International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

FG AVA TR

Version 1.0
(10/2013)

Focus Group on Audiovisual Media Accessibility

Technical Report

Part 3: Using audiovisual media – A taxonomy of participation

FOREWORD

The procedures for establishment of focus groups are defined in Recommendation ITU-T A.7. The ITU-T Focus Group on Audiovisual Media Accessibility (FG AVA) was proposed by ITU-T Study Group 16 for creation in-between TSAG meetings and it was established on 22 May 2011. The Focus Group was successfully concluded in October 2013.

Even though focus groups have a parent organization, they are organized independently from the usual operating procedures of ITU, and are financially independent. Texts approved by focus groups (including Technical Reports) do not have the same status as ITU-T Recommendations.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Technical Report may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU-T Focus Group participants or others outside of the Technical Report development process.

Table of Contents

Summary

This Technical Report of FG AVA was prepared by Working Group F “Participation and digital media” and it presents a glossary that FG AVA had foreseen as a future work item for ITU-T Study Group (SG16) “Multimedia” and ITU-R SG6 “Broadcasting service”. This Technical Report is taxonomy of participation within the context of audio visual media. It considers stakeholders related to production, transmission and dissemination of media and uses that to consolidate a set of case studies of participation. The case studies lead to taxonomy of participation that is used to identify requirements from various access services and assistive technologies.

Keywords

Accessibility, adaptable interface, inclusive design, multimodal interface, participation taxonomy, personalization, usability, user model.

1 Participation taxonomy

Older adults are the fastest growing demographic group, while currently around 15% of the total world's population, or roughly 900 million people, live with a disability. Modern research in intelligent interactive systems can offer valuable assistance to the elderly and people with disabilities by helping them to engage more fully in the world. However, many users find it difficult to use existing interaction devices because of physical or aging-related impairments.

As an example of digital exclusion, statistics show about 70% users between 65 and 74 years old had never used Internet and 39% cannot use mobile phones in European countries. Existing research on intelligent voice recognition, adaptable pointing, browsing and navigation, effect and gesture recognition can hugely benefit them.

Additionally, systems and services developed for the elderly or people with disabilities often find useful applications for their able-bodied counterparts – a few examples are mobile amplification control, which was originally developed for users with hearing problems but helpful in noisy environments, audio cassette version of books originally developed for blind users, and standards of subtitling in television for deaf users and so on. Furthermore, many important technical achievements could not yet be implemented at the industrial level, mostly due to lack of awareness among industrial developers, and missing software and guidelines support during design and development. Existing research on intelligent systems often works for ‘average’ users and excludes a certain portion of the population who finds it difficult to use the existing systems and may benefit from intelligent adaptation of the interface.

However, accessibility practitioners and other computing professionals often fail to understand each other and come up with the wrong solution. Lack of knowledge about the problems of people with disabilities and elderly users has often led designers to develop non-inclusive systems. On the other hand, accessibility research often focuses on developing tailor-made products for certain types of disability and lacks portability across different platforms and users.

The ITU-T FG AVA Working Group F on participation media analysed the present situation and proposed a vision for 2020 and a plan to attain this vision. At first, this working group aimed to scope the terms ‘digital media’ and ‘participation’. The term digital media covers a wide array of digital interactive devices and services including:

– Devices:
  - digital television;
  - computers;
  - gaming consoles;
  - smartphones and so on.
– Services:
  - television broadcast;
  - software;
  - web-based systems;
  - hybrid TV applications;
  - computer games;
  - social networking services and so on.

The term participation includes:

– authoring;

– interacting; and

– disseminating contents.

In the context of accessibility, the main target users will be elderly people and users with different ranges of abilities consisting of visual, cognitive, auditory and motor impaired users. However, neo-literate or illiterate users and immigrants – with a meaning of users speaking a foreign language - are also considered in this text.

Considering the scope of the terms, the following key stakeholders were identified:

– people with disabilities/elderly people;

– suppliers of audiovisual media;

– suppliers of interaction devices/assistive technologies;

– academic/research institutes;

– legislators and regulators.

It will be necessary to look at the actual value networks for the media chosen and the kinds of participation that are included.

A generic value network diagram is shown in Figure 1.

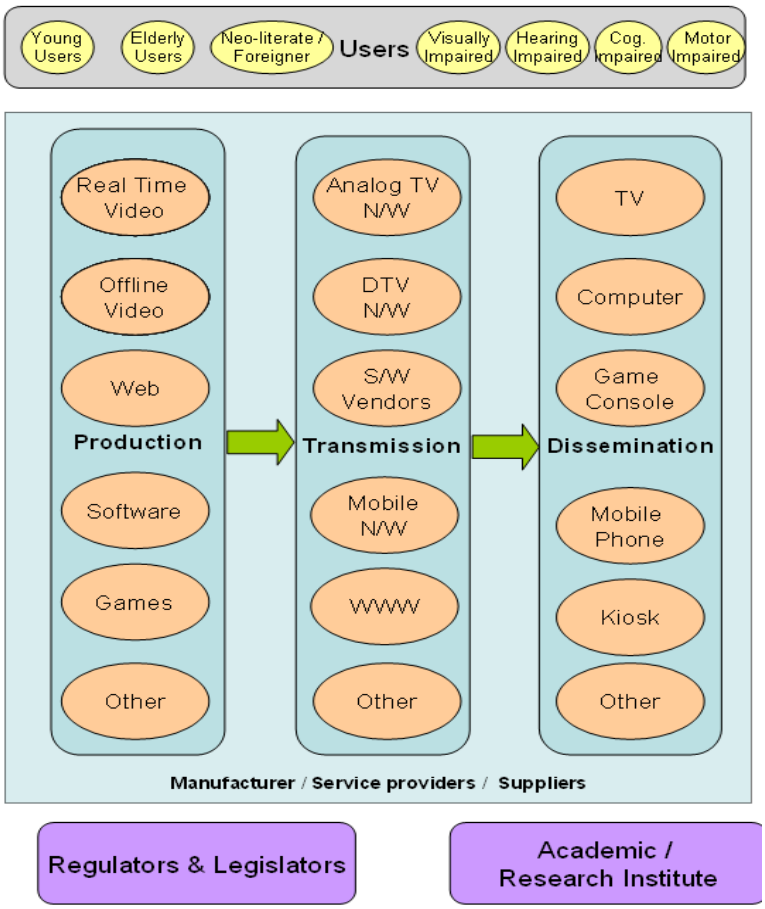


Figure 1 – Value network diagram

2 Analysis based on taxonomy

The value network diagram depicts a picture of the whole spectrum of participation. It has been used to identify a set of participation scenarios (Box 1) and in light of that, propose taxonomy (Box 2) to identify issues with existing accessibility products and services.

Based on these participation scenarios, a set of questionnaire, as shown in the example below, has been prepared:

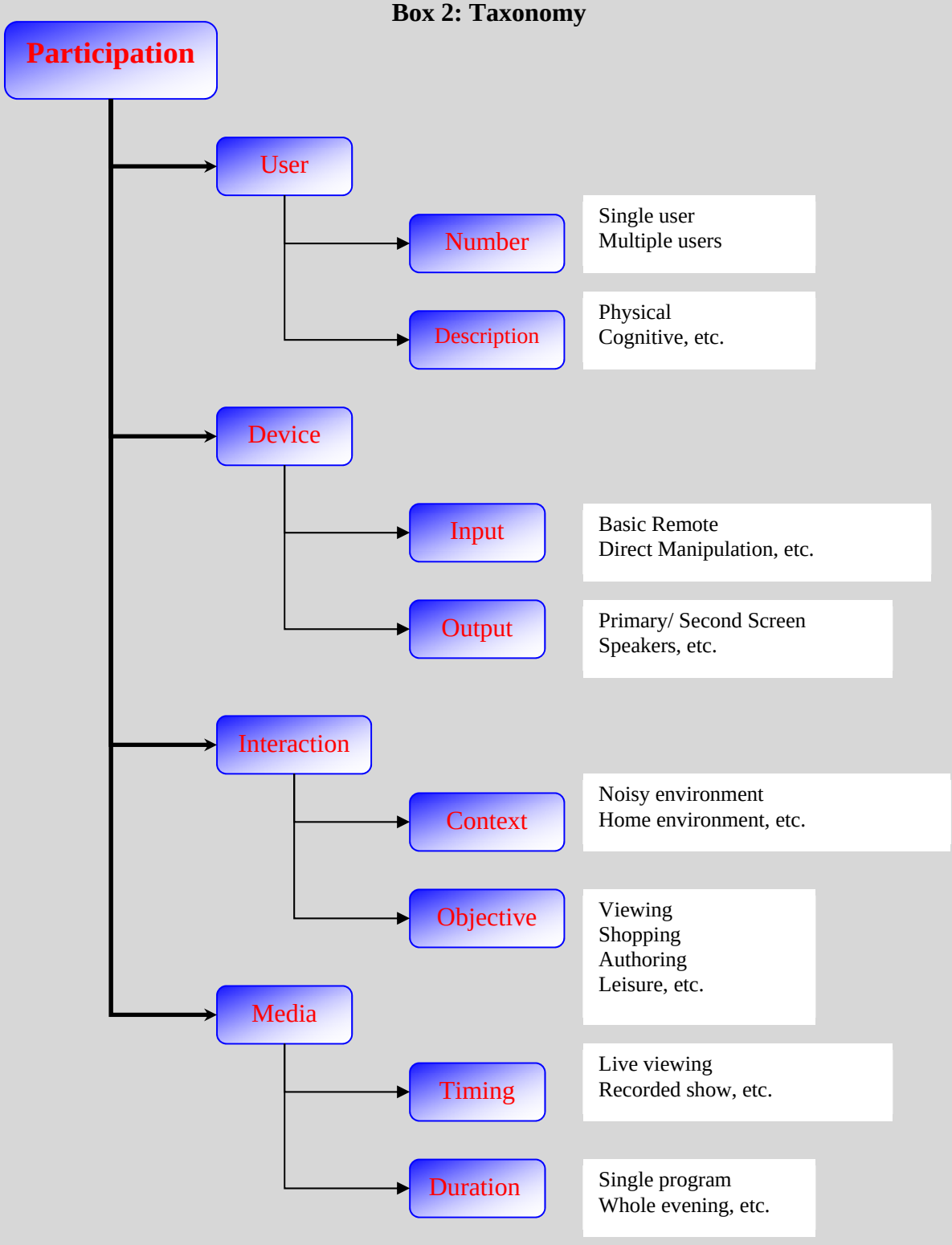
- How many people are watching: One, or more than one?
- What are they watching: A single programme (and then switch off); More than one programme?
    - If more than one programme, is it: A series of programmes viewed sequentially; More than one programme simultaneously.
    - Are the programmes on now, but not "live" (e.g. the scheduled broadcast time for a soap); On now and live (e.g. a football match); Previously recorded (e.g. on a digital video recorder/digital versatile disk (DVR/DVD), etc.).
    - How is the user participating in the viewing: Watching only; Voting (e.g. X Factor); Shopping; Asking a question; Getting further information (e.g. via text); Selecting viewing options (e.g. camera angles for a football match).
- Where is the user: At home in the living room; At home in the kitchen; etc.
- What equipment is being used: Large screen television (TV); personal computer (PC); mobile phone; iPlayer, etc.

The set of questionnaire is visualized in the chart in Box 2.

Box 1: Participation scenarios

1. Viewer alone watching news in his drawing room
2. Viewer watching news with multiple viewers and one of them requires subtitles
3. Viewer changing settings of a digital TV to turn on audio caption
4. Viewer watching a live show and participating in a voting process with basic remote control
5. Viewer watching a live show and participating in a voting process with mobile phone and image recognition technology
6. Viewer watching a live show in a connected TV and shopping over Internet using basic remote
7. Viewer watching a live show in a connected TV and shopping over Internet using direct manipulation with a gyroscopic remote
8. Viewer watching a live show in a connected TV and tweeting about it over social media using a second screen
9. Viewer using an electronic program guide with hand gestures
10. Viewer using an electronic program guide using a second screen on a tablet
11. Viewer recording a program from a live TV show
12. Viewer watching a movie in 3-dimension
13. Viewer playing a game in a connected TV with augmented reality system
14. Viewer using a video conferencing application in a hybrid TV
15. Viewer discovering content in a connected TV
16. Viewer authoring content for a tele-learning application to be run on a hybrid TV

Box 2: Taxonomy



This taxonomy is used to analyse existing accessibility services and compare and contrast them. Subtitles or captions, audio captions and visual screen modifications were considered as a starting point. The following clauses identify issues with these with respect to taxonomy. In the future, accessibility features can be considered and therefore the analysis can be extended.

2.1 Caption

1. **Multiple users**: Caption will be fine with multiple users even though someone does not need it. This depends on whether the viewers in a country where captioning is something with which they are familiar from foreign language programmes[1]. It depends also on the amount of text that has to be read, and the speed with which the captions are refreshed.
2. **Impaired user**: If a user has visual impairment, the font size and colour contrast needs to be adjusted. If a user is a slow processor of information or a speaker of a foreign language, the delay in providing caption needs to be predictable. It should also stay on the screen for a longer time. There are already well-established norms based on empirical studies that govern this. The norms vary from one country to another, and also from one medium (TV) to another (movies).
3. **Input device**: It would be good if the input device like a TV remote can be used to change the size of the font.
4. **Output device**: Captions are suitable for large screen displays; the font size needs to be adjusted for small screen displays such as portable TV or laptop screen based on viewing angle, screen size and resolution.
5. **Context**: Colour contrast may need adjustment if the ambient light is too bright or may be bright like a public display facing sunlight.
6. **Objective**: If a caption bears important information, it should use highlighting either using bigger fonts or different colours to attract attention. While the user is participating in a particular purpose such as online shopping or checking the weather using captions, highlighting or different colour contrasts should be used as in written media for important features and steps (as in during the confirmation of an electronic transaction or alerting about a natural disaster and so on).
7. **Timing**: The delay should be predictable and small during live viewing.
8. **Duration**: The duration of the sequence should not affect if the delay is predictable overall.

2.2 Audio caption

1. **Multiple users**: Studies have shown that non-deaf family members do not feel disturbed with audio caption while watching TV with a deaf member. Audio captioning is often also helpful for people with cognitive impairment besides visual impairment.
2. **Impaired user**: It is suitable for the blind user, but if blind users have also age related or other form of hearing impairment, they may not hear certain words properly; in this case, the speaker or the producer of the audio caption may need digital signal processing for digital compression and/or a spectral contrast enhancement.
3. **Input device**: Users needing audio caption may not use a keypad (such as a keyboard or a remote control) or direct manipulation devices (such as a mouse, a trackball, a gesture controller, etc.) due to visual impairment. Therefore, audio caption should be accompanied by a voice input system with good speech recognition capability, if possible in the users' native language.

[1] This is not the case in dubbing countries.

4. **Output device**: Speakers with some in-built digital signal processor for noise cancellation. Transmitting the main signal and background signal in different channels (e.g. Swedish public radio) may be useful as users can adjust the signal-to-noise (SNR) themselves.
5. **Context**: The speaker volume needs to be adjusted in a noisy environment. The volume should not always be higher, which may create a worse hearing experience due to loudness recruitment; good noise cancellation and multi-channel compression algorithms should be used instead (e.g. the European Broadcasting Union (EBU) recommendation for Loudness R 128 normalization for public and private broadcasters).
6. **Objective**: If a caption bears important information, it should be repeated or slowly spoken. While the user is participating in a particular purpose such as online shopping or checking the weather using audio captions, special features like slow or repeated messages or audio icons should be used for important features and steps (as during the confirmation of an electronic transaction or alerting about a natural disaster and so on). If the speaker has any in-built digital signal processors, those should be adjusted based on purpose, for example, a noise cancellation system will create a worse experience while listening to music.
7. **Timing**: The delay should be predictable and small during live viewing.
8. **Duration**: The duration of the sequence should not affect if the delay is predictable overall.

2.3 Screen modifications (font size, colour contrast adjustment, visual magnifier)

1. **Multiple users**: While multiple users have different requirements, careful adjustment of font size, colour contrast, types of input and output modalities, sizes of menus and buttons in screen are required. An algorithm for merging user profiles based on user models should be used to resolute conflicts in requirements. The resolution also depends on the purpose of viewing – for example, profile merging will be attained in different ways for public displays and home entertainment systems.
2. **Impaired users**: Visual adjustments are mainly suitable for partially sighted users. However, certain visual adjustment may be avoided by carefully designing the screen, for example, using bigger font, and suitable colour contrast and so on. Simulation of visual impairment and adjusting font size and colour contrast for mild to moderate visual impairment will be useful and should be used during the development of application interfaces.
3. **Input device**: If the user uses a direct manipulation device (such as a mouse, a trackball or a gyroscopic remote), the pointer size, colour and speed should be adjusted based on the visual adjustment required by the user.
4. **Output device**: For small screen devices, visual adjustment may be difficult for lack of screen space. Alternate modalities such as voice output or tactile feedback can be invoked in similar situations.