Co-funded by the Horizon 2020
Framework Programme of the European Union
Grant Agreement Number 644771

FREME
OPEN FRAMEWORK OF E-SERVICES FOR MULTILINGUAL AND SEMANTIC ENRICHMENT OF DIGITAL CONTENT

Introducing FREME: Deploying Linguistic Linked Data

Felix Sasaki, Tatiana Gornostay, Milan Dojchinovski, Michele Osella, Erik Mannens, Giannis Stoitsis, Phil Ritchie, Kevin Koidl, Thierry Declerck

4th Workshop on the Multilingual Semantic Web
@ ESWC 2015, 1st June 2015, Portoroz

THE FREME PROJECT

- Two year H2020 Innovation action; start February 2015
- Industry partners leading four business cases around digital content and (linked) data
- Technology development bridging language and data
- Outreach and business modelling demonstrating monetization of the multilingual data value chain

Objectives

FREME is building an open innovative commercial-grade framework of e-services for **multilingual and semantic enrichment of digital content**.

By digital content we understand any type of content that exists in a digital form (text, video, audio, image, and others) and that is stored in structured or unstructured, standardized or proprietary formats.

The e-services provided by FREME will be capable to process (harvest and analyze) content, capture datasets, and add value throughout content and data value chains across sectors, countries, and languages.

E-Services

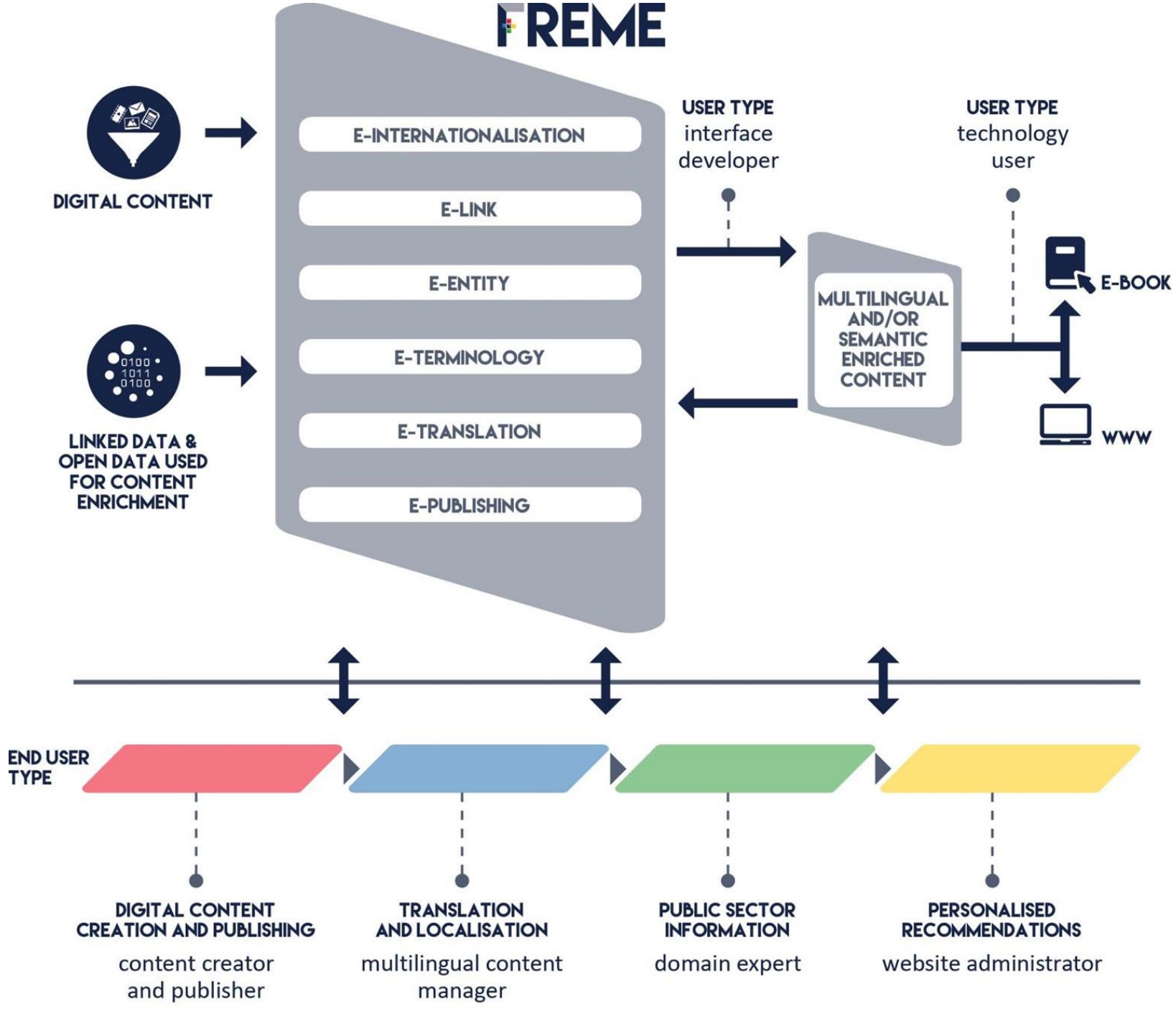
The FREME framework for multilingual and semantic enrichment will integrate the following e-services based on existing and mature technologies:

- e-Internationalisation, to provide metadata for multilingual digital content enrichment. It is based on Internationalisation Tag Set (ITS) version 2.0.
- e-Link, to enrich content with additional information from Linked Data. It is based on the Natural Language Processing Interchange Format (NIF), the DBpedia Ontology and RDF.
- e-Entity, to recognise, link and classify entities in multilingual texts. It is based on entity recognition software and existing linked entity datasets from Linked Data and Open Data.
- e-Terminology, to identify terms for providing domain information and translation suggestion. It is based on cloud terminology services for terminology management and terminology annotation web services.
- e-Translation, to translate content from a source language to a target language. It is based on cloud machine translationservices for building custom machine translation systems.
- e-Publishing to package and export content in the open EPUB3 format, and improve discoverability. It is based on cloud content authoring environment and its export for publishing in standardised Electronic Publication (EPUB3) format.

Business Cases

The FREME e-services are being designed on the requirements provided by four business cases:

- digital publishing, integrating enrichment into publishing workflows with open digital content formats
- translation and localisation, localising intelligent content with increased discovery and multimedia integration
- agricultural & food domain data management, enhancing the cross language sharing and access to open agricultural and food data
- personalisation, providing multilingual and personalised Web content recommendations


FREME
DIGITAL CONTENT
LINKED DATA & OPEN DATA USED FOR CONTENT ENRICHMENT
E-INTERNATIONALISATION
E-LINK
E-ENTITY
E-TERMINOLOGY
E-TRANSLATION
E-PUBLISHING
USER TYPE interface developer
USER TYPE technology user
MULTILINGUAL AND/OR SEMANTIC ENRICHED CONTENT
E-BOOK
WWW
END USER TYPE
DIGITAL CONTENT CREATION AND PUBLISHING
content creator and publisher
TRANSLATION AND LOCALISATION
multilingual content manager
PUBLIC SECTOR INFORMATION
domain expert
PERSONALISED RECOMMENDATIONS
website administrator

Linguistic Linked Data and FREME

- The data value chains that are built with FREME rely heavily on linguistic linked datasets (LLD). The LIDER project is crucial in providing the basis for a linguistic linked data cloud. LIDER fosters LLD as a basis for content analytics tasks of unstructured multilingual cross-media content. FREME e-Services, especially e-Entity, e-Link and E-Terminology, can be seen as prototypical examples of content analytics tasks.

- By providing these services together with e-Translation and several metadata items relevant or translation workflow information (via e-Internationalisation), FREME provides a technology stack that spans across content analytics and machine translation

The relevance of LIDER work for FREME

The relevance of LIDER work on LLD can be seen in three areas:

- creation of LLD data sets,
  - But not all data necessarily in LLD format, but all data items should have a URI (path to LLD). LDD representation essential for exchange but nor for processing.
- best practices on multilingual linguistic linked data
  - For the conversion to LLD resources, off-the-shelf tooling is crucial. In the realm of the BPMLOD group, a TBX2RDF converter has been created. This implementation will help FREME to tackle for example conversion tasks of the data hosted by Tilde in TaaS framework.
- deployment of the LIDER reference architecture.
  - e-Services as LLD aware services. By using NIF as the interchange format between e-Services, FREME provides e-Services as LLD aware services. In the terminology of the reference architecture the e-Services allow to constitute LLD based workflows

Thanks for your attention!!