Discussion:
Convergence of the Multilingual Web and Linked Open Data
11 June 2012

Dave Lewis
Felix Sasaki
MLW and LOD

Language Resource Community

Language Technology Community

Multilingual Web

Localisation Industry

Draft of the Language Engineering Loci (LOE) Language Engineering Loci (LOE) as of February 2012

Languages

WSDL

WordNet

Lentis

BiGrama

Sociolinguistic Resources

LOHA Association

POMA

Sociolinguistic Resources

Localisation Industry

TM +TB

Multilingual Web

Language Resource Community

Language Technology Community

TM +TB
Roadmap - Opportunities

• Sharing open language resources as LOD
  – Interlinking LR and stand off QA, e.g. For training SMT or Text analytics

• Pooling TM & TB as enterprise linked data
  – Improve data management and integration

• Quality Meta-data when harvesting language resources from the web
  – Provenance annotation

• Localising Linked Data
  – Cross language interlinks
Roadmap - Challenges

• Use and business cases
• Multiple paths:
  – LEMON, NIF, LMF, TBX/RDF....
• Sourcing LR resources
  – Public sector bootstrap + commercial value add?
• Accessibility of LR and LT to SMEs
  – Simplicity, tools, training, best practice
• Progressive access control
  – Enterprise data access to open data access
• Different velocities
  – 2 yrs = 0.5 phd = 8 quarterly reports
Discuss?