

## **LT-Innovate Dinner**

**26 June 2013**

**Roberto Cencioni**

**European Commission, DG CONNECT**

5 years ago, when I was brought back into the language business of INFISO, I started with my colleagues a process meant to raise the profile of language and speech technologies within European programmes.

That process has yielded more than 60 projects and a few collaborative forums, in particular LTi and META. Quite a few more projects are due to start later this year, some of them will be presented tomorrow.

As you may know, META was designed to bring closer together a rather diverse research community, to look at core technologies and their current gaps and weaknesses, to set priorities and define a widely-accepted research path towards a new generation of systems.

LTi as you know since you are here, was (and still is) geared towards commercial players, technology developers and vendors/integrators, the aim being to create the first business forum cutting across traditional language-processing families and industrial sectors.

The simple idea behind it was to stimulate a productive dialogue between you people, between you and your actual or prospective buyers, between you and your sponsors/investors, ..., a dialogue to be translated into a meaningful innovation agenda – highlighting collaboration and commitment ways beyond European funding.

The end-result was to be a strong coalition between research and industry that my institution - and indeed national agencies - would recognize as a trusted partner under future programmes.

However, something else happened, which provides an even better opportunity for you, and the field at large.

While language (and speech) processing was a relatively small, often neglected domain, Big Data became very rapidly a popular theme for ICT, administration, media gurus and politicians alike, something that Information Management never managed to do in 30 years.

So last year the Commission decided to bring language (and multi-language) matters under the Big Data wings, which in my view provides a much-needed boost and recognition for what you do.

Data, open data, big data... are now recognized in a growing range of EU programmes. And the fact that data happens to come in different languages has finally surfaced. Very much the same way as handling unstructured content, extracting knowledge, making sense of social media... using – also – language processing techniques.

So I sincerely believe that the new set-up namely supporting *research into and use of Big Data*, in particular *across borders and languages*, does provide an excellent opportunity for you all, in business terms as well as in relation to European programmes.

Tomorrow my colleagues, the deputy-director general of DG CONNECT and the head of the Value Data Chain unit, will outline where we stand and where we intend to go in the next year or two. So I will not spoil the surprise.

But let me give some basic information:

- the new R&I programme, H2020, is now approaching adoption following extensive discussions between the European Parliament, Council & Commission
- the ICT part of H2020 which is most relevant for your business, the so-called LEIT i.e. Industrial Leadership is advancing well; the 2014-2015 budget envelope is nearly final; the workprogramme is available in a draft form and will be presented at the ICT 2013 conference planned for early November, after discussions with the MS
- as we speak, the budget provisionally earmarked for the "data / content / media / knowledge / language" *conglomerate* amounts to some 200 Meuro over two years, more than half of which is allocated to i) Big Data, ii) Automatic Translation and iii) Multimodal Interaction. The Innovation element is well present, support is foreseen for SMEs
- so you see that the 3 strands that we tried to bring together within LTI, namely information management + automated translation + speech[+], are well represented

- but the fight is not over, neither for the start of the programme, where discussions continue and national delegations will try and influence the budget allocation, nor for later programme operations, after 2015
- lastly, let me mention another programme of some relevance here. A few weeks ago the EC has transmitted (to EP and CO) an amended proposal for the Telecommunications part of the Connecting Europe Facility; the core of the programme are the so-called Digital Service Infrastructures, which are to enable cross-border exchanges and cooperation through EU-funded platforms and services
- this is a much smaller programme, with 1 billion euro over 7 years. the proposal includes a number of "building blocks" upon which pan-European exchanges can be built, and one such building block is automated translation (most likely a mix of MT and CAT); contrary to H2020, the purpose is not to develop new technologies, but rather to deploy operational services either directly or within hosting online services
- the discussion of the proposal by the EP and MS representatives will start in a few days. We will know in the autumn what it becomes. Meanwhile the EC services have started looking at the implementation of the programme.

So, let me conclude by saying that

although a dedicated focal point for language and speech technologies does not appear any more in the CONNECT organisational chart,

you will not be left out in the cold, the new Data paradigm provides excellent opportunities, both within our DG and in other parts of the Commission.

you'll just need to remind my colleagues from time to time, so they don't forget you!