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

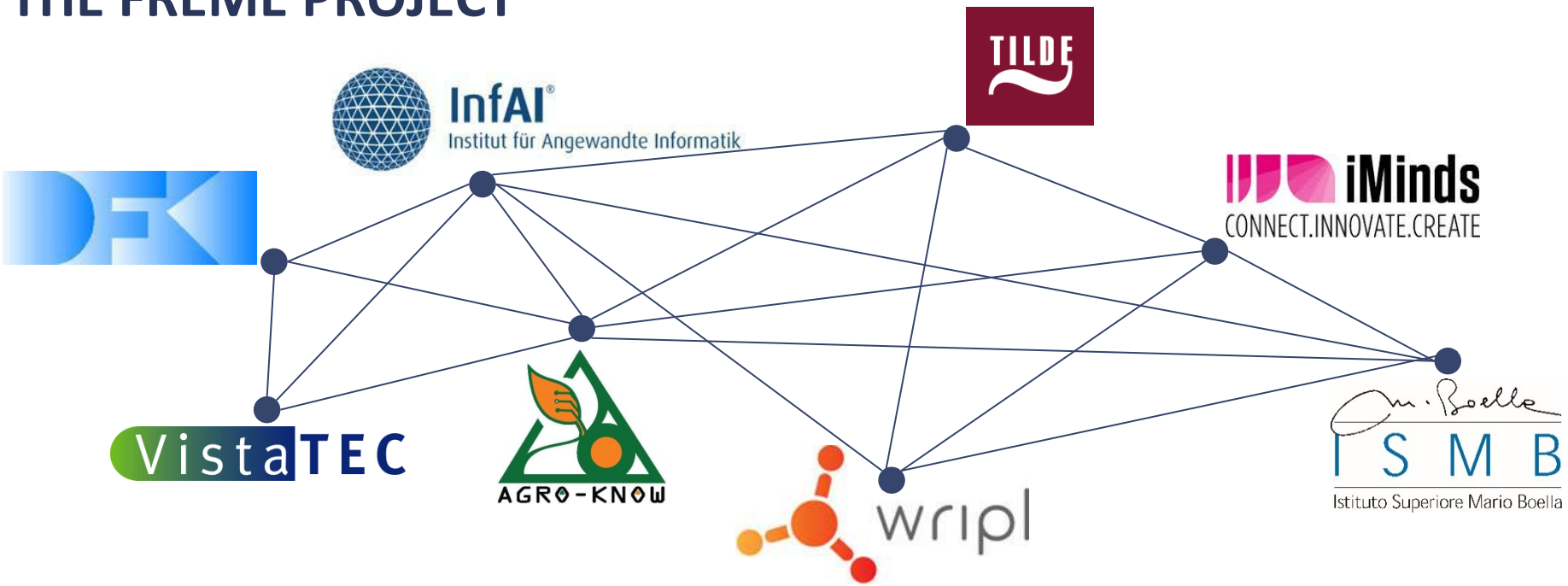
PUTTING LANGUAGE AND DATA STANDARDS AND BEST PRACTICES INTO ACTION: THE FREME FRAMEWORK

META-FORUM, 5 JULY 2016

www.freme-project.eu

Presented by Felix Sasaki (DFKI / W3C Fellow)
On behalf of the FREME consortium

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 & transfer bridging language and data
- Outreach and business modelling demonstrating monetization of the multilingual data value chain

Imagine you want to develop a Web service

Imagine you want to develop a Web service

Would you implement a data transfer protocol yourself?

Imagine you want to develop a Web service

Would you implement a data transfer protocol yourself?

Hopefully not – you would rely on

- **Standards: HTTP**
- **Existing (open source) implementations: e.g. Apache server**

CLAIMS

- Language and data technologies are mature enough to benefit from a common multilingual & semantic interoperability layer
- Don’t try to develop this layer on your own!
- Develop it in an open and distributed manner
- This will benefit your research and business too!

A GOOD EXAMPLE TO REPLICATE

- European Data Innovator Award 2016, going to Ralf Klingenberg, Co-funder of Rapidminer
- Márta Nagy-Rothengass (EC) asking: “What is the secret of success?”
- Ralf: ”I have shared core modules of R.M. ☺”

A GOOD EXAMPLE TO REPLICATE

- European Data Innovator Award 2016, going to Ralf Klingenberg, Co-funder of Rapidminer
- Márta Nagy-Rothengass (EC) asking: “What is the secret of success?”
- Ralf: ”I have shared core modules of R.M. ☺”
- **Important: right level of sharing**
  - **E.g.: Let’s have inline tag handling probably done ☺**

A GOOD EXAMPLE TO REPLICATE

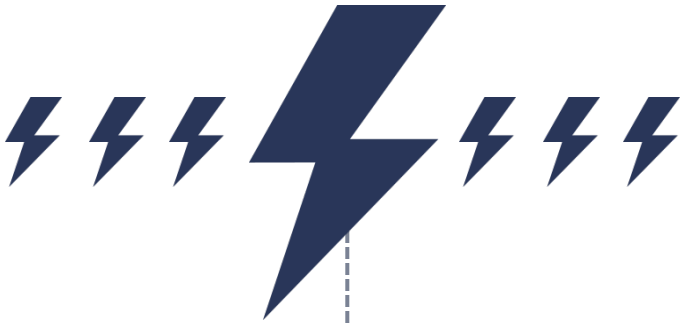
- European Data Innovator Award 2016, going to Ralf Klingenberg, Co-funder of Rapidminer
- Márta Nagy-Rothengass (EC) asking: “What is the secret of success?”
- Ralf: ”I have shared core modules of R.M. ☺”
- Important: right level of sharing

➢ **Base your research and business on standards and open source**

➢ **Pledge for an open framework for LT and data**

CURRENT STATE OF LT AND DATA SOLUTIONS

Machine translation, terminology annotation, ...



Linked data creation & processing

GAPS THAT HINDER Deployment:

- Plethora of formats
- Adaptability and platform dependency
- No easy to use interfaces for LT and LT
- Usability “The right tool for the right person in given and new enterprises”

FREME TO THE RESCUE

Machine translation, terminology annotation, ...

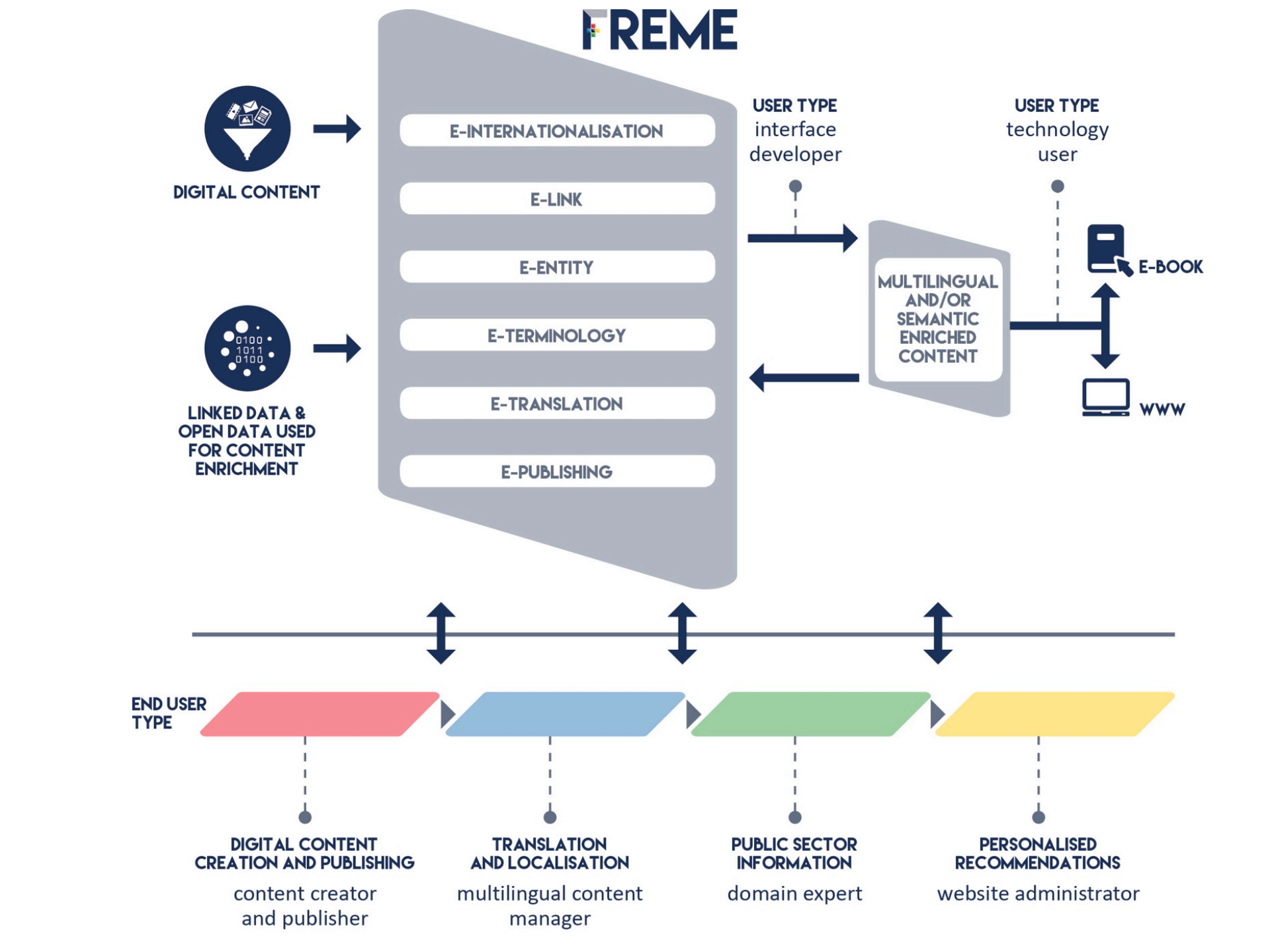
Linked data creation & processing

A SET OF INTERFACES* - DESIGN DRIVEN BY BUSINESS CASES

* Graphical interfaces
* Software Interfaces

LT and LD as first class citizens on the Web

LT and LD for various user types: (application) developer, content architect, content author, …


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

BASIS 1: STANDARDS, FORMATS, RESOURCES

Standards deployed in FREME

ITS 2.0
Metadata to integrate automated processing of human language into core Web technologies

NIF 2.0
Format to achieve interoperability between NLP tools, language resources and annotations

OntoLex / Lemon
Model to publish lexica or terminological data on the web as linked data

Formats for FREME Enrichment

- **HTML5**
- General XML
- Selected **XML Vocabularies**
  - DocBook, TEI, XLIFF, ODF, …
- **Linked Data** as turtle, json-ld etc
- **W3C Web Annotation**
- **More to come – depending on your needs!**

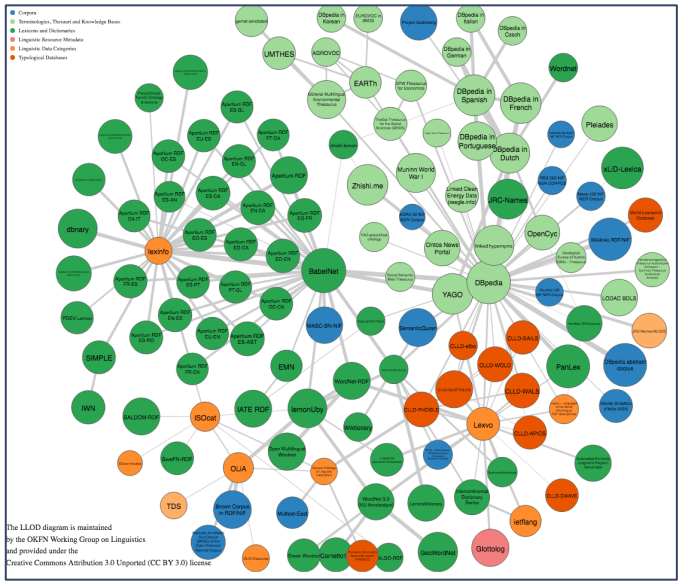
➢ Processing of a growing set of formats via the **Okapi framework**

➢ **Roundtripping** – storing enrichment in the original content

BASIS 1: STANDARDS, FORMATS, RESOURCES

Linguistic Linked Data Cloud

- Corpora
- Terminologies, Thesauri and Knowledge Bases
- Lexicons and Dictionaries
- Linguistic Resource Metadata
- Linguistic Data Categories
- Typological Databases



- Deploying FP7 project results from LIDER and FALCON: Standards, best practices, tooling, …
  basis for decentralized data and language value chains
- Cf. presentation from Ian Wood on MixedEmotions

BASIS 2: DEVELOPMENT GOING BEYOND THE PROJECT

Contributions to the framework – going beyond the project

You here: access via the framework to your R&D outcomes!

BASIS 2: DEVELOPMENT GOING BEYOND THE PROJECT

Contributions to the framework – going beyond the project

Digitale Kuratierungs-technologien

Okapi Framework

You here: access via the framework to your R&D outcomes!

DKT content can be processed with FREME e-Services out of the box – and vice versa
No hard wired integration needed

BASIS 2: DEVELOPMENT GOING BEYOND THE PROJECT

Contributions to the framework – going beyond the project

Digitale Kuratierungs-technologien

Okapi Framework

You here: access via the framework to your R&D outcomes!

Next steps within this year

1) Finalisation of FREME Beta = version 1.0

2) Gather feedback by hundreds of beta users 3) Foster multilingual interoperability topic in **BDVA group on standards and (upcoming) on**

FREME FROM A TECHNICAL PERSPECTIVE

How to access FREME – several options:

- A life version 0.7 including documentation at http://api.freme-project.eu/doc/current/
- A development version at http://api-dev.freme-project.eu/doc/
- A Java / maven software package;
  see the documentation for installation instructions
- Source code in a GitHub project
  https://github.com/freme-project/
- The framework is available under Apache 2.0 license to ease commercial use
- Underlying services have various licensing conditions

CONTACTS

E-mail:

info@freme-project.eu

felix.sasaki@dfki.de

CONSORTIUM

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

PUTTING LANGUAGE AND DATA STANDARDS AND BEST PRACTICES INTO ACTION: THE FREME FRAMEWORK

META-FORUM, 5 JULY 2016

www.freme-project.eu

Presented by Felix Sasaki (DFKI / W3C Fellow)
On behalf of the FREME consortium