



EuRED : XML-TEI in database structure

François Vignale

► To cite this version:

François Vignale. EuRED : XML-TEI in database structure. TEI Conference 2016, Sep 2016, Vienna, Austria. hal-01483798

HAL Id: hal-01483798

<https://univ-lemans.hal.science/hal-01483798>

Submitted on 7 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

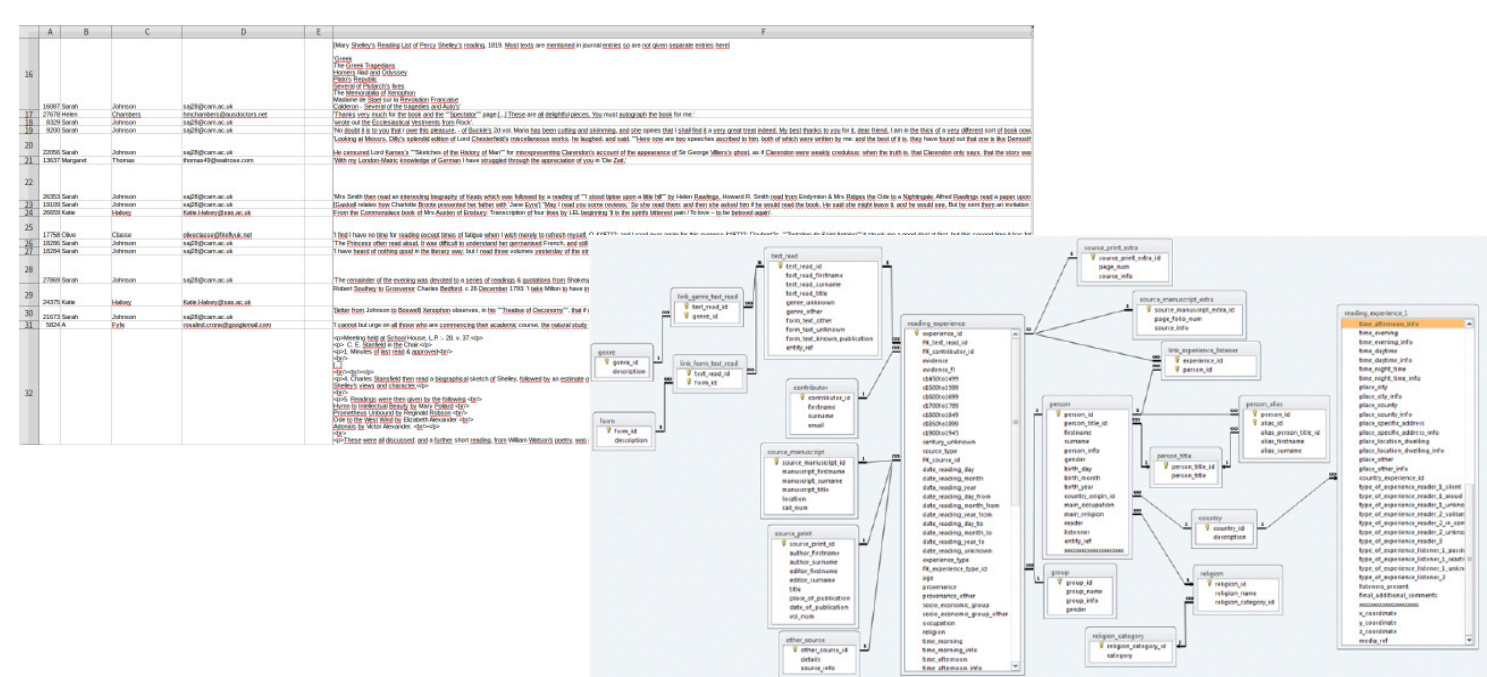
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

EuRED : XML-TEI in database structure

François VIGNALE, 3LAM, Université du Maine, Le Mans, France



The EuRED collects testimonies of European Reading Experiences from the 15th century until today through text mining, public and scholarly crowdsourcing and contribution interface. This multilingual, multimedia, open access and interoperable database, inspired by the 2006 British UK-RED, includes inquiries, geographical and chronological extensions. While addressing new sources (CMCs, audio transcriptions) and interests by the general public, policy makers, educators and scholars, this innovative tool aims to offer access to the evolving journey of reading.

Technical context and specifications

- Proof of concept
- Enrichment of existing data
- Interoperability
- Linked open data
- Open science principles

Methods

- Definition of what is a reading experience
- Work on a new data model
- Choice of CollectiveAccess as CMS
- Choice of XML-TEI for data encoding

XML-TEI in EuRED

- Creation of a new class (model.experienceDescPart)
- Creation of 23 new elements mostly describing the reading experience
- Description of the reading experience placed into the <teiHeader/>
- Use of pointers into <text/> referring to the reading experience

What has been done

- New data model based on XML-TEI
- Creation of specific thesaurii
- Conversion of UK-RED data into TEI
- Linking with referentials (ISO, Geonames ...)

What has to be done

- Full disambiguation
- Automatic enrichment
- Linking with more referentials (ISNI, Dbpedia)
- Crowdsourcing process

More about the project :
<http://eured.univ-lemans.fr>
<http://github.com/EuRED>