may be relevant for future projects with similar features and limited technical teams in-house.

Maintenance difficulties

Software lifecycle and compatibility

MediaWiki has major releases every six months; every two years a long-term support version (LTS) is released. Each new version, and subsequent minor updates, is launched first through selected Wikipedias, then to all Wikipedias, and only later to the wider public as a product ready for self-hosting. Each stable release's lifecycle lasts a year, while the LTS versions are supported for three years. Providing security and stability for longer periods, LTS versions tend to be rather useful for projects that do not need cutting-edge technology and/or with small tech teams, such as eViterbo. eViterbo was initially set up using 1.34 version of MediaWiki and later migrated to 1.35, the current LTS. However, Wikibase only runs adequately on the most up-to-date versions of MediaWiki. This means that even though MediaWiki will continue providing support for 1.35 until September 2023, eViterbo must be frequently updated in order to ensure that the Wikibase part runs smoothly - as an example, one of extensions typically used with Wikibase, the UniversalLanguageSelector (which allows users to select a language and configure its support in an easy way), is only compatible with MediaWiki 1.37 or higher.

While Wikibase's update cycle tends to be in sync with MediaWiki, new releases were less frequent during 2021 and 2022, as WMDE put extra focus on parallel projects. During 2022, most Wikibase releases were regular maintenance ones, such as security patches and bug fixes, and the delayed update cycle generated some technical-debt on the stack. For instance, the recommended version of

Elasticsearch, an optional yet particularly useful search engine software to use with Wikibase, had been deprecated. At the time of writing, this discontinued version is still listed as the recommended one, against general good practice in software maintenance - deprecated versions should be replaced with their updated counterparts.

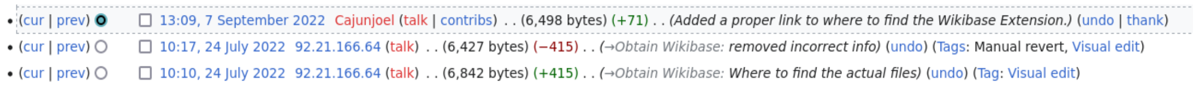
Additionally, some extensions need others to run properly and the compatibility map of each extension, MediaWiki and Wikibase versions is quite hard to navigate, even by experienced engineers. It is also worth mentioning that bundled extensions may not receive as extensive support as MediaWiki, even if included in the LTS versions.

Poor documentation, insufficient support and the generosity of volunteers

Documentation about Wikimedia tools tends to be scattered and often insufficient. Most official information (meaning documentation provided by the developers of the tools) for the technology used on eViterbo is split across three main parts: 1) MediaWiki documentation website, 2) dedicated documentation for specific extensions (these can be on MediaWiki documentation website and/or on separate websites), and 3) Wikibase's documentation website. While all those websites are aimed at a technical audience, the latter is far more advanced than the first two, explaining the full architecture of the system in deep detail.

Many extra features are only possible thanks to the effort and generosity of a wide developer community that shares free and open source software with the other users - from MediaWiki extensions to code libraries for specific use cases. The documentation available varies greatly: while some projects have extensive documentation, in most cases it is only very basic. However, many of these pieces are crucial for basic operations, such as manipulating data in bulk once inside Wikibase.

In most cases, the available documentation summarizes how to install the given piece of software and, occasionally, enumerates some troubleshooting procedures to fix the most common errors. However, often there are important steps or configuration details missing. For instance, in August 2022, when installing a brand new instance of MediaWiki plus Wikibase to set up a developer's environment (inexistent up until that point), the team struggled to find a link for downloading Wikibase. All the instructions available, including on MediaWiki and Wikibase's websites, stated: "Obtain Wikibase", but none of them had a link that pointed towards the download (this was later fixed, as per image below).



- (cur | prev) ● ☐ 13:09, 7 September 2022 Cajunjoel (talk | contribs) . . (6,498 bytes) (+71) . . *(Added a proper link to where to find the Wikibase Extension.)* (undo | thank)
- (cur | prev) ○ ☐ 10:17, 24 July 2022 92.21.166.64 (talk) . . (6,427 bytes) (−415) . . *(→Obtain Wikibase: removed incorrect info)* (undo) (Tags: Manual revert, Visual edit)
- (cur | prev) ○ ☐ 10:10, 24 July 2022 92.21.166.64 (talk) . . (6,842 bytes) (+415) . . *(→Obtain Wikibase: Where to find the actual files)* (undo) (Tag: Visual edit)

Image 11 - Revision history of Wikibase documentation, illustrating the missing link for downloading Wikibase

When error messages are not included in the more generic troubleshooting, it is sometimes possible to find similar bugs reported on Phabricator, Wikimedia's own bug reporting and tracking system. The platform is open, allowing users to read through previously reported issues, both fixed and unresolved, but navigation can be confusing, especially for those less familiar with software development routines. Similarly, valuable information can be sometimes found on Gerrit, Wikimedia's code review system, but it is not really a ticketing or support tool.

MediaWiki's Project:Support_desk serves as a main support channel but it should be stressed that the focus of this forum is MediaWiki, meaning that issues with specific extensions (including Wikibase) may be out of scope. While mailing lists, bug reports on Phabricator and dedicated Telegram channels are all available for help regarding MediaWiki and Wikibase, most actual support often comes on the basis of goodwill from both WMF engineers and community volunteers, who will gladly help people struggling with some of the tools. However, sometimes questions are left unanswered or are answered unclearly, not fixing the problem raised. Thus, this approach to software support is very limited since it does not provide a reliable system for users that need to ensure stable solutions and may have minimal engineering teams in-house.

Significant information, including detailed configuration settings, is often found on the web through the form of tutorials, once again these are provided in a non-systematic way by volunteers who may not be able to help beyond this. Additionally, it can be difficult to tell if these tutorials are still relevant for current versions.

Last but not least, terminology is not always consistent across software and documentation, making instructions often confusing. This issue is amplified by the fact that, while both MediaWiki and Wikibase are localized in a variety of languages, including Portuguese, the documentation is mostly in English and having multiple terms for the same thing can be rather disorienting, making it harder to troubleshoot some issues.

Dynamic infoboxes

As detailed on eViterbo's Templates and Infoboxes section, eViterbo uses Templates to create dynamic infoboxes that display information based on the Wikibase identifier of a given agent, be it People or Institutions. These boxes are created using a mix of Wikitext and Lua functions. Wikitext (also known as Wikicode) is the markup language of Wikimedia projects; it essentially enables easy text formatting. Lua is a lightweight high-level programming language; it is the language used for scripting in Wikimedia platforms. "Scripts" are sets of instructions for a computer to follow, files of code that can be used to perform a variety of tasks - from data transformation to software updating. Lua scripts are run through Scribunto, a MediaWiki extension (bundled with MediaWiki since version 1.34) which enables script embedding.

For eViterbo's dynamic boxes, a Lua script (Module) was created by the masters fellow with a series of functions to fetch specific information from a given Wikibase item. The way this works is: the code snippet of the Templates invokes a function from the Lua Module, the function then returns the requested data from Wikibase. But while this gathers information from some statements, the information related to other statements rely solely on the keywords provided by Wikitext, a far simpler language.

Dynamic boxes that retrieve data from Wikidata are used in Wikipedia and other Wikimedia platforms. On Wikipedia, the Module used to do this has nearly 1900 lines of code. It is arguable whether eViterbo, a much smaller project, would need such a long script. However, with the equivalent Module on eViterbo having only approximately 150 lines, it is no wonder that many of the possibilities enabled on Wikipedia's dynamic boxes are not on eViterbo. Simple details contemplated on the Wikipedia-Wikidata version, such as adding or removing links from the text in the boxes, or hiding statements with 'unknown values' are currently not available.

This intricate structure along with the lack of technical documentation (for example, the functions in eViterbo's Module were not described anywhere and were not even commented on the code) made it very difficult to make the boxes work as the researchers intended once the masters fellow left. Regardless, it was possible to fix the Lua errors that populated some of the boxes, following the migration from Cargo to Wikibase.

Wikibase limitations for managing data in research databases in social sciences

Wikibase provides an easy-to-use interface for creating and editing entities on a LOD environment, theoretically enabling non-technical users to interact with the data in an uncomplicated manner. Nonetheless, using Wikibase to create and manage a research database comes with its own set of challenges.

Building ontologies under Wikibase

As previously mentioned, one of the advantages of Wikibase in comparison to other software products is the curation of data. This means that users have complete control over the ontology and schemas attached to the data, being able to set constraints and narrow possible entries from the start.

Since May 2021, the federated architecture allows for a direct connection with other Wikibases, Wikidata being the main one. For every property created, it is possible to set the corresponding one on Wikidata. With more than eleven thousand properties (and counting), aligning Wikibase properties with their Wikidata counterparts can become a significant challenge. For instance, the field "Also known as", which allows for each item to include alternative labels (different spellings, full names, etc.) for better disambiguation, is included on every item by default. However, in eViterbo, this was added as a dedicated property. This was only highlighted as a potential issue during the ontology revision session with the WiR. But while for Wikidata this may be redundant, for the granular approach required by historians for eViterbo none of the properties provided by Wikidata really matched what they were hoping for (see: Ontology (v1 - 11.2022). This scenario illustrates how it might also be the case that this alignment is not wanted. Wikidata's ontology is governed by its own community, whose interests and needs may not coincide with those of researchers or other groups - compiling knowledge about everything is quite different from compiling highly specialized knowledge on a very specific topic. Yet, research communities can also present an opportunity to reflect upon properties and items in Wikidata, which sometimes overlap and exhibit some ambiguity. Nevertheless, a community-driven approach can also mean that, over time, some properties might be replaced with others.

Wikibase also requires some wide-scale knowledge of the Wikimedia universe, as some of its intricacies are not always clear in the documentation. As *e*Viterbo system was assembled by an informatics engineering student with no previous knowledge of MediaWiki, sometimes basic rules of the Wikidata system were not used in *e*Viterbo. One striking example: while each type of item had its own set of properties, "instance of" was not included in the main ontology. This is arguably the main property of Wikidata, a defining one for the system, as plenty of extensions and tools - such as OpenRefine - rely on it to reconcile data with separate data sets. This was identified on the initial review of the project by the WiR and Wikimedia Portugal,and set in *e*Viterbo by the end of the project.

Constraining properties

As previously mentioned, by allowing the addition of new properties at any point, Wikibase is more flexible than some of the other options for handling data in MediaWiki, but it still struggles with the rigidity of data types. Once a data type is assigned to a property, it is not possible to change it; the only solution is to create a new property, move the data and delete the old one. Later in this report, we will highlight how moving data comes with its own setbacks.

The ability to curate data that the researchers were hoping for when migrating to Wikibase fell a bit short of expectations. Constraining properties on eViterbo is only set up at the data type level. This means that, unlike Wikidata, values are not constrained to specific topics. For instance, it is technically possible to add an institution as a value for the property Gender, since Gender accepts Wikibase-objects (feminin and masculin) and an institution on eViterbo is also a Wikibase object. While this is a far-fetched example, as it is very clear that an institution is not an appropriate value for gender, it is a genuine challenge when working with concepts with similar names. There are multiple items around "engineering" - Engenharia (Q93) is a military branch, Engenharia Civil (Q70) is training, Engenheiro militar (Q16) is a military official in the branch of engineering and Engenheiro (Q2903) is a position held in administration.

Additional constraints are only possible using an extra extension, WikibaseQualityConstraints. Handling this extension demands advanced technical knowledge, as the process of setting up a new constraint type requires changing multiple system files. As such, it was never used in eViterbo but should be included in a second upgraded version of *e*Viterbo.

Basic features often served as extensions or separate software

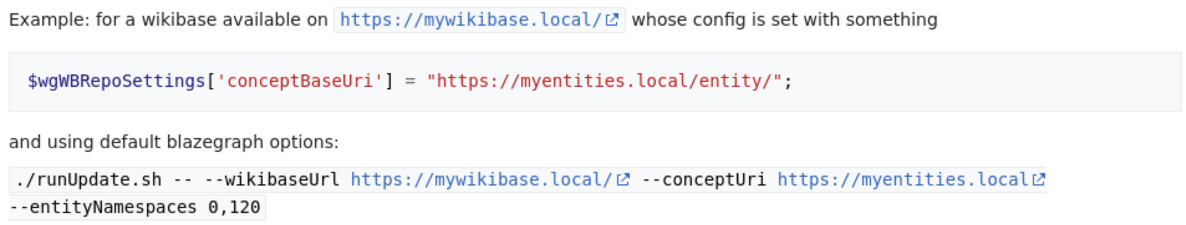
The WikibaseQualityConstraints serves as a case-in-point for one of Wikibase's most evident weaknesses: basic features of Wikidata are often provided as extensions or separate services. This can be said for Wikidata Query Service, a crucial service that allows users to query their datasets in SPARQL, facilitating data extraction and visualization, has to be set up separately as an external service. It is possible to query the data through Wikibase's API as an alternative, but even for those coming from a Wikidata background, where SPARQL queries are the standard, it may not be as easy to navigate.

Exporting the data from Wikibase to a JSON dump is another basic feature made possible through installing a separate software, namely Wikibase Dump Filter. Without that, one is limited to the MediaWiki XML export file, which includes every single piece of a rather complex system - from system files to lists of users and page contents. Despite being possible to filter the XML export, by selecting only Wikibase properties and entities to be in the dump file, manipulating such XML still poses a significant challenge, especially for those with low technical proficiency. For programmers, handling that XML is made easier by using a specific Python library, called MediaWiki XML Processing. However, this data still needs to be heavily manipulated to get the latest revision of each item.

Data modelling in Wikibase

Remodeling data once in Wikibase can be troublesome. In order to remodel data in a research context, one needs: 1) access to the full dataset or to a filtered version of that set and 2) a way to edit multiple entities at once. For specific cases, like eViterbo, it is also necessary to ensure that the tools used are user-friendly enough for people without programming knowledge. This initially ruled out Wikibase API; according to the documentation, it "allows querying, adding, removing and editing information on Wikidata or any other Wikibase instance". However, it also requires some familiarity with MediaWiki Action API, as well as using APIs in general - something rather basic for a developer or a computer engineer, but a huge leap for a standard user. Consequently, the solutions available were, once again, provided by external tools, namely OpenRefine and/or QuickStatements.

Due to its familiar graphic interface, OpenRefine would be a great option even for researchers with lower technical skills. OpenRefine is described as "a free and open source tool for working with messy data". Its interface resembles any regular spreadsheet-handling software, which is likely to be recognizable by most users. OpenRefine works with Wikidata out of the box and is prepared to work with any self-hosted Wikibase. In order to do so, it needs a working reconciliation service which in turn needs an associated SPARQL query service. Designed to work specifically with Wikidata, Wikidata Query Service (WQS) is made available with warnings that it may not function properly outside WMF's infrastructure, as many things are set as hard coded. Another limitation is the need to constantly update the dataset to ensure up-to-date data, instead of having a live service. Regardless, to ensure that OpenRefine could be used, the engineer installed WQS on eViterbo. However, the default settings did not work with eViterbo's concept-URI. Despite multiple attempts to set WQS up and even to change eViterbo's concept-URI, this was not possible.



Example: for a wikibase available on https://mywikibase.local/ whose config is set with something

```
$wgWBRepoSettings['conceptBaseUri'] = "https://myentities.local/entity/";
```

and using default blazegraph options:

```
./runUpdate.sh -- --wikibaseUrl https://mywikibase.local/ --conceptUri https://myentities.local
--entityNamespaces 0,120
```

Image 12 -Default settings for running Wikidata Query Service outside-WMF's infrastructure

Developed by Magnus Manske, QuickStatements (QS) is a tool aimed at bulk-editing Wikidata. Through a graphic user interface, a variety of tasks (such as: adding and removing statements, adding references, creating or merging items, among others) can be performed on batches of entities, by using a simple syntax formula. This tool can be installed in self-hosted Wikibase instances, and it even comes bundled with Docker's image of Wikibase. Docker is a software that enables containerized application development, meaning that said applications can be built and run safely in closed environments, minimizing potential conflicts and side effects of software writing and debugging; it also enables easily-reproducible machine setups. Wikibase has a Docker version, but during most of *e*Viterbo's project lifecycle this image was not full-featured and was not maintained as regularly; so, QS was manually installed on *e*Viterbo. Despite the successful installation, the OAuth module that allows user authentication of connected apps was stuck on a loop, preventing the usage of the tool. This was later fixed, after a maintenance update of Wikibase, but issues regarding the OAuth connection lingered as the team struggled to have QS running on the production environment, despite it working on the dev one. So, we ended up transforming the dev environment into the production environment so that it would be possible to use QS adequately. It is also worth noting that QS only works with public wikis, a serious limitation considering that eViterbo, much like many research projects, was a private wiki while researchers were adding and tweaking data and content.

Analysis and visualization of data are a crucial part of the researchers' work. It is necessary for their ongoing work, but also as a way of finding gaps and problems in data. Data dumping (in a variety of formats and ways) and reinserting data in specific software used for analysis and visualizations, such as Nodegoat the software used by TechNetEMPIRE team, was made manually. This was a very time-consuming solution that presented major challenges each time. So, the difficult access to data had a big impact on the team’s work and possible outputs.

There should be an easy way for researchers who may not have advanced technical knowledge to access and manipulate data once inside Wikibase. Without that it is hard to make the case for Wikibase as a viable solution for research teams without significant software engineering support, on a permanent basis.

Digital literacy and savviness across the research team

Many of the researchers often felt uneasy while editing on the platform, as they were not used to working with similar technologies on their previous projects. Apparently simple tasks like including code snippets, such as the infobox templates, proved to be overwhelming to most of the team. As the team had no technical support for some time, the hardship of using a new tool was only made worse by the persistent errors (e.g. in the infoboxes) and the inability to fix them.

Throughout the project, there were several hands-on and training sessions focusing on Wikimedia tools and other digital tools used within the project, namely Zotero and Nodegoat. Additionally, as previously detailed, a multitude of “How to” guides were created throughout the project (and its multiple iterations) in order to help researchers. For instance, there were guides for the pages of people and institutions created in May 2019, before the “Visual Editor” was included on MediaWiki. Another set of “how to”guides were created in September 2019, when both the Visual Editor and Wikibase were first installed. So many changes in the system caused a certain discomfort, which became more apparent when creating and completing Wikibase items.

In order for the researchers to know what information should be added (properties and qualifiers), most of the items were created with all their properties set as ‘unknown values’, to create a placeholder. This is against the use-cases/expected behavior of Wikibase. By default, the system is prepared in a way that only properties with assigned values should be created for a given item. This later posed an issue, as the infoboxes on the encyclopedia-pages were configured to display all the information available from the Wikibase item, which included those “unknown values”. Regardless, the creation of these placeholders was of extreme importance while working in a collaborative way within a team of 20 researchers and several students, coming in and out of the team as school semesters go by, and often with incomplete data. The introduction of properties and qualifiers, one by one by each user, is prone to create inconsistency on the data set collected - specifically regarding missing data, since some researchers will remember to fill in all the properties but others will not. Predefined fields serve as a reminder of the data that should be filled. Moreover, they are important for the future evolution of the data. Since eViterbo platform will not be closed within TechNetEMPIRE research project, but will continue and new data added by other projects with different researchers, such placeholders can be a valuable indication of the historical data that is still unknown.

Conclusions

Throughout the project, as is mentioned in this report, it became clear that the MediaWiki-encyclopedia plus Wikibase can be a powerful tool for collaboration in a research setting, including in social and human sciences, where data sets and models often need to be rather flexible. It is also obvious that the growing interest in academia regarding open science (open data, open access, open software) can widen possibilities for researchers.

Nonetheless, it is important to highlight that without appropriate full-time technical support, such a solution is not viable. "Appropriate full-time technical support" means having an engineer that can set up and maintain the system adequately and in a timely manner, (our experience tells us that a system administrator is probably the most essential role in this context; PHP and Lua knowledge also help greatly), but it goes beyond the general IT support. Having a deep understanding of Wikimedia platforms plays an important part, since these pieces of software are primarily built for them. Even if, at times, her technical knowledge was not enough to safely solve some more complex problems, having a Wikimedian in Residence was particularly helpful in troubleshooting and identifying underlying issues with eViterbo, navigating available documentation, advising on possible solutions and tools, and overall ensuring that the system operated as intended. Issues arising from the lack of wiki-specialized knowledge are notorious when looking at some of the technical decisions made during the early stages of eViterbo, which ended up causing errors later on (e.g. the way the infoboxes were set up with limited functions) and problems with data modeling (e.g. no "instance of" property being used from the beginning).

Looking at other projects using Wikibase at an institutional level, it is striking how different chapters of Wikimedia tend to be listed as partners on most use-cases, particularly in Europe (from Wikimedia Deutschland to Wikimedia Italia and, now, Wikimedia Portugal). Following the experience with eViterbo, this information does not come as a surprise. This type of support may be crucial for adopting these tools, as the learning curve is far too steep even with such guidance.

Despite WMDE's efforts to make it easier to install and setup Wikibase (via the Docker image and Wikibase.cloud), as it stands today, it is hard to envision using a self-hosted Wikibase instance without this type of support, particularly in less technically-inclined settings, like Human and Social Sciences tend to be. Furthermore, it would be good if WMDE and WMF could improve the way they provide support for using their software, both on the documentation part and direct support.

Software and technology change very rapidly, but universities often do not. Without adequate technical infrastructure and support, without a digital research strategy at their faculties. For instance, the creation of digital labs for Social Sciences that include information and system engineers, as well as people that can build the bridge between the technical and non-technical sides, could be beneficial. Until then, researchers in Social Sciences will continue to struggle with technical challenges instead of being able to focus mostly on their domain of expertise and research processes.

System Details (v1 - 11.2022)

Main software stack

Product	Version
MediaWiki	1.38.4
PHP	7.4.3 (fpm-fcgi)
MariaDB	10.3.34-MariaDB-0ubuntu0.20.04.1
ICU	66.1
LuaSandbox	3.0.3
Lua	5.1.5

Installed extensions

Special pages				
Extension	**Version**	**License**	**Description**	**Authors**
Data Transfer	1.4 (59e2f97) 06:08, 21 September 2022	GPL-2.0-or-later	Allows for importing and exporting data contained in template calls	Yaron Koren and others

Delete Batch	1.8.1 (67c86f4) 06:27, 14 September 2022	GPL-2.0-or-later	Delete a batch of pages	Bartek Łapiński and others
Nuke	–	GPL-2.0-or-later	Gives administrators the ability to mass delete pages	Brion Vibber and Jeroen De Dauw
Replace Text	1.7-alpha	GPL-2.0-or-later	Provides a special page to allow administrators to do a global string find-and-replace on all the content pages of a wiki	Yaron Koren, Niklas Laxström and others
UserMerge	1.10.1 (01c1631) 06:01, 12 October 2022	GPL-2.0-or-later	Merges references from one user to another user in the wiki database - will also delete old users following merge. Requires *usermerge* privileges	Tim Laqua, Thomas Gries and Matthew April

Editors				
Extension	Version	License	Description	Authors
VisualEditor	0.1.2	MIT	Visual editor for MediaWiki	Alex Monk, Bartosz Dziewoński, C. Scott Ananian, Christian Williams, David Lynch, Ed Sanders, Inez Korczyński, James D. Forrester, Moriel Schottlender, Roan Kattouw, Rob Moen, Subramanya Sastry, Thalia Chan, Timo Tijhof, Trevor Parscal and others
Parser hooks				
Extension	Version	License	Description	Authors
Cite	–	GPL-2.0-or-later	Adds <ref> and <references> tags for citations	Ævar Arnfjörð Bjarmason, Andrew Garrett, Brion Vibber, Ed Sanders, Marius Hoch, Steve Sanbeg, Trevor

				Parscal and others
InputBox	0.3.0	MIT	Allow inclusion of predefined HTML forms	Erik Moeller, Leonardo Pimenta, Rob Church, Trevor Parscal and DaSch
ParserFunctions	1.6.0	GPL-2.0-or-later	Enhance parser with logical functions	Tim Starling, Robert Rohde, Ross McClure and Juraj Simlovic
Scribunto	–	GPL-2.0-or-later AND MIT	Framework for embedding scripting languages into MediaWiki pages	Victor Vasiliev, Tim Starling and Brad Jorsch
TemplateData	0.1.2	GPL-2.0-or-later	Implement data storage for template parameters (using JSON)	Timo Tijhof, Moriel Schottlender, James D. Forrester, Trevor Parscal, Bartosz Dziewoński, Marielle Volz and others
Variables	2.5.1 (b28b78e) 10:28, 24 May 2022	ISC	Parser functions allowing to work with dynamic variables in an article scoped context	Rob Adams, Tom Hempel, Xiloynaha, Daniel Werner, MGChecker and others

Wikibase				
Extension	**Version**	**License**	**Description**	**Authors**
WikibaseClient	– (1892948) 06:01, 12 October 2022	GPL-2.0-or-later	Client for the Wikibase extension	The Wikidata team and others
WikibaseRepository	– (1892948) 06:01, 12 October 2022	GPL-2.0-or-later	Structured data repository	The Wikidata team and others
Other				
Extension	**Version**	**License**	**Description**	**Authors**
OATHAuth	0.5.0	GPL-2.0-or-later AND GPL-3.0-or-later	Provides authentication support using HMAC based one-time passwords	Ryan Lane, Robert Vogel <vogel@hallowelt.com> and Dejan Savuljesku <savuljesku@hallowelt.com>
OAuth	1.1.0 (e2fe6a1) 05:59, 12 October 2022	GPL-2.0-or-later	Allows usage of OAuth 1.0a and OAuth 2.0 for API authorization	Aaron Schulz, Chris Steipp, Brad Jorsch, Robert Vogel and Dejan Savuljesku

UniversalLanguageSelector	2022-01-19 (fb452e4) 06:01, 12 October 2022	GPL-2.0-or-later OR MIT	Gives the user several ways to select a language and to adjust language settings	Alolita Sharma, Amir Aharoni, Arun Ganesh, Brandon Harris, Kartik Mistry, Niharika Kohli, Niklas Laxström, Pau Giner, Santhosh Thottingal, Siebrand Mazeland and others

Ontology (v1 - 11.2022)

As mentioned, in the eViterbo's system description section, most of the information on eViterbo can be divided into one of the following groups: content, data, navigation and system pages. As previously stated, by 'content', we mean the encyclopedia-like entries; by 'data' we mean the structured data items (wikibase) associated with each one of the content entries or to its related data. Thus, an entry in eViterbo contains two pages: the content one and the data one, both linked bilaterally.

The basic ontology of the system was created considering these links and established by types of content pages guides: people, institutions, locations (see eViterbo content pages above). It was established as "how to write" script for each content page, where the edition rules for the content pages, the controlled vocabularies for Wikibase and categories, as well as taxonomies for categories were created. Locations were created by a core team members as well as the Wikibase-items created to fill certain properties such as: sex, religion, military arm, post, etc.

As already described in section eViterbo wikibase, a list of items was created in order to guide researchers, see: Guia de itens. Some of these items probably should/could have been established as qualifiers.

Complementary to the content pages templates in the next pages a detailed overview of data (wikibase pages) is presented in tables of properties and properties equivalents in Wikidata. These tables mirror the system details V1 - 11.2022. It should be noted that, during the project, it became clearer that some of these properties may be slightly redundant (P34, a person's name, is already included on Wikibase as the label for the entity). Shared properties are marked with an asterisk. Whenever appropriate, the equivalent property on Wikidata is listed, but that was not possible for every case. There are also some special fields: label, description, also know as (aliases), link to Wikipedia and other Wikimedia platforms. On eViterbo, these interwiki boxes are used to link the item to its counterpart on Wikidata.

In the last column, “Notes and suggested improvements”, we detail possible improvements to the system’s ontology, for a new iteration of eViterbo (see *e*Viterbo v2 (to be continued…)).

eViterbo PEOPLE **Vocabularies: Guia de Redação - Pessoas**				**EQUIVALENT PROPERTY**	**Notes and suggested improvements**
Name		**URL**	**Data Type**	**Wikidata**	
Instância de		https://eViterbo.fcsh.unl.pt/wiki/Property:P15	Object/Item	https://www.wikidata.org/wiki/Property:P31	
Nome	**P34**	http://eViterbo.fcsh.unl.pt/wiki/Property:P34	String	https://www.wikidata.org/wiki/Property:P2561	Since each item's basic identification includes a label, this property is redundant. To be deleted in V2.
Nome Completo	**P17**	http://eViterbo.fcsh.unl.pt/wiki/Property:P17	String	https://www.wikidata.org/wiki/Property:P1477	
Outras grafias	**P20**	http://eViterbo.fcsh.unl.pt/wiki/Property:P20	String	“Also known as”	Since each item's basic identification includes an “also known as” field, this property may be redundant. This is tied to the different spellings of a given name through time.

Pai	**P22**	http://eViterbo.fcsh.unl.pt/wiki/Property:P22	Object/Item	https://www.wikidata.org/wiki/Property:P22	
Mãe	**P23**	http://eViterbo.fcsh.unl.pt/wiki/Property:P23	Object/Item	https://www.wikidata.org/wiki/Property:P25	
Irmãos	**P25**	http://eViterbo.fcsh.unl.pt/wiki/Property:P25	Object/Item	https://www.wikidata.org/wiki/Property:P3373	
Cônjuge	**P24**	http://eViterbo.fcsh.unl.pt/wiki/Property:P24	Object/Item	https://www.wikidata.org/wiki/Property:P26	
Data de nascimento	**P2**	http://eViterbo.fcsh.unl.pt/wiki/Property:P2	Point in time	https://www.wikidata.org/wiki/Property:P569	
Local de nascimento	**P4**	http://eViterbo.fcsh.unl.pt/wiki/Property:P4	Object/Item	https://www.wikidata.org/wiki/Property:P19	

Data de morte	**P3**	http://eViterbo.fcsh.unl.pt/wiki/Property:P3	point in time	https://www.wikidata.org/wiki/Property:P570	
Data de início*	**P9**	http://eViterbo.fcsh.unl.pt/wiki/Property:P9	point in time	https://www.wikidata.org/wiki/Property:P580	Start date. Used as a qualifier to indicate the point in time a given statement has validity.
Data de fim*	**P10**	http://eViterbo.fcsh.unl.pt/wiki/Property:P10	Point in time	https://www.wikidata.org/wiki/Property:P582	End date. Used as a qualifier to indicate the point in time a given statement stopped being valid.
Instituição	**P12**	http://eViterbo.fcsh.unl.pt/wiki/Property:P12	Object/Item		Name of the institution of training or where a position was held.
Residência	**P26**	http://eViterbo.fcsh.unl.pt/wiki/Property:P26	Objecto/item*	https://www.wikidata.org/wiki/Property:P551	
Formação	**P48**	http://eViterbo.fcsh.unl.pt/wiki/Property:P48	Object/Item*	https://www.wikidata.org/wiki/Property:P8258	Professional training

Local de formação	**P28**	http://eViterbo.fcsh.unl.pt/wiki/Property:P28	Object/Item		Since Locations are already referenced using their geocoordinates, this should be replaced with the "Location" property, but used as a qualifier.
Posto	**P50**	http://eViterbo.fcsh.unl.pt/wiki/Property:P50	Object/Item*	https://www.wikidata.org/wiki/Property_talk:P410	
Arma	**P52**	https://eViterbo.fcsh.unl.pt/wiki/Property:P52	Object/Item	https://www.wikidata.org/wiki/Property:P241	
Cargo	**P8**	http://eViterbo.fcsh.unl.pt/wiki/Property:P8	Object/Item*	https://www.wikidata.org/wiki/Property:P39	
Instituição	**P12**	http://eViterbo.fcsh.unl.pt/wiki/Property:P12	Object/Item*		Can be used as a “cargo” qualifier.
Local do Cargo	**P11**	http://eViterbo.fcsh.unl.pt/wiki/Property:P11	Object/Item*	https://www.wikidata.org/wiki/Property:P131	Since Locations are already referenced using their geocoordinates, this should be replaced with the "Location" property, but used as a qualifier.

Actividades	**P56**	http://eViterbo.fcsh.unl.pt/wiki/Property:P56	Object/Item*	https://www.wikidata.org/wiki/Property:P101	
Local de Actividade	**P57**	http://eViterbo.fcsh.unl.pt/wiki/Property:P57	Object/Item	https://www.wikidata.org/wiki/Property:P276	Since Locations are already referenced using their geocoordinates, this should be replaced with the "Location" property, but used as a qualifier.
url eViterbo local	**P30**	http://eViterbo.fcsh.unl.pt/wiki/Property:P30	URL		
local enterramento	**P33**	http://eViterbo.fcsh.unl.pt/wiki/Property:P33	Object/Item	https://www.wikidata.org/wiki/Property:P119	

eViterbo INSTITUTIONS Vocabularies: Guia de Redacão - Instituições				EQUIVALENT PROPERTY	Notes and suggested improvements
Name		**URL**	**Data Type**	**Wikidata**	
Instância de	**P15**	https://eViterbo.fcsh.unl.pt/wiki/Property:P15	Object/Item	https://www.wikidata.org/wiki/Property:P31	
outras denominações instituição	**P68**	http://eViterbo.fcsh.unl.pt/wiki/Property:P68	String		Since each item's basic identification includes aliases, this property is redundant. To be deleted in V2.
sigla	**P69**	http://eViterbo.fcsh.unl.pt/wiki/Property:P69	String	https://www.wikidata.org/wiki/Property:P1813	
tipo instituição	**P36**	http://eViterbo.fcsh.unl.pt/wiki/Property:P36	String		Should be "subclass of" institution. Data type should be object.
data fundação	**P62**	http://eViterbo.fcsh.unl.pt/wiki/Property:P62	Point in time	https://www.wikidata.org/wiki/Property:P571	

data extinção	**P63**	http://eViterbo.fcsh.unl.pt/wiki/Property:P63	Point in time	https://www.wikidata.org/wiki/Property:P576	
paralisação início	**P63**	http://eViterbo.fcsh.unl.pt/wiki/Property:P63	Point in time		
paralisação fim	**P70**	http://eViterbo.fcsh.unl.pt/wiki/Property:P70	Point in time		
Nome de localidade	**P45**	http://eViterbo.fcsh.unl.pt/wiki/Property:P45	Object/Item		Since each item's basic identification includes a label, this property is redundant. To be deleted in V2.
Local instituição	**P35**	http://eViterbo.fcsh.unl.pt/wiki/Property:P35	Object/Item		Since Locations are already referenced using their geocoordinates, this should be replaced with the "Location" property, but used as a qualifier.
Data de início*	**P9**	http://eViterbo.fcsh.unl.pt/wiki/Property:P9	Point in time		
Data de fim*	**P10**	http://eViterbo.fcsh.unl.pt/wiki/Property:P10	Point in time		

antecessora	**p60**	http://eViterbo.fcsh.unl.pt/wiki/Property:P60	Object/Item	https://www.wikidata.org/wiki/Property:P1365	
sucessora	**P61**	http://eViterbo.fcsh.unl.pt/wiki/Property:P61	Object/Item	https://www.wikidata.org/wiki/Property_talk:P1366	
url eViterbo local	**P39**	http://eViterbo.fcsh.unl.pt/wiki/Property:P30	**URL**		

eViterbo LOCATIONS **See: Guia de Locais**				**EQUIVALENT PROPERTY**	**Notes and suggested improvements**
Name		**URL**	**Data Type**	**WIKIDATA**	
Instância de	**P15**	https://eViterbo.fcsh.unl.pt/wiki/Property:P15	Object/Item	https://www.wikidata.org/wiki/Property:P31	
Nome localidade	**P45**	http://eViterbo.fcsh.unl.pt/wiki/Property:P45	String		Since each item's basic identification includes a label, this property is redundant. To be deleted in V2.
url eViterbo local	**P30**	http://eViterbo.fcsh.unl.pt/wiki/Property:P30	URL		
coordenadas	**P29**	http://eViterbo.fcsh.unl.pt/wiki/Property:P29	Geocoordinates	https://www.wikidata.org/wiki/Property:P625	

Categories

Below, find a list with the categories related to TechNetEmpire, as stated in the various guides to writing and entry. It does not include maintenance or system categories.

CATEGORIAS				
	PESSOAS	PROFISSÕES	...	
		OBRAS	...	
	INSTITUIÇÕES	TIPO DE INSTITUIÇÃO	ADMINISTRAÇÃO	CENTRAL
				LOCAL
			ENSINO	MILITAR
				CIVIL
				RELIGIOSO
			RELIGIOSA	
			CULTURAL	
			MILITAR	
			PROFISSIONAL	
	ACTIVOS EM (...)	SÉCULOS	...	
		GEOGRÁFICAS	...	
	GLOSSÁRIO	...		
	PROJECTO	...		

eViterbo v2 (to be continued…)

With all the content made available under a Creative Commons license, eViterbo's content and data can be reused with barely any restrictions, narrowing the distance between academia and society. eViterbo is now a valuable resource not only for the history of architecture and engineering during the Portuguese colonial empire era, but also for art historians or anyone interested in the history of Portugal and related countries. Despite being open to the general public for reading, it should be noted that new editors for eViterbo must be approved ensuring appropriate peer reviewing of any content added.

As the project developed, it became evident that a second improved version of eViterbo should be created in the near future. A version encompassing the several recommendations spread throughout this report, that we resume briefly in this last section of this report the most important ones:

- The advantages and disadvantages of a large platform for art history or a smaller dataset dedicated to built environment (i.e. architecture, engineering, urban, construction and landscape history) related themes should be considered, as the navigation through multiple projects of multiple areas can be confusing;
- Content pages about locations should disappear, since their content is redundant, existing in its entirety in the data pages. Any further content can be moved, a be added to the navigation page of geographical categories;
- Data ontology should be reviewed. Deleting redundant properties and using qualifiers in a systematic way are among the most important transformations;
- "Empty placeholders" of Wikibase items should continue to be used, but new infobox templates must be created to fix the display of missing information;
- Install Wikibase Quality Constraints, to ensure the quality of data;
- Install QuickRefine and other ways of "easily" modeling and exporting data;
- Explore the development of tools for data manipulation and extractions, such as dedicated bots, or scripts to automate these pipelines, using the Wikidata JSON

Dump filter or the API.

- Transform the glossary into structured data, using the Wikibase Lexeme extension and linking it with Wikidata.

- Exploring other ways of linking *e*Viterbo with Wikimedia Foundation tools for example Wikisource or Wikimedia Commons, continuously developing eViterbo as an example of possible use of these tools in academia.

- Explore the possibility of using Wikibase.cloud as a new home for eViterbo and/or as a mirror or sandbox for the project.

This is a proposal and should serve only as a starting point for future developments.

Reference documentation for MediaWiki, Wikibase and other tools:

MediaWiki Documentation - https://www.mediawiki.org/wiki/MediaWiki
MediaWiki FAQ: https://www.mediawiki.org/wiki/Manual:FAQ
MediaWiki's version lifecycle: https://www.mediawiki.org/wiki/Version_lifecycle
MediaWiki Action API documentation: https://www.mediawiki.org/wiki/API:Main_page

Wikibase Documentation- https://doc.wikimedia.org/Wikibase/master/php/
Wikibase FAQ: https://www.mediawiki.org/wiki/Wikibase/FAQ
Wikibase API documentation: https://www.mediawiki.org/wiki/Wikibase/API

Phabricator: https://phabricator.wikimedia.org/
Gerrit: https://gerrit.wikimedia.org

Extensions, services and tools: Elasticsearch support matrix, UniversalLanguageSelector, Scribunto, WikibaseQualityConstraints, Wikibase Dump Filter, MediaWiki XML Processing, OpenRefine, QuickStatements, Wikidata Query Service