# Language Technology and the Publishing and Media Industries

## Innovation Partnering Focus Group

12 April 2013
09:00 to 14:00
Berlin, Germany

## *Venue*

Ahoy Berlin, Windscheid Str. 18 (zweites Hinterhaus)
D - 10627 Berlin
Tel: +49 173 7079 148

## *Purpose*

The purpose of the LT-Innovate Focus Groups is to launch a dialogue between Language Technology vendors and buyers in key segments of today's content industries that are massively shifting to digital agendas. This dialogue should provide feedback and input of two sorts:

a)  Insight into what the buyer industry (in this case publishing/media) needs and expects from LT in the next 5 to 10 years. This will help LT vendors prepare development plans or reflect on new approaches to staying competitive;

b)  Information for the publishing/media industry on what LT vendors are already capable of producing.

There should thus be learning on both sides, but the publishing/media industry will be encouraged to talk as frankly as possible about how they *evaluate* LT today, where the strengths and weaknesses are in current technology, what software or other language-related standards (if any) are required, and what changes are needed.

The Focus Group will not be a one-off event, but marks the launch of a process over the entire year, with additional highlights, such as follow-up webinars, Innovation Focus sessions at the LT-Innovate Summit on 26-27 June in Brussels and another Focus Group event in the second part of the year. Additionally, LTi will be publishing the outcome of the vendor/buyer dialogue in the 'LT Industry Innovation Agenda' which will be distributed to all participants and interested parties towards the end of the year.

## *Key LT focus points for publishing/media*

There are various 'functions' in a publishing/media company that could benefit from LT. The idea would be to ask media/publishers to look ahead 5 to 10 years to the key touch points in their value chain, and find out what needs they foresee that must to be satisfied for them to stay competitive in a business environment where

global mobile, social media, cloud (software as a service) and multilingual big data become the names of the game.

The following is a set of notes and questions that touch on LT 'contact points' for a) the industry as a whole, b) the publishing/media enterprise as an organisation, and much more importantly, c) the media/publishing player in the business of selling content.

1. **Across the publishing/content industry**

   a. Does the publishing/media industry have professional organisations that examine digital technology futures for the various types of publishing?
   b. If so, is language technology in general mentioned in these discussions, reports, etc?
   c. More generally, how is technology purchasing organised by the industry? Are there committees or other forms of shared knowledge gathering about technology?

2. **Inside the enterprise** (not specific to this industry)

   a. LT can help to optimise workflows in various ways in any company
      i. Speech tech for improving inputting at different workstations
      ii. Are there internal translation needs for the enterprise?
      iii. Do they use social media analytics to track their company reputation, sentiment about the company products etc?
      iv. Do they need to speed up document production – authoring tools, compliance with standards etc?
      v. Are they interested in looking at the role of LT in workplace optimisation and CRM in general?

3. **As part of the company's key products** (content in the publishing industry)

   a. Making documents and content easier to search as knowledge containers
      i. Metadata added automatically to documents, pages,
      ii. 'Semantic' searching
      iii. Document classifiers (so that you can automatically find documents on a given subject)
      iv. Are standards in metadata an issue?

   b. Automating knowledge collection and publication:
      i. Automatically searching for information, discovery
      ii. Organising it into readable structure
      iii. Correcting, editing and publishing with minimum human intervention
      iv. Are there plans to integrate video with 'book' publishing? If so what needs might this create for speech technology (and translation)?
      v. Will user generated knowledge play a role in publishing strategies and if so how can language tech help improve quality etc?

   c. Adding value to content with speech and video/image technologies
      i. Text to speech to make content 'speakable' (for e-readers ?) or rather,

      ii.   How do tablets and e-readers impact the technology stack for publishers?

      iii.  Adapting content to mobile/tablet formats for specific types of use (educational, infotainment, etc). What challenges will ubiquitous tablets/e-readers raise?

           1.   Will publishing involve the making of special apps for mobiles?

      iv.  How will publishing/media evolve with respect to introducing video and speech content sources into standard content products?

   d.   <u>Making content easier to translate</u>

      i.   How important is translation as a *process* to publishing/media?

           1.   Is translation to be done "at source" - i.e. by the publisher, or do they expect people to use Google Translate for free on a website?

           2.   If to be kept in house, which languages and how much volume do they expect to be handling in 5 or 10 years?

           3.   If translation is kept in-house, will they outsource to a translation company, or make their own translation engine customised to their specific domains of knowledge? If so what competences do they need?

      ii.   Tools for improving the readability standards of content to simplify translation

---

## *Expected take aways*

For <u>buyers:</u>
- A clearer, richer picture of the opportunities that LT can bring their industry as a whole, and some aspects of their company processes in particular;
- A reference point for finding out more about the importance of LT in digital futures.

For <u>vendors:</u>
- A clearer, richer picture of the kinds of issues that publishers face in their work;
- A sense of how the publishing industry evaluates or regards LT in general and some specific products;
- A picture of possible touch points each individual buyer could pursue as an innovation scenario.