



XLIFF Workflow and Multilingual Provenance in Linked Data

David Lewis

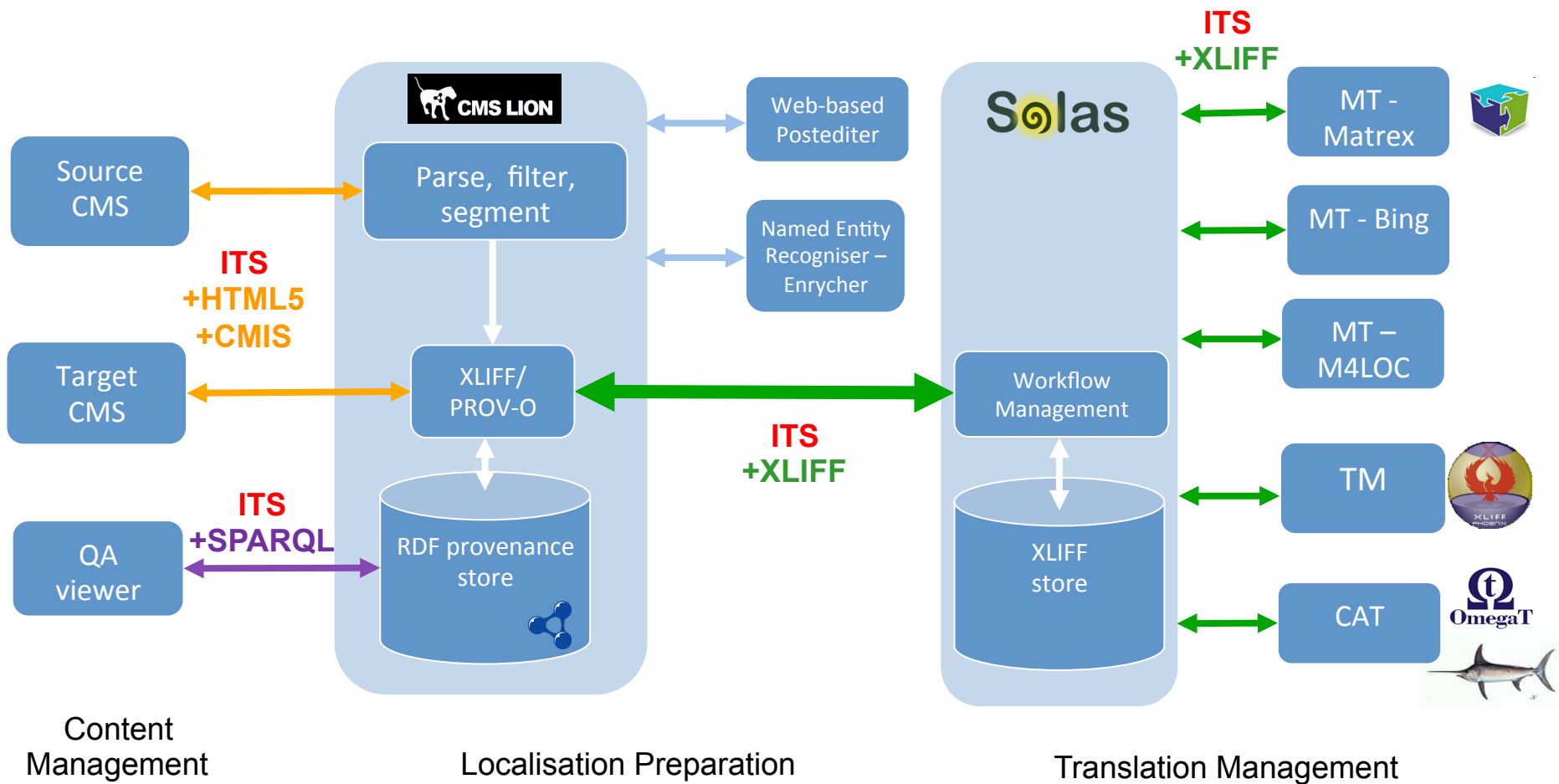
CNGL at Trinity College Dublin

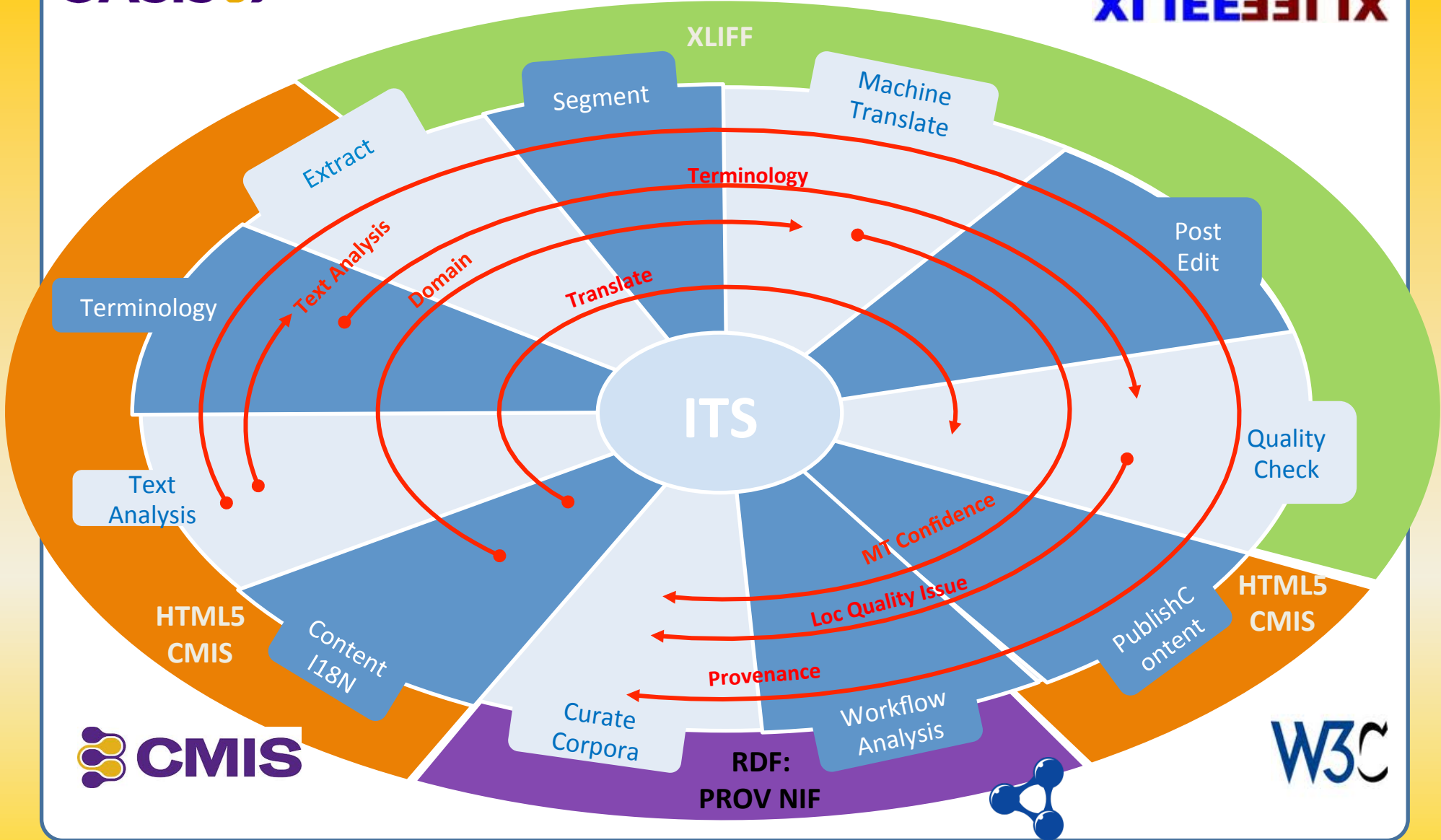
Co-chair W3C MLW-LT WG

Curation of Language Resources

- Language Resource Curation
 - Relies on specific funded actions
 - Episodic funding, sustainability and coverage concerns
- LOD promises:
 - Improved sharing and annotation of existing resources
 - Some opportunistic curation, e.g. mining Dbpedia
- Can there be better synergy with commercial language services?
 - Use case: Localisation workflows

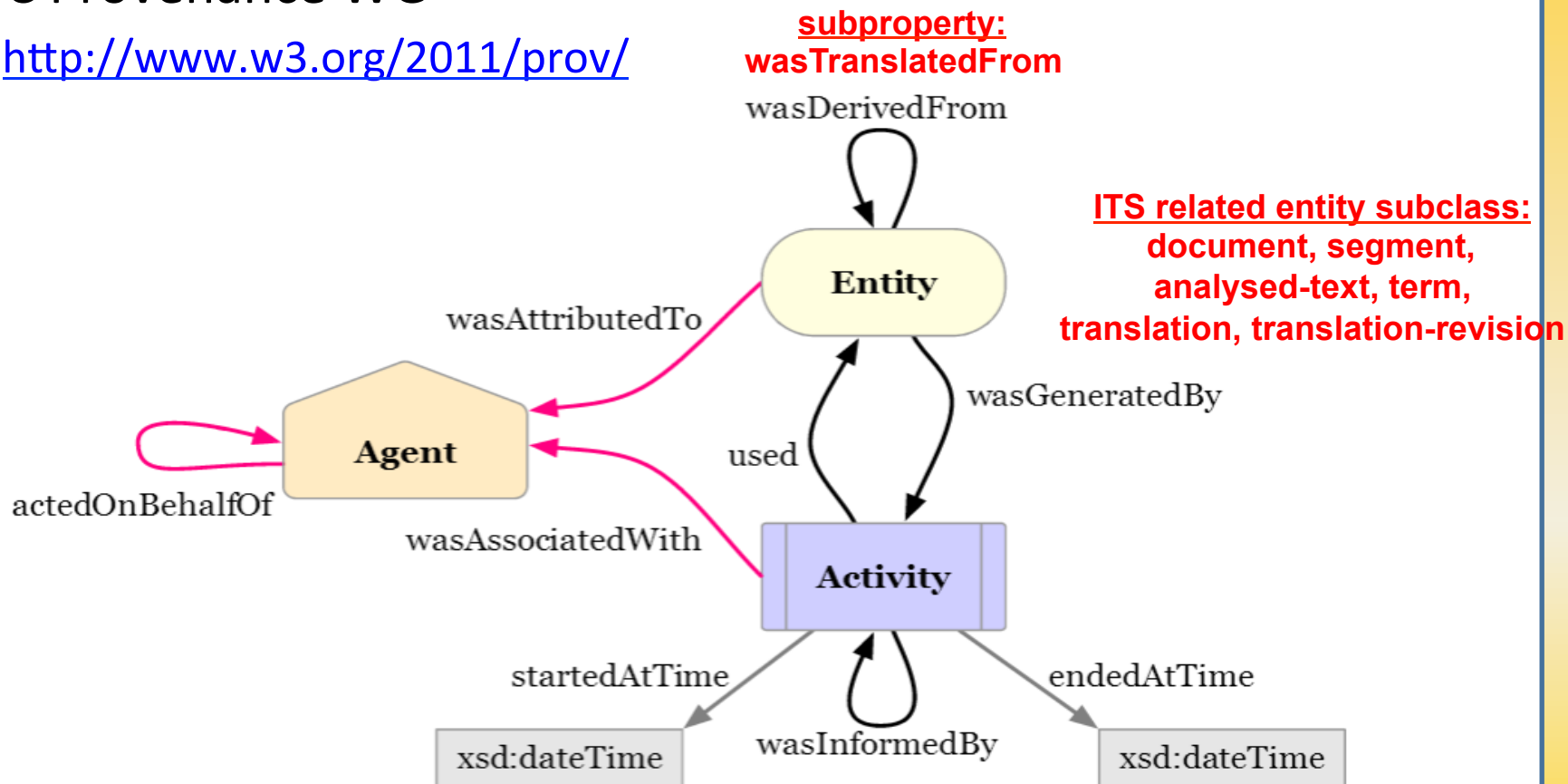
System Architecture





Provenance Data: RDF-based logging

- W3C Provenance WG
 - <http://www.w3.org/2011/prov/>



From: <http://www.w3.org/TR/prov-primer/>

The MultilingualWeb-LT Working Group receives funding by the European Commission (project name LT-Web) through the Seventh Framework Programme (FP7) in the area of Language Technologies. Grant Agreement No. 287815.

Active Curation of L3Data

- Generate Linked Language and Localisation Data ('L3Data')
- Pooling to boost EU's many SME language service providers
- Encourage commercial users/generators/curators of L3Data
- EC and other public bodies could bootstrap open data
- Current Gaps:
 - Standards based instrumentation and analytics – ITS
 - Attribution/license annotation and access control
 - Open source tool chain - linked data generating CAT/TMS/CMS
 - More academic-industrial collaboration on use cases
- Follow ITS-XLIFF-PROV-O Use Case at:
 - http://www.w3.org/International/multilingualweb/lt/wiki/Provenance_Best_Practice

Questions?