|  |  |
| --- | --- |
| **Working Group B:  Audio/Video description and spoken captions** | **Coordinators:  Pilar Orero, Universitat Autonoma de Barcelona, Spain Aline Remael, University College Artesis, Belgium E-mail: pilar.orero[at]uab.cat E-mail: aline.remael[at]artesis.be** |

The structure of the Input Document from each WG

The scope and objectives of input document should lead to an explanation of how the WG is going to answer the following 4 questions:

**A. What is the state of the art of Audio/Video description and spoken captions?**

* + The value network and the key stakeholders
  + Existing work practices in the area/domain
  + Examples of good practice

**B. What is the vision for 2015 and 2020 for Audio/Video description and spoken captions?**

* + A scenario which explains how the needs and interests of persons with disabilities (in its broadest sense) would be addressed
  + The design and production processes
  + The key technologies
  + The business models underpinning accessible digital media
  + The legislative and regulatory frameworks needed to support these visions

**C. What are the barriers in 2011 that currently prevent these visions from becoming a reality?**

* + Barriers related to currentdesign and production processes
  + Barriers related to currenttechnologies
  + Barriers related to current business models for accessible digital media
  + Barriers related to current legislative and regulatory frameworks needed to support these visions
  + Other barriers (add those you feel are missing)

**D. What actions are needed to break down current barriers in order to make the 2015 and   
 2020 visions a reality?**

* + Actions related to currentdesign and production processes
  + Actions related to currenttechnologies
  + Actions related to current business models for accessible digital media
  + Actions related to current legislative and regulatory frameworks needed to support these visions
  + Other actions (add those you feel are missing)

**Annex 1  
A hypothetical example of the kind of issues to be addressed by Working Group A - Captioning**

**A. What is the state of the art of Captioning?**

* + *The value network and the key stakeholders*

The document is expected to contain a Value Network diagram for captioning (intralingual subtitles) and foreign language captioning (interlingual subtitles), both pre-prepared and live captioning, indicating the main stakeholders.

* + *Existing work practices in captioning*

The document describes how you would adapt and extend the descriptions in the forthcoming ITU-D report “Making Television Accessible” – the latest draft can be downloaded here: <http://dl.dropbox.com/u/4655124/Making_TV_Accessible_V09.doc>

Note any significant differences in practices from one country to another.

* + *Examples of good practice*

The document contains references to good specifications, papers, reports etc that the WG feels are examples of good practice and worthwhile emulating elsewhere.

**B. What is the vision for 2015 and 2020 for Captioning?**

* + *A scenario which explains how the needs and interests of persons with disabilities (in its broadest sense) would be addressed*

The scenario addresses significant demographic changes foreseen in the period to 2015 and 2020. My own vision for 2020 is the following:

“Viewers can watch compelling programs on a TV screen, a phone or any other digital device they have at their disposal. There are captions for *all* same language and foreign language programming, including live broadcasts. The broadcaster has produced captioning at source that is close to a verbatim transcription. The viewer can select captions that match interest and capability. Those who need and want captions with considerable language condensation can select this to match their required reading speed. The on-screen presentation offers a default that can be changed by the viewer to match her needs and requirements (size, contrast and color of the text). The caption display time can be increased, as R&D to seamlessly vary the delivery speed of on-screen voices without increasing the playing time of the program has been widely deployed. TV programming is now easier to follow for the eldest citizens. Immigrants can use captioning as a means of improving their command of the national language(s). Prime-time programming is increasingly offered with foreign-language options to promote social cohesion.”

* + *The design and production processes*

“The captioning process still requires the human touch but many of the basic processes build on speech recognition, intelligent mark-up of the text and the ability to parse the captions in the receiver under viewer control to match their needs and expectations. Captions accompanying a given program can be readily exchanged as they all are subsets of the same W3C-TT standard.”

* + *The key technologies*

Stenography has almost completely been replaced by re-speaking solutions that are available for national and some regional languages around the globe. Intelligent parsing powers the delivery of captions in television receivers, computers, smartphones and tablets”

* + *The business models underpinning accessible digital media*

Basic captioning is so widely used that the technologies to display them are available in all mainstream ICT products including TVs, computers, smartphones and tablets. Intelligent parsing is on the increase, but is currently available as assistive technology on demand.

As captioning still requires a combination of speech recognition technology and human “captioners”, various business models including public subsidies, cross-funding from accessible advertising and sponsorship are in use.

* + *The legislative and regulatory frameworks needed to support these visions*

There are clear targets covering the supply and perceived quality of captioning. National consultation panels agree on the Key Performance Indicators that cover not only targets approaching 100% of programming on main channels but also metrics for user awareness, use and enjoyment of captions.

**C. What are the barriers in 2011 that currently prevent these visions from becoming a reality?**

* + *Barriers related to current**design and production processes*

No agreed metrics for discussing what constitutes “quality” in relation to same-language and foreign language captions. Where explicit metrics exist, these are often simplistic in that they equate quality with verbatim transcriptions.

* + *Barriers related to current**technologies*

Speech recognition technologies are not available for many national and almost all regional languages. Public-private partnerships are needed to do the basic work on language modelling so that these language models can be developed and maintained, and then licensed to commercial and public players in the captioning field.

* + *Barriers related to current business models for accessible digital media*

Poor awareness of business models for devices and captioning, in particular the implications of captioning for special interest channels and/or the provision of captioning in small or poor national markets.

* + *Barriers related to current legislative and regulatory frameworks needed to support these visions*

A lack of evidence-based targets, metrics and key performance indicators reached by consensus by the main stakeholders.

* + *Other barriers (add those you feel are missing)*

**D. What actions are needed to break down current barriers in order to make the 2015 and 2020 visions a reality?**

* + *Actions related to current**design and production processes*
  + *Actions related to current**technologies*

National and regional public-private partnerships to put in place language models for speech recognition systems that power re-speaking systems.

* + *Actions related to current business models for accessible digital media*
  + Actions related to current legislative and regulatory frameworks needed to support these visions
  + Other actions (add those you feel are missing)