

Five discussion lead points for the LTO Terminology charrette, Vienna, July 2015

LTO wishes to consult with members of the terminology and special language community to learn about the availability and usability of “formal terminology” in training and enhancing MT systems (largely statistics-based) in any European language today.

LTO is a project tasked with mapping the public availability of language resources (LRs) and enhancing their usability in practical translation situations involving machine translation by various public authorities in the coming years. It plans to deliver a rich information source about such LRs and help streamline non-specialist access to them.

In this respect, LTO wishes to look at the following five “discussion lead points” concerning terminology. We call this resource “formal” terminology because unlike the text documents in many LRs, terminologies are complex human artefacts that embody coded forms of knowledge: terminology as language data that has already been processed into accessible structures for humans (using various technologies). The critical issue is to understand how to ensure this artefact can become an easily-integrated resource for machine processing.

Point 1: Finding terminology LRs: where, how centralised, public/private availability?

Point 2: Evaluating terminology LRs: who, how, why, which standards?

Point 3: Maintaining terminology LRs: who, how, where?

Point 4: Technical features: formats, management systems, standards

Point 5: New terminology developments/practices on the horizon that might impact MT usability: research projects etc.