qtleap

quality translation with deep language engineering approaches

European Commission
FP7, Call 10
Strep
#610516
http:/qtleap.eu

List of partners	
	Faculty of Sciences, University of Lisbon, Portugal (coordinator)
	German Research Centre for Artificial Intelligence, Germany
	Charles University in Prague, Czech republic
	Bulgarian Academy of Sciences, Bulgaria
	Humboldt University of Berlin, Germany
	University of Basque Country, Spain
	University of Groningen, The Netherlands
	Higher Functions, Lda, Portugal

Project duration: November 2013 — October 2016

Summary

The incremental advancement of research on Machine Translation (MT) has been obtained by encompassing increasingly sophisticated statistical approaches and fine-grained linguistic features that add to the surface level alignment on which these approaches are ultimately anchored.

The goal of this project is to contribute for the advancement of quality MT by pursuing an approach that further relies on semantics and opens the way to higher quality translation.

We build on the complementarity of the two pillars of language technology — symbolic and probabilistic — and seek to advance their hybridization. We explore combinations of them that amplify their strengths and mitigate their drawbacks, along the development of three MT pilot systems that progressively seek to integrate deep language engineering approaches.

The construction of deep treebanks has progressed to be delivering now the first significant Parallel DeepBanks, where pairs of synonymous sentences from different languages are annotated with their fully-fledged grammatical representations, up to the level of their semantic representation.

The construction of Linked Open Data and other semantic resources, in turn, has progressed now to support impactful application of lexical semantic processing that handles and resolves referential and conceptual ambiguity.

These cutting edge advances permit for the cross-lingual alignment supporting translation to be established at the level of deeper semantic representation. The deeper the level the less language-specific differences remain among source and target sentences and new chances of success become available for the statistically based transduction.