



Question(s): 2/20

Geneva, 18-28 July 2022

## TD

Source: Editors

Title: Output text of draft Recommendation ITU-T Y.ACC-UI-req “Accessibility requirements for user interface of smart applications supporting IoT”, Q2/20 meeting (Geneva, 18-28 July 2022)

**Contact:** Yong J Lee  
Center for Accessible ICT  
Korea (Rep. of)

Tel: + 82 70 4907 4208  
E-mail: [ylee@caict.re.kr](mailto:ylee@caict.re.kr)

**Contact:** Hark Sohn  
SCE Korea Inc.  
Korea (Rep. of)

Tel: +82 70 4354 2015  
E-mail: [daniel.hark.sohn@gmail.com](mailto:daniel.hark.sohn@gmail.com)

**Contact:** Andrea J. Saks  
G3ict  
United States

Tel: +44 1242 820 800  
E-mail: [andrea@andreasaks.onmicrosoft.com](mailto:andrea@andreasaks.onmicrosoft.com)

**Abstract:** The draft Recommendation ITU-T Y.ACC-UI-req has been developed to define “Accessibility requirements for user interface of smart applications supporting IoT” since it was accepted as a new work item in the Q2/20 meeting, Virtual, 17-27 May 2021. This draft Recommendation is aiming to specify accessibility requirements for user interfaces of smart applications supporting IoT for persons with disabilities, persons with age-related disabilities and those with specific needs to utilize the benefits of IoT services.

This is the updated output document of draft Recommendation “Accessibility requirements for user interface of smart applications supporting IoT” (Y.ACC-UI-req), output of Q2/20 meeting, Geneva, 18-28 July 2022.

It is generated based on TD2485(GEN/20) output of Q2/20 meetings, Virtual, 11-21 October 2021, and the meeting agreements on the following contribution:

No.	Source	Title	Proposal	Discussion and results
C24	Korea (Republic of)	Y.ACC-UI-req on “Accessibility requirements for user interface of smart applications supporting IoT”, proposed modification of Section 7	The contribution proposes to modify sections 7.1 and 7.2.	It was agreed with modification.

The proposed update is attached as below:

## **Draft new Recommendation ITU-T Y.ACC-UI-req**

### **Accessibility requirements for user interface of smart applications supporting IoT**

#### **Summary**

The use of IoT may increase quality of life among persons with disabilities, persons with age-related disabilities and those with specific needs when properly designed. There are many possible IoT services in various environments that provide accessibility services as well. The IoT also can be used to create tools for persons with many types of disabilities and specific needs, including physical, visual, hearing and cognitive disabilities. IoT service interact with user through the user interface. To ensure an IoT service be accessible, the user interface must be accessible. Accessible user interface must take into account a user's physical, audio and visual capabilities, and consider compatibility with any assistive technology used by the user. This Recommendation outlines essential requirements that a user interface must consider in order to secure the accessibility of the smart applications.

#### **Keywords**

Accessibility, Persons with Disabilities, Smart Applications, User Interface, Universal Design  
*[TBA]*

#### **Introduction**

The concept of universal design, as addressed in the UN Convention on the Rights of Persons with Disabilities (UNCRPD), is that the design of services shall be usable by all people, to the greatest extent possible. It specifically stresses the concept of inclusive design, without the need for adaptation or specialized design. However, the universal design should not exclude assistive devices for particular groups of persons with disabilities where this is needed. This means that all the smart applications supporting IoT must be designed to accommodate both persons with and without disabilities. However, in many IoT services, it is a reality that developers are not creating smart applications that support the needs of persons with disabilities due to lack of knowledge on accessible application or preconceived notion that it is not very necessary. This Recommendation outlines essential requirements that a user interface must consider in order to secure the accessibility of the smart applications.

## Table of Contents

	Page
1 Scope.....	4
2 References.....	4
3 Definitions.....	4
3.1 Terms defined elsewhere.....	4
3.2 Terms defined in this Recommendation .....	5
4 Abbreviations and acronym .....	5
5 Conventions .....	6
6 Accessible smart applications to support IoT .....	6
7 User interface accessibility requirements for smart applications.....	7
Appendix I.....	13

*[TBD]*

## Draft new Recommendation ITU-T Y. ACC-UI-req

### Accessibility requirements for user interface of smart applications supporting IoT

#### 1 Scope

This Recommendation specifies accessibility requirements for user interface of smart applications supporting IoT.

The scope of this Recommendation includes:

- accessibility requirements for user interfaces of smart applications supporting Internet of things (IoT) in order for persons with disabilities, persons with age-related disabilities and those with specific needs to utilize the benefits of IoT services.

#### 2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

- [ITU-T F.790] Recommendation ITU-F.790 (2007), *Telecommunications Accessibility Guidelines for Older Persons and Persons with Disabilities*.
- [ITU-T Y.4000] Recommendation ITU-T Y.4000/Y.2060 (2012), *Overview of Internet of Things*.
- [ITU-T Y.4204] Recommendation ITU-T Y.4204 (2019), *Accessibility requirements for Internet of things applications and services*.

#### 3 Definitions

##### 3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

**3.1.1 accessibility** [b-ITU-T F.791]: The degree to which a product, device, service or environment (virtual or real) is available to as many people as possible.

**3.1.2 accessibility feature** [b-ITU-T F.791]: An additional content component that is intended to assist people hindered in their ability to perceive an aspect of the main content.

**3.1.3 assistive technology** [b-ITU-T F.791]: Piece of equipment, product system, hardware, software or service that is used to enable, maintain or improve functional capabilities of individuals with disabilities.

**3.1.4 Internet of things (IoT)** [b-ITU-T Y.4000]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

NOTE 1 – Through the exploitation of identification, data capture, processing and communication capabilities, the IoT makes full use of things to offer services to all kinds of applications, whilst ensuring that security and privacy requirements are fulfilled.

NOTE 2 – From a broader perspective, the IoT can be perceived as a vision with technological and societal implications.

**3.1.5 person with age-related disabilities** [b-ITU-T F.791]: A person with cognitive or physical disabilities caused by the aging process. Examples are impaired eyesight, deafness in varying degrees, reduced mobility or cognitive abilities.

**3.1.6 person with specific needs** [b-ITU-T F.791]: Includes persons with disabilities (PWDs), persons who are not literate, those with learning disabilities, children, indigenous people, older persons with age-related disabilities, and anyone who has a temporary disability.

**3.1.7 platform accessibility feature** [b-ITU-T F.791]: Accessibility functionality provided as standard on a particular hardware or software platform.

**3.1.8 profile setting** [b-ITU-T F.791]: The ability for users to store and retrieve multiple profiles containing sets of user interface preference settings without having to reset them each time, including accessibility settings.

**3.1.9 screen reader software** [b-ITU-T F.791]: Software application used by a person who is blind or cannot easily read print to identify and interpret what is shown on a video display and read aloud using speech synthesis.

**3.1.10 specific needs** [b-ITU-T F.791]: This replaces the use of the term 'special needs'. This term refers to a wide range of categories including women, children, youth, indigenous people, older persons with age-related disabilities, persons with illiteracy, as well as persons with disabilities (PWDs), see [b-ITU PP Res.175], [b-WTDC Res.58], and [b-WTDC AP] and clause 6.39 (of [b-ITU-T F.791]).

**3.1.11 universal design** [b-ITU-T Y.4211]: The design of products, environments, programmes and services to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design. "Universal design" shall not exclude assistive devices for particular groups of persons with disabilities where this is needed.

NOTE – Paraphrased from [b-UNCRPD].

## **3.2 Terms defined in this Recommendation**

*None*

## **4 Abbreviations and acronym**

This Recommendation uses the following abbreviations and acronyms:

AAC	Alternative Augmentative Communication
ASR	Automated Speech Recognition
AT	Assistive Technology
ICT	Information and Communication Technology
IoT	Internet of Things
PWD	Persons with Disabilities
TTS	Text-to-Speech

## UI                    User Interface

### 5    Conventions

In this Recommendation:

The expression "**is required to**" indicates a requirement that must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The expression "**is recommended**" indicates a requirement that is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

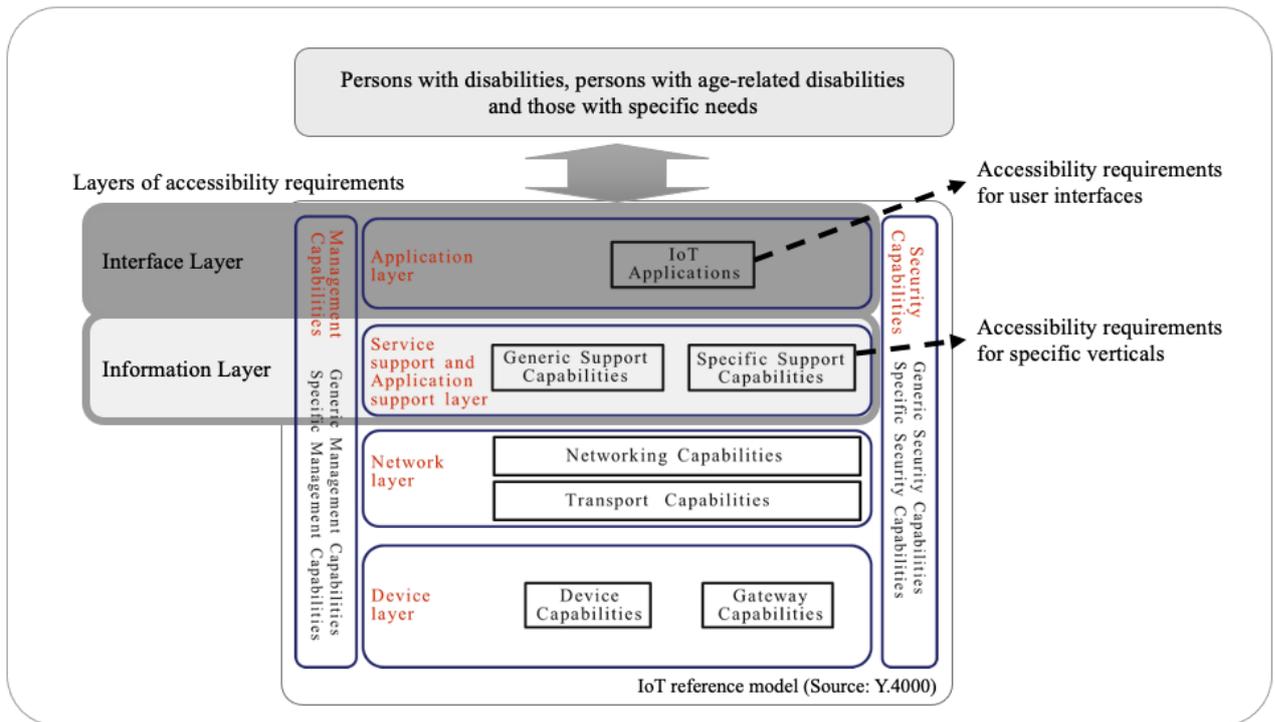
The expression "**can optionally**" and "**may**" indicates an optional requirement that is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with this Recommendation.

### 6    Accessible smart applications to support IoT

The concept of universal design, as addressed in the UN Convention on the Rights of Persons with Disabilities (UNCRPD) [b-UNCRPD], is that the design of services shall be usable by all people, to the greatest extent possible. It specifically stresses the concept of inclusive design, without the need for adaptation or specialized design. However, the universal design should not exclude assistive devices for particular groups of persons with disabilities where this is needed. This means that all the smart applications supporting IoT must be designed to accommodate both persons with and without disabilities.

However, in many cases, it is a reality that developers are not creating smart applications that support the needs of persons with disabilities due to lack of knowledge on accessible application or preconceived notion that it is not very necessary. This Recommendation outlines essential requirements that a user interface must consider in order to secure the accessibility of the smart applications.

As identified in Recommendation ITU-T Y.4204 [ITU-T Y.4204], there are many possible IoT services in various environments that provide accessibility services such as home automation services, IoT for work environments, transport services, etc. The use of IoT may increase quality of life among persons with disabilities, persons with age-related disabilities and those with specific needs when properly designed. The IoT also can be used to create tools for persons with many types of disabilities and specific needs, including physical, visual, hearing and cognitive disabilities.



**Figure 1 Concept of Accessible IoT services**

The figure 1 depicts the concept of accessible IoT services. As introduced in Recommendation ITU-T Y.4211 [ITU-T Y.4211], concept of two-layer structure of accessible IoT service is presented here. The information layer of accessible IoT service is mainly concerned within the service support and applications support layer of IoT reference model [ITU-T Y.4000]. The accessible information should be prepared concerning the user's physical, audio, visual, linguistic and cognitive abilities. For public transport, accessibility requirements are addressed in [ITU-T Y.4211].

On the other hands, the interface layer of accessible IoT service, is mainly concerned with the application layer of IoT reference model [ITU-T Y.4000]. When IoT service prepares accessibility information, it should be presented to the user through the user interface. To ensure an IoT service be accessible, the user interface must also be accessible. Accessible user interface must take into account a user's physical, audio and visual capabilities, and consider compatibility with any assistive technology used by the user. This Recommendation directly addresses the requirements for user interfaces of smart applications supporting IoT.

Smart applications may include, but not limited to, applications for mobile, tablets, wearables, embedded panel and interactive digital signage that are part of IoT services. However, physical interface is not of interest although it may be a part of IoT services.

[TBA]

## **7 User interface accessibility requirements for smart applications**

In this clause, user interface accessibility requirements are listed. Requirements are categorized according to [ITU-T Y.4204].

## 7.1 UI requirements to perceive information

This clause identifies requirements of UI on users' ability to perceive information and that the users equally recognise all application contents regardless of disabilities.

- 1) Alternative text is required to be provided along with non-text content.

NOTE 1 - Alternative text is a textual substitute for non-text content. The alternative text should be specific and succinct. The alternative text may be converted into a voice using screen reader software and be used as a basis for sign interpretation.

- 2) Both caption and sign interpretation are required to be provided along with the audio contents or visual contents with auditory information.

NOTE 2 - A display with a small screen, such as a wearable device or an embedded panel, may exceptionally omit the sign interpretation.

NOTE 3 - Since it is difficult for persons with hearing disabilities to recognize information provided by audio, caption or sign interpretation is required. It is necessary to provide both because each person may have different specific needs.

- 3) Caption and sign interpretation are required to maintain its physical size large enough for a user to recognize it.

NOTE 4 - When caption and sign interpretation are displayed on a small screen, securing a size that the user can sufficiently recognize is essential. In addition, visual content and its accompanying captions and sign interpretation should be arranged together to be recognized on one screen.

- 4) Caption and sign interpretation are required not to obscure important information on the screen.

NOTE 5 - Caption and sign interpretation need to have separate spaces to not overlap with visual content. Still, if they need to be overlapped due to screen size constraints, it must be ensured that important visual information is displayed well.

- 5) When text appearing on the screen in visual content has meaning, a separate voice commentary explaining the text content is required to be provided.

NOTE 6 - In particular, when the text information displayed on the screen requires a separate explanation, it is necessary to explain it rather than just stating it.

- 6) Automated speech recognition (ASR) captioning is recommended to be used only in limited situations where a human captioner cannot.

NOTE 7 - ASR captioning cannot wholly replace human captioners and should be regarded as complementary, not alternative.

- 7) Sign interpretation using avatars is recommended to be used only in limited situations where real human sign language interpreters cannot.

NOTE 8 - Sign intervention using avatars cannot clearly express nonmanual signals, which are essential elements of sign language, and have limitations in not being interactive. Like ASR, it should be considered complementary, not alternative.

- 8) It is required to provide a voice commentary explaining visual content provided without audio or text information.

NOTE 9 - Since it is difficult for persons with visual disabilities to recognize visual information of visual contents, voice commentary is required if there is text information on the screen that is not explained by audio information.

- 9) All information displayed on the screen is required to be recognizable regardless of colour.

NOTE 10 - All information provided by the content should be made recognizable to users who cannot distinguish a particular colour, users of black-and-white displays, users who view black-and-white prints, and users in high contrast mode.

- 10) All information displayed on the screen is required to be kept in a sufficient brightness contrast ratio so that foreground and background colours can be distinguished.

NOTE 11 - All user interface components and text displayed on the screen are required to be provided with a brightness contrast ratio of 4.5:1 or higher so that foreground and background colours can be distinguished [b-Yan].

- 11) Instruction for a user to take a specific action is required to be provided distinctly even when the user utilizes alternative technologies.

NOTE 12 - When users recognize information using alternative text (i.e., using screen readers), the names of terms and components used in the alternative text should be consistently provided so that instructions can be clearly communicated. For example, when creating buttons for user consent, the name of the buttons should be consistent throughout the applications. Possible candidates of the button name, in this case, are "ok", "yes" and "accept". However, only one of them should be used consistently within an application to avoid confusion.

- 12) Instruction for a user to take a specific action is recommended to be recognizable regardless of shape, size, location, orientation, colour or sound.

NOTE 13 - In particular, it is essential to ensure consistency so that users with various specific needs can easily understand instructions when using recognized and clearly identified symbols.

- 13) Notification information is required to be provided in various ways, such as using screen display, sound, vibration, etc.

NOTE 14 - For users with cognitive disabilities, simultaneous use of various senses such as hearing, vision and touch for important notification information is preferred. In addition, to meet the diverse specific needs of different users, it is desirable to provide notification information so that the user may select a method suitable for the user.

- 14) When providing audio content, the appropriate volume and speed are recommended to be maintained.

NOTE 15 - It is not suitable for the user to continuously adjust the volume to identify the audio content. It is also necessary to maintain an appropriate audio content level so that it is not too fast or too late. To this end, the opinions of various users must be reflected.

15) When providing audio content, it is recommended to minimise noise.

NOTE 16 - It is necessary to minimise noise, especially for users who use hearing aids, headsets, etc.

## 7.2 UI requirements to understand information

This clause identifies the requirements of UI on the user's ability to understand the information. It concerns providing the user understands the content provided in the application regardless of disabilities.

1) When using the input form, a method to prevent or correct input errors is required to be provided.

NOTE 1 - The input format is required to provide alternative information to identify its intended use and purpose. Also, explanatory information on the input content should be provided to prevent input errors in the input format with a different input method. When there is an error in the user input value, a method should be provided to inform the user of the error content to be understood and corrected. Moreover, it is desirable to move the focus to the point where the error occurred so that it is easy to correct the error contained in the input form.

2) Controls with similar functions within the application are required to be provided consistently.

NOTE 2 - For those with cognitive disabilities, different controls with similar functions within the application may confuse users. The same type of controls should be used throughout the application to avoid confusion.

3) It is recommended that no blinking or flashing content be provided.

*Editor's note: Need to add a reference regarding "3 to 50 cycles per second" below.*

NOTE 3 - Content that has a blinking effect on the screen should be avoided, especially 3 to 50 cycles per second. Inevitably used, content that provides a blink is required to be informed in advance and equipped with a way to avoid it.

4) It is recommended that no automatic background sound be used.

NOTE 4 - Using the background sound that is automatically played should be avoided. In the case of using a background sound, a means for the user to control it, such as stopping, pausing, and controlling volume should be provided.

5) When a screen is switched or an event such as a pop-up is executed, a method of predicting it is required to be provided.

NOTE 5 - Especially when a user is using a screen reader, the user needs to know the transition of the screen when a screen is switched automatically or at an event such as a pop-up is executed.

- 6) If an application is switched to an external application, it is required to provide a way to predict it.

NOTE 6 - Users using screen readers may not be aware of the transition of the application in a visual way. In this case, confusion can be reduced only when the screen is clearly notified before switching.

- 7) The content placed on the screen is recommended to be consistently arranged in the same order.

NOTE 7 - When the arrangement of content is dynamically changed, the user should be able to see that the structure of the content is dynamically changed. In addition, content that dynamically changes its placement should allow the user to know the placement criteria.

- 8) All text used on the screen are required to use clear expressions, symbols, and abbreviations.

NOTE 8 - Clear expressions, symbols, and abbreviations should be used for users with developmental disabilities or who speak a non-native language. In particular, explanations should be included when using abbreviations so that the original meaning can be grasped.

- 9) Visual contents such as images and photos are required to be easily understood without an additional sentence description.

NOTE 9 - In particular, even when text descriptions are included in visual contents such as images and photos, the meaning should be grasped only by the image and the picture itself.

- 10) If two or more functions need to be used continuously to perform a single action, a clear explanation is required to be given to users.

NOTE 10 - When providing the ability to perform an action using sequential functions, an explanation is required to be provided so that the user knows the meaning of each function and the consequences of each detailed action step.

- 11) If feedback is provided to the user, the functionality is required to be provided to clarify that the user is aware.

NOTE 11 - Identifying the user's perception is not just a wrong message, but it is necessary to determine whether it is recognised through the user's secondary feedback.

- 12) When using voice recognition technology, a means for the user to verify the recognised content should be provided.

NOTE 12 - Since speech recognition technology faces inherent error possibilities, it is essential to present the recognised content to the user to determine whether it is recognised as intended by the user.

- 13) When providing a complementary Alternative Augmentative Communication (AAC) function, a training information is required to be provided so that the user can learn the AAC function.

NOTE 13 - AAC can be used as an alternative for users with speech disabilities. When using AAC as a UI, the application must allow the user to master the function thoroughly.

### **7.3 UI requirements to perform operation**

*Editor's Note: This clause identifies requirements of UI on user's Ability to perform the required operations*

*e.g.*

- *Keyboard input*
- *Time to complete action*
- *Navigating through information (focus, identification, context change, etc.)*

[TBA]

### **7.4 UI requirements to use with assistive technology**

*Editor's Note: This clause identifies requirements of UI on user's ability to use suitable assistive technology*

*e.g.*

- *Compatibility of assistive technologies*
- *Operating user interface using assistive technologies*

[TBA]

## **Appendix I**

### **Use cases indicating the needs for accessible user interface of smart applications supporting Internet of things (IoT) that are required for persons with disabilities and those with specific needs**

(This appendix does not form an integral part of this Recommendation.)

In this appendix, use cases illustrating the needs for accessibility requirements for user interface of smart applications supporting Internet of things (IoT) in order for persons with disabilities, persons with age-related disabilities and those with specific needs to utilize the benefits of IoT services will be presented here as examples.

#### **I.1. Providing alternative texts**

When smart devices are used on IoT services, persons with visual disabilities have difficulty reading and understanding the information displayed on the screen. Also, it is difficult for them to locate buttons or enter information using screen digitizer of the smart device. To address this problem, smart applications need to provide an alternative means to replace the information displayed on the screen so that it can be understood without visually perceiving the screen. As an alternative means of converting visual information into auditory information, a screen reader software can be used in smart devices. This screen reader software uses the text-to-speech (TTS) function to read the screen to the user. Many mobile operating systems provide a built-in screen reader software. Users also can install and use a third-party screen reader software of their choice.

In general, screen reader software can convert textual information into auditory information using the TTS function. In addition, for non-text information such as images and videos, the TTS functionality can be used by adding appropriate alternative information as annotations to images or videos.

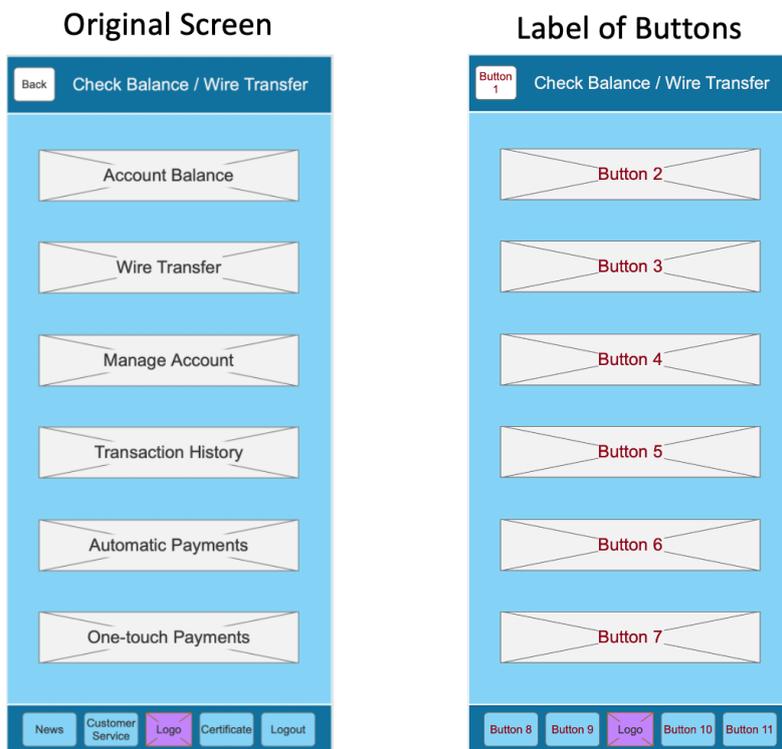
In order to express appropriate alternative auditory information for visual contents, appropriate principles must be consistently applied. Also, semantic equivalence between the visual information provided through the screen and auditory alternative through the screen reader software should be maintained.

In case of one-to-one conversion that simply relying on the function of TTS, the alternative auditory output may result in missing, distorted or implicitly expressed information or instructions. Consider a screen that receives user ID and password input for membership authentication. The visual screen can consist of simple two input components that ask a user ID and password. However, converting it to audio using the screen reader software may not clearly explain what each input components are for. Hence, extra instructions may need to be provided for clarification.

In some cases, an icon of magnifying glass may be used for searches. Utilizing icons is very effective at compactly presenting information visually, but when converting to auditory information, it is often difficult to transcribe information clearly without appropriate explanation.

Therefore, during the development process of the smart applications, the contents displayed on the screen needs to ensure the semantic equivalence between the two aspects -- one, the visually

provided information, and, two, the auditory alternative information provided by the screen reader software.



**Figure I-1 Example of incorrect button labelling**

The example in Figure I-1 shows a person without disability who can choose what he wants by looking at the buttons listed on the screen. A person with visual disabilities cannot perceive the visual contents of the screen and must read the screen through the screen reader software. Since the buttons are images, the TTS function reads the label of the buttons. If an inexperienced developer leaves the label of said buttons unmodified to the initial value of a “button”, the screen reader software will read meaningless information. Examples what the user will hear would be only “Button 1, Button 2, Button 3, etc.” Then, the user would no longer be able to proceed with the application.

## **I.2. Providing proper focus structure**

When the auditory information is used as an alternative for visual information, several problems may arise due to the differences between the visual and auditory characteristics.

Visual information is represented in two dimensions. This allows multiple information (i.e. detailed components), to be expressed at one time. Usually, visual information is expressed utilizing a variety of visual elements such as text, icons, and pictures. The information is provided to reflect the service provider's intentions through the spatial arrangement of these detailed components.

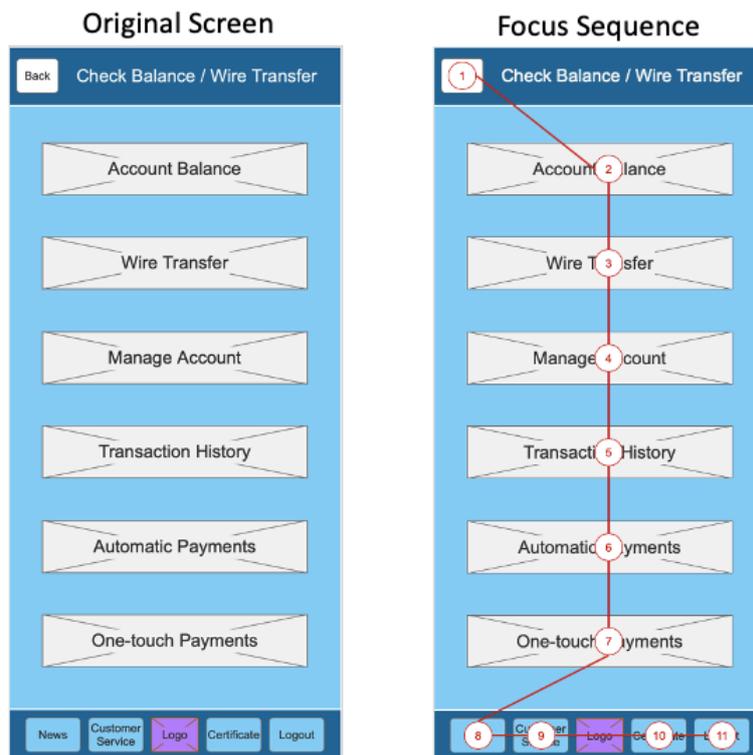
On the other hand, auditory information displays information in chronological order. Auditory information is mostly commentary expression utilizing various linguistic expressions and various sound signals. Due to the nature of the auditory information in which information is provided over time, the order in which information is presented is important to reflect the intention of the service provider. Auditory alternative information is dependent on the visual information on the screen. Alternative auditory information is provided for the visual

information. The goal should be not only to reflect sensory differences between visual and auditory information, but also to allow the user to understand the screen and communicate with the device regardless if the screen is not visually perceivable.

There are two ways to utilize the screen reader. The first method is a method in which the screen reader reads the content of a specific part when the user touches a specific part of the screen. The second way is for the screen reader to read the screen in order from beginning to end. The issue of the focus structure arises when using the screen reader in the latter way.

To express a two-dimensional screen through a screen reader chronologically, the sequential order in which the screen is read is important. In order for the user to clearly understand the screen, this order should follow the logical order of the information rather than the specific positional order of the screen (i.e., top-to-bottom or left-to-right). In addition, to ensure the semantic equivalence of visual and auditory information, meaningless visual embellishments (e.g., background images or blank images to secure space) should not be read, but, at the same time, visual components with meaning or function should not be omitted.

In which order the screen components are read through the screen reader software is determined by the order in which the components receive focus. Therefore, when composing a screen, the order of focus should be set in consideration of the logical order in which the screen reader software reads the screen.



**Figure I-2 Logical Focus Sequences**

Figure I-2 shows an example of a focus sequence when the screen reader reads the whole page. The flow is from the top left to the bottom right consistent with the screen layout. Note that the unnecessary embellishment (i.e. the logo image) at the bottom pane is intentionally ignored.

### **I.3. Providing caption and sign interpretation**

*Editor's note: text below (till end of I.3) has not been fully reviewed at the October Q2/20 meeting.*

It is difficult for persons with hearing disabilities to recognize auditory information of multimedia contents, and for persons who do not speak the language of the meeting as their native language. Consequently, alternative information such as the use of captions or sign interpretation (either in real time or recorded) of the content should be provided to assist in the recognition of the content.

There is a difference between captions that are automated and real time human generated captions. The latter is preferred as it is more explicit in giving an exact replication of all spoken contents accurately in real time. ASR, automated speech recognition captioning, is less accurate because it cannot differentiate well unusual accents i.e., foreign accents or the voices of persons who are deaf and hard of hearing. Human captioners are more accurate and faster.

Real human sign language interpreters are preferred over the use of avatars. A wider choice of languages can be provided by human sign language interpreters spontaneously and since there is at least one sign language for every country, more accurately. Avatars are only able to communicate one way and have to be pre-recorded. Though there is a work being done, avatars cannot operate in real-time two-way situations, and can be expensive to create. Human sign language interpreters work just like spoken language voice interpreters. They are able to interpret for individuals and larger audiences such as meetings or one-to-one arrangements. Avatars cannot interpret for a live human being i.e. such as a participant in a large gathering asking a question. Real human sign language interpretation is a two-way communication.

There is a signed communication called international sign (IS), which is not a sign language as it is created on the spot by highly trained sign language interpreters and which changes by virtue of the different sign language requirements of the participants for every single meeting.

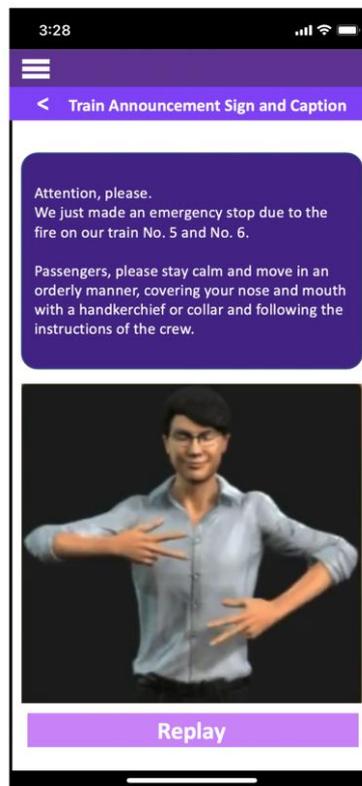
Such alternative information can be helpful at noisy environments and when persons with disabilities or persons whose language is not the language of the meeting in attendance. It is also helpful to have quiet environments where extraneous and unnecessary sound is not present.

Pre-produced content as in pre-recorded videos and live broadcasts that have pre-planned content can be provided with captions or sign interpretation in advance. However, in the IoT and Smart City environments, some video information may be inserted and used spontaneously. This content can be obtained from various sources. It may include various content produced and provided by the users. Independently provided contents by users may make it difficult to insert captions or sign interpretation retrospectively.

Abstract expressions such as background sounds or audible alarms can be difficult to express with captions unless they are human real-time produced captions and not ASR captions. It is possible to express background sound and audible alarm in real-time human sign language.

In addition, it is insufficient to provide alternative information only in one interpretation, i.e. captions or sign interpretation. Both captions or sign interpretation must be provided as some users may not read captions or understand sign interpretation.

For smart devices, captions or sign interpretation may be difficult to understand visually when they are displayed in a small corner of the screen. The limitations of the physical size of the display may make it difficult for some to understand captions or sign interpretation. In addition, especially for ASR and automatically generated content, care needs to be taken to ensure that captions and sign interpretation does not obscure important information on the screen.



**Figure I-3 Captions and sign language services through mobile applications**

Figure I-3 is an example of a service that an announcement is made when an accident occurs while using a train service. That in-cabin announcement could be provided as a video with captions and sign interpretation along with audible interpretation using mobile applications. Persons with disabilities can either see, read or hear the announcement with captions, sign language, and audible interpretation using the mobile applications to understand the situation.

In the IoT environment, diverse interfaces may coexist and provide alternative means to each other. Various services may be provided using these functions. In the case of railway services, multilingual audio and text services may also be provided for foreigners. In the same context, smart services using IoT may include the ability to combine international sign (IS) or provide multiple versions of sign interpretation in different sign languages to provide services for foreigners with hearing disabilities. Audible announcements are important in multiple languages for persons who has visual disabilities.

[TBA]

## Bibliography

- [b-ITU-T F.790] Recommendation ITU-T F.790 (2007), *Telecommunications accessibility guidelines for older persons and persons with disabilities*.
- [b-ITU-T F.791] Recommendation ITU-T F.791 (2018), *Accessibility terms and definitions*.
- [b-ITU-T F.921] Recommendation ITU-T F.921 (2018), *Audio-based indoor and outdoor network navigation system for persons with vision impairment*.
- [b-ITU-T F.922] Recommendation ITU-T F.922 (2020), *Requirements of information service system for visually impaired persons*.
- [b-ITU-T F.930] Recommendation ITU-T F.930 (2018), *Multimedia relay services*.
- [b-ITU-T H.702] Recommendation ITU-T H.702 (2020), *Accessibility profiles for IPTV systems*.
- [b-ITU-T H-Sup.17] ITU-T H Series Recommendation – Supplement 17 (2014) | ISO/IEC Guide 71:2014, *Guide for addressing accessibility in standards*.
- [b-ITU-T Y.4000] Recommendation ITU-T Y.4000/Y.2060 (2012), *Overview of the Internet of things*.
- [b-ITU-T Y.4204] Recommendation ITU-T Y.4204 (2019), *Accessibility requirements for Internet of things applications and services*.
- [b-ITU-T Y.4211] Recommendation ITU-T Y.4211 (2020), *Accessibility requirements for smart public transport services*.
- [b-ISO/IEC TR 29138-1] ISO/IEC TR 29138-1:2009, *Information technology – Accessibility considerations for people with disabilities – Part 1: User needs summary*.
- [b-ISO/TR 22411] ISO/TR 22411:2021, *Ergonomics data for use in the application of ISO/IEC Guide 71:2014*.
- [b-UNCRPD] United Nations (2006), *Convention on the rights of persons with disabilities*. New York, NY: United Nations. 37 pp. Available [viewed 2021-02-10] from: [https://treaties.un.org/doc/Publication/CTC/Ch\\_IV\\_15.pdf](https://treaties.un.org/doc/Publication/CTC/Ch_IV_15.pdf)
- [b-W3C WCAG 2.1] W3C WCAG 2.1 (2018) *Information technology – W3C Web Content Accessibility Guidelines (WCAG) 2.1*.  
<http://www.w3.org/TR/WCAG21/>
- [b-W3C UAAG 2.0] W3C UAAG 2.0 (2015), *W3C User Agent Accessibility Guidelines (UAAG)*.  
<http://www.w3.org/TR/UAAG20/>
- [b-Yan] Yan, S., and Ramachandran, P. G. (2019), *The Current Status of Accessibility in Mobile Apps*, *ACM Transaction on Accessible Computing*, Vol. 12, No. 1, pp. 1-31.

[TBA]