Multilingual Value Chain Challenges for the Digital Single Market

LT-Innovate Summit

Brussels, 25 Jun 2015

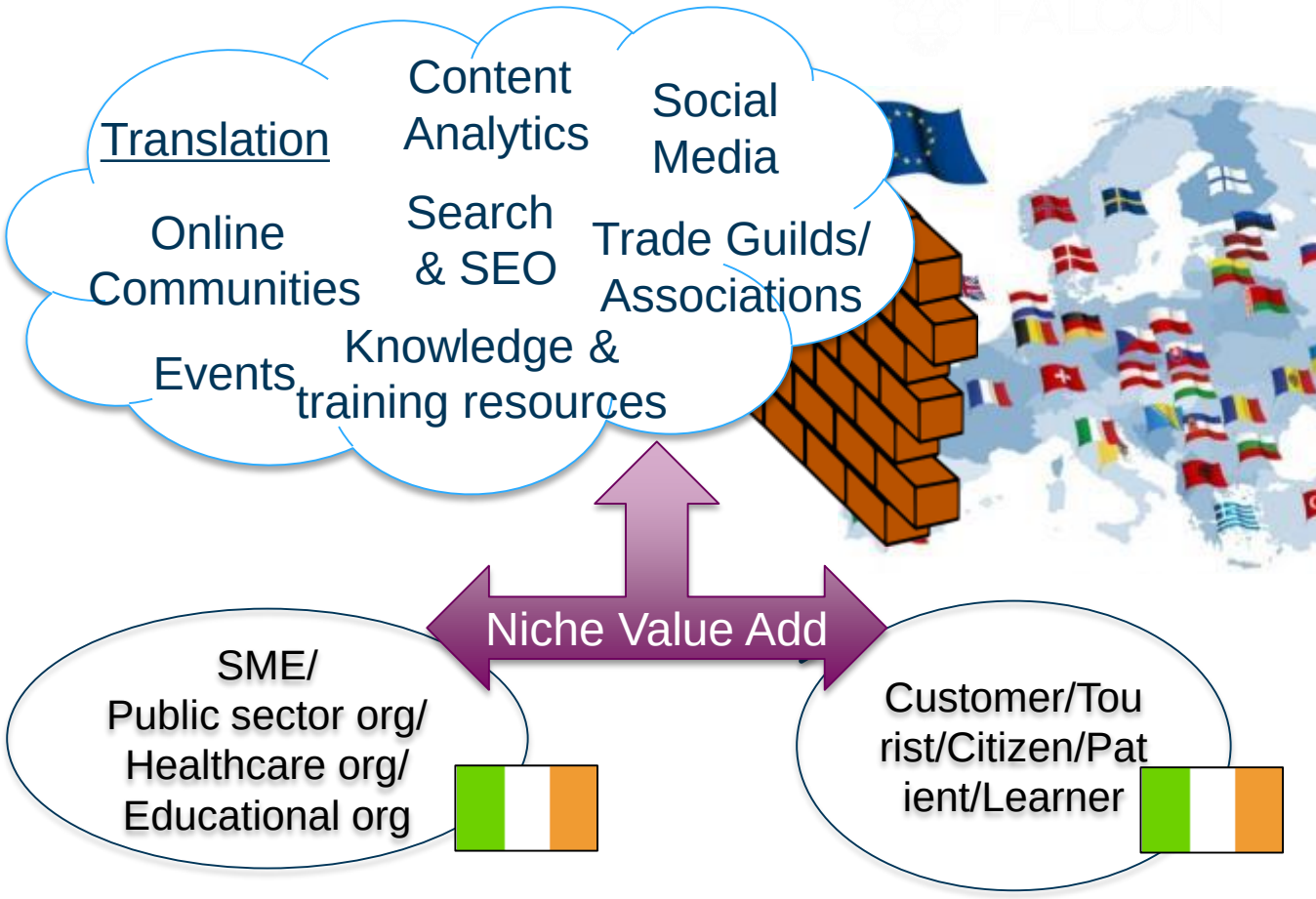
Dave Lewis, ADAPT @TCD
dave.lewis@scss.tcd.ie

Engaging Content
Engaging People

LT for a Digital Single Market

Customer Engagement Ecosystem

- Digital Engagement is **bi-directional, individual and collective**
- **LT** in engagement must allow org **to excel at it niche**
- Must become systematically **multilingual**
- Must serve **micro-domains** that allow Orgs to add value



Challenges matching supply and demand

Technical - lexicons vs. domain terminology/ontologies

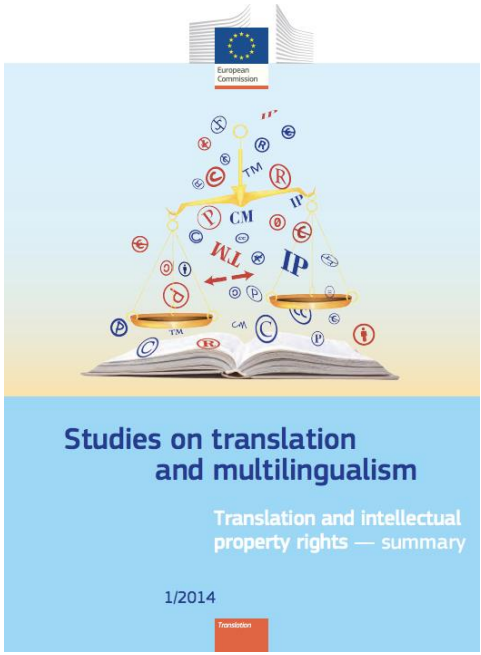
- Needs mechanisms to grow, merge and maintain lexical-semantic graph across value chain
- LT Service Chains must handle overlapping annotations – text and media

Legal – Complex rights and provenance for LR

- Complex matrix of copyright, moral rights, DB Directive, Data Protection, Public sector exemptions
- Uncertain, expensive, ignored?

Business – end-to-end value

- Value creators: authors, linguists, translators, reviewers are poorly incentivised
- LT impact on end user value not measured: conversions, problems solved, knowledge found, advocates nurtured -> A/B Testing

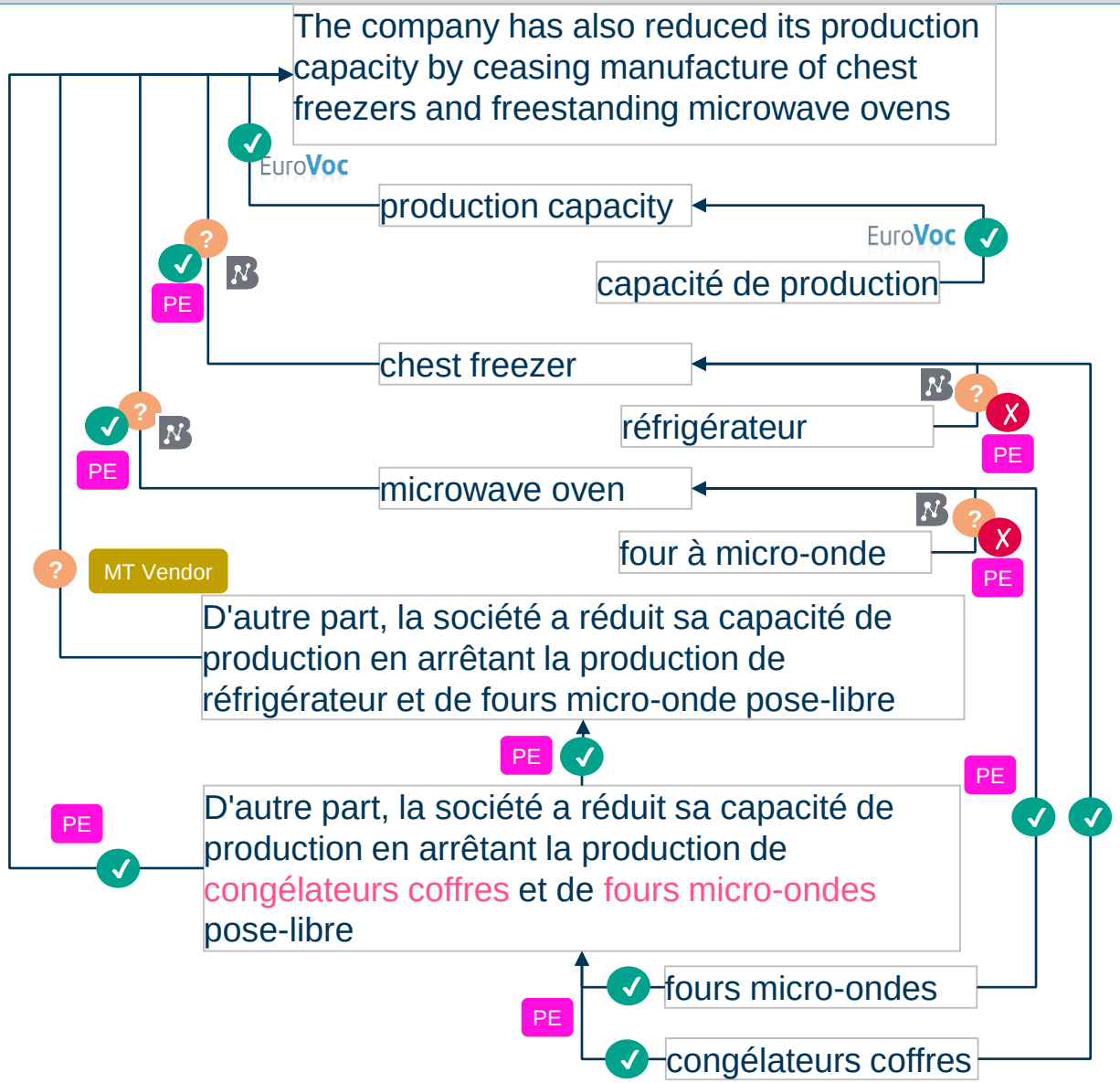


LT Cloud

- Services must integrate easily into existing and future systems
  - Web content management systems
  - Web analytics
  - Site search and SEO
  - Translation management systems
- Open APIs – Basic lexical services
  - Spell checkers
  - Part of speech taggers
  - Sentence splitters
  - Stemmers
- Open Text Analysis Services – lexical-semantic
  - Automated Term Extraction
  - Named Entity Recognition
  - Work Sense Disambiguation

META FORUM 2015

Example: Adding LR value through LT usage



Extraction & Segmentation

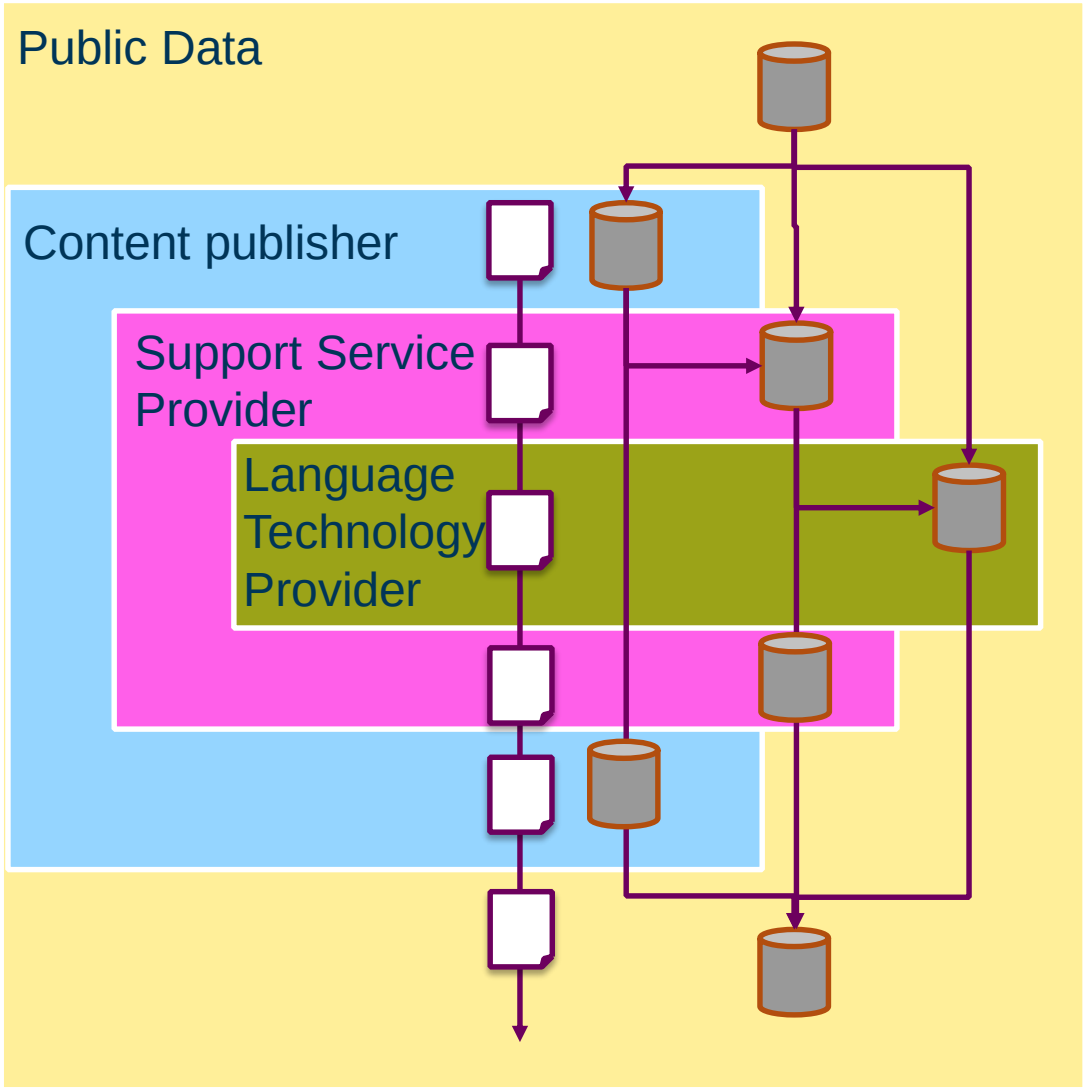
Annotation with Existing Terms

Auto suggestion from Babelfy/Babelnet

Machine Translate with Term Translations

Postedit and capture terms in context

Integrated Content/Data Value Chain



- Protect & pool niche knowledge
- Interlink corpora and lexical-conceptual resources
- Measuring ROI at each point in value chain
- Manage ownership, rights and rewards
- Privacy by Design for Social Media data resources
- Open data for NLP Shared Tasks

- THANK YOU!