
**Information technology — User
interface component accessibility —**

Part 15:

**Guidance on scanning visual
information for presentation as text in
various modalities**

iTeh STANDARD PREVIEW

(standards.iteh.ai) *Technologies de l'information — Accessibilité du composant interface
utilisateur —*

*Partie 15: Recommandations relatives à la numérisation des
informations visuelles en vue d'une présentation sous forme de texte
selon différentes modalités*

<https://standards.iteh.ai/data/standards/sis/a7a50237-90cb-470c-8762-83ab36c0b843/iso-iec-ts-20071-15-2017>



iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC TS 20071-15:2017

<https://standards.iteh.ai/catalog/standards/sist/a7a30237-90cb-47be-87b2-8f3ab56c6684/iso-iec-ts-20071-15-2017>



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2017, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

	Page
Foreword	v
Introduction	vi
1 Scope	1
2 Normative references	2
3 Terms and definitions	2
4 Conformance	3
5 Overview of scanning visual information for presentation as text in various modalities	3
5.1 General	3
5.2 Contexts for scanning	4
5.3 Framework for scanning	4
5.4 Types of devices	5
5.5 Type of software	6
5.6 Scanning modes	7
6 Guidance on scanning visual information for presentation as text in various modalities	8
6.1 General guidance	8
6.1.1 Provide scanning results according to the purpose	8
6.1.2 Provide scanning results equivalent to the object being scanned	8
6.1.3 Provide focus adjustment	8
6.1.4 Provide exposure adjustment	8
6.1.5 Provide position adjustment	9
6.1.6 Providing distance adjustment	9
6.1.7 Provide orientation adjustment	9
6.1.8 Provide accessible initiation of scanning	9
6.1.9 Provide notification during scanning progress	10
6.1.10 Provide the scanning results in a textual representation	10
6.2 User preference settings	11
6.2.1 Enabling user preference settings	11
6.2.2 Selecting scanning modes	11
6.2.3 Selecting purpose of use	11
6.2.4 Selecting image resolution	11
6.2.5 Selecting image enhancement	11
6.2.6 Controlling presentation of visual guidance	11
6.2.7 Selecting post-processing components	12
6.2.8 Selecting modalities for textual representation	12
6.2.9 Selecting other application(s) to be invoked	12
6.2.10 Storing and retrieving user preferences	12
6.3 Scanning input	12
6.3.1 General	12
6.3.2 Scanning/capturing appropriate images	13
6.3.3 Enhancing image quality	13
6.3.4 Recording scanned images	13
6.3.5 Naming scanned images	14
6.3.6 Retrieving scanned images for processing	14
6.4 Processing	14
6.4.1 General	14
6.4.2 Presenting results of processing	14
6.4.3 Pre-processing: selecting visual objects of interest	14
6.4.4 Post-processing: respecting the context	15
6.5 Output	15
6.5.1 General	15
6.5.2 Presenting accessible scanning results	15
6.5.3 Invoking other applications: passing the scanning results	15

6.5.4 Privacy protection and security.....	16
Annex A (informative) Scanning visual information for presentation as text in various modalities — Use cases.....	17
Annex B (informative) Checklist of recommendations	21
Bibliography	24

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC TS 20071-15:2017](https://standards.iteh.ai/catalog/standards/sist/a7a30237-90cb-47be-87b2-8f3ab56c6684/iso-iec-ts-20071-15-2017)
<https://standards.iteh.ai/catalog/standards/sist/a7a30237-90cb-47be-87b2-8f3ab56c6684/iso-iec-ts-20071-15-2017>

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by ISO/IEC JTC 1, *Information technology, SC 35, User interfaces*.

A list of all parts in the ISO/IEC 20071 series can be found on the ISO website.

Introduction

There are a wide variety of visual objects in a user's environment that provide information that could be scanned and processed to output text-based information about or related to the object. Providing text-based information can provide accessibility to diverse users in various contexts of use, including:

- persons who cannot see the information (due to vision or environmental limitations);
- persons who cannot approach the information closely enough to see it (due to physical or environmental limitations);
- people who cannot understand the information (due to cognitive or linguistic limitations);
- where the information is provided in a format that human cannot directly understand (e.g. barcodes, QR codes);
- where information content, beyond what is in the scanned image, can be obtained through additional processing.

This document provides guidance on various aspects of the user interface of applications that scan visual information. This includes selecting the types of information that are of interest to the user, helping the user locate visual objects of interest, helping the user position the object or device used for scanning, scanning the information, processing the scanned image, and outputting textual information in various modalities.

The guidance contained in this document can be applied to a wide range of devices, applications, and contexts of use.

STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC TS 20071-15:2017
<https://standards.iteh.ai/catalog/standards/sist/a7a30237-90cb-47be-87b2-8f3ab56c6684/iso-iec-ts-20071-15-2017>

Information technology — User interface component accessibility —

Part 15: Guidance on scanning visual information for presentation as text in various modalities

1 Scope

This document provides guidance on various aspects of the user interface of applications that scan visual information that are used directly by humans, including:

- initiating the scanning application;
- setting user's preferences and configuring the scanning application;
- identifying the types of information currently of interest to the user;
- locating visual objects of interest to the user;
- creating a static image via scanning the visual object;
- identifying the information content provided by the visual object;
- processing scanned information and outputting the results to the user.

This document provides increased accessibility by addressing the user accessibility needs of diverse users in diverse contexts.

This document contains guidance that can be applied to a variety of devices, including:

- specialized devices that are dedicated to scanning and processing visual information;
- mobile devices (such as smartphones and tablets);
- general purpose computers with camera capabilities;
- office machines with scanning functions.

This document contains guidance that can be applied to various types of software, including:

- stand-alone scanning applications;
- applications including scanning functionalities;
- (scanning) applications that interoperate with other applications.

This document contains guidance that can be used for outputting scanned information in various modalities, including:

- audio outputs;
- visual outputs;
- tactile outputs;
- storing information for future use within the application performing the scanning;

- electronic outputs (to other applications, systems, or devices including those directly connected and those connected via the Internet).

This document does not:

- apply to fully automated scanning that is not under direct human control;
- apply to applications that scan visual information for editing an image or just displaying it;
- provide guidance on the design of specific hardware devices involved in scanning;
- provide guidance on the specific objects that can be recognized or the specific software used to recognize these objects;
- provide guidance on the internal functioning of software that recognizes specific types of objects.

This document is intended for use by developers of applications that include user controlled scanning functionalities. It does not expect that an application includes all of these functionalities. It can be used for those functionalities that an application does provide.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at <http://www.electropedia.org/>
- ISO Online browsing platform: available at <https://www.iso.org/obp>

3.1 scanning

use of a camera or similar function to obtain image data

3.2 recognition

<scanning> processing of *scanned image* (3.6) to locate and identify visual *objects of interest* (3.4)

3.3 visual information

recognizable information presented by a visual object

3.4 object of interest

type of object that has been chosen by a user

3.5 information of interest

type of information that has been chosen to be processed by a *scanning* (3.1) system

3.6 scanned image

data resulting from *scanning* (3.1)

3.7

scanning application

software that scans visual objects, recognizes and processes the *objects of interest* (3.4), and outputs the information about the objects

3.8

cropping

<scanning> cutting off part(s) of a *scanned image* (3.6) so that its dimensions, and possibly its information content, are changed

[SOURCE: ISO/IEC 5127:2017, 3.5.9.1.08, modified — “a document or graphic” has been replaced by “a scanned image”.]

4 Conformance

It is recognized that all of the functionalities recommended in this document might not be implemented.

If an organization claims conformity of a system with this document, then it shall document:

- a) the decisions of which recommendations from this document are and are not applicable to the system;
- b) how it addresses each of the applicable recommendations in this document or the justifications for any course of action that deviates from any of the applicable recommendations.

In making such a claim, an organization should be specific about the basis on which it is made and should have evidence to support the claim. The claim can be based on self-assessment or by an assessment carried out by another party.

(standards.iteh.ai)

[Annex B](#) provides assistance in documenting conformance to this document.

ISO/IEC TS 20071-15:2017

5 Overview of scanning visual information for presentation as text in various modalities

5.1 General

Humans recognize a multitude of objects in the real world and in scanned image, despite the fact that their vision of the objects can vary somewhat from different viewpoints, in different sizes and scales or even when they are translated or rotated.

As a type of computer vision technology, object recognition is the task of finding and identifying the visual objects of interest. Utilizing this technology allows for a wide range of applications that can provide information associated with the object. It can identify, locate, and retrieve information of the physical features of the object, and search for detailed information relevant to the scanned image.

Utilizing this technology can help users with diverse capabilities in diverse contexts to identify and understand objects as well as environments. Users, who will be aided by this technology include:

- users who cannot see the visual object being scanned;
- users who cannot directly understand the visual object being scanned;
- users who cannot sufficiently approach the visual object being scanned.

Contextual examples of use that can be provided by this technology include:

- where the information is not provided in a format that a human can directly understand (e.g. barcodes, QR codes);
- where the information content beyond what is in the scanned image can be obtained through additional processing.

In order to be of use to the widest range of users, the results will be provided by textual or other informative forms of presentation.

NOTE [Annex A](#) contains example use cases for scanning visual information for presentation as text.

5.2 Contexts for scanning

A scanning application captures (or scans) visual objects, identifies the visual objects of interest in the scanned images, finds out the visual information of interest, and provides the scanning results to the user.

In everyday life, people can encounter a variety of applications that use scanning functionalities.

EXAMPLE 1 When the customer places the printed barcode of a product under the barcode reader installed on the wall of the store, the customer can get the product information recognized by the reader through the display attached to the reader (see [A.1](#)).

EXAMPLE 2 By connecting the document scanner and the refreshable braille display to a PC and activating the optical character recognition (OCR) application, the user can read the content of the document in braille (see [A.3](#)).

EXAMPLE 3 By installing and activating a colour-aware application on the mobile device, a person who cannot distinguish colour can recognize the colour of the selected garment without the help of others (see [A.4](#)).

EXAMPLE 4 By installing and activating an application on the mobile device that identifies the number of visual objects specified by the user, it is easy to determine how many objects exist in the direction the mobile device camera points (see [A.5](#)).

EXAMPLE 5 Someone who cannot see waits for a bus putting on a wearable scanning system that recognizes shapes and text. This person can identify the bus he/she wants to take through the information the system reads by voice (see [A.6](#)).

EXAMPLE 6 When a traveller captures a landmark image using a mobile device, on which the scanning application analyses the image and automatically opens the web browser to search information about the landmark, the traveller enjoys the pleasures of travelling more than he/she would without the mobile device and the scanning application (see [A.7](#)).

EXAMPLE 7 When a person using the address book management software purchases a business card scanner and scans business cards, he or she then conveniently stores the name and phone number without typing them (see [A.8](#)).

5.3 Framework for scanning

The procedure of scanning, which inputs visual objects by scanning, identifies visual objects of interest and generates visual information of interest, shown in [Figure 1](#).

Scanning can be done on a wide range of visual objects including (but not limited to) visual text, graphic codes and/or other identifiable visual objects. Visual text includes text on a printed material, a signboard, a road sign or a business card, etc. Graphic code includes (but not limited to) barcodes, icons, emoticons, traffic signs, etc. Other identifiable visual objects include daily goods, transportation, landmark, etc.

The scanning process consists of three steps.

First is the scan input step, which scans the visual object to obtain the scanned image. In this step, the user positions the visual object of interest within the camera angle, adjusts the size of the object. It might be necessary to adjust the exposure or focus the object to get a good quality image. The scanned image will be used in subsequent steps or can be stored for future use.

Second is the processing step, which extracts the visual information of interest of the scanned image. This step is the core of the scanning application.

Pre-processing is the initiating phase of the processing step, which identifies the visual objects of interest in the scanned image according to the purpose of the scanning application. Subsequently, main-processing extracts the visual information from the scanned (visual) objects.

Post-processing is optionally performed by the user's behest to carry out additional tasks (e.g. translation, database search).

Typically, the processing step proceeds without the user's intervention, once initiated.

Third is the output step, which provides the user with textual information about the visual information of interest extracted in the processing step. Textual information is presented in suitable modalities for user characteristics.

NOTE Some modalities include visual, audio, and tactile presentation.

In some cases, the textual information of the output step is used as an argument when the scanning application invokes other applications without providing the scanning results directly to the user. In this case, the user is provided with the result of the callee instead of the caller.

These three steps can be repeated by request of the users or until the scanned result presents the desired outcome.

Setup defines the user preference settings, required during the scanning process. Specifying the type of visual object of interest is a part of the user preference setting. Defining the output modality is also considered as the user preference setting.

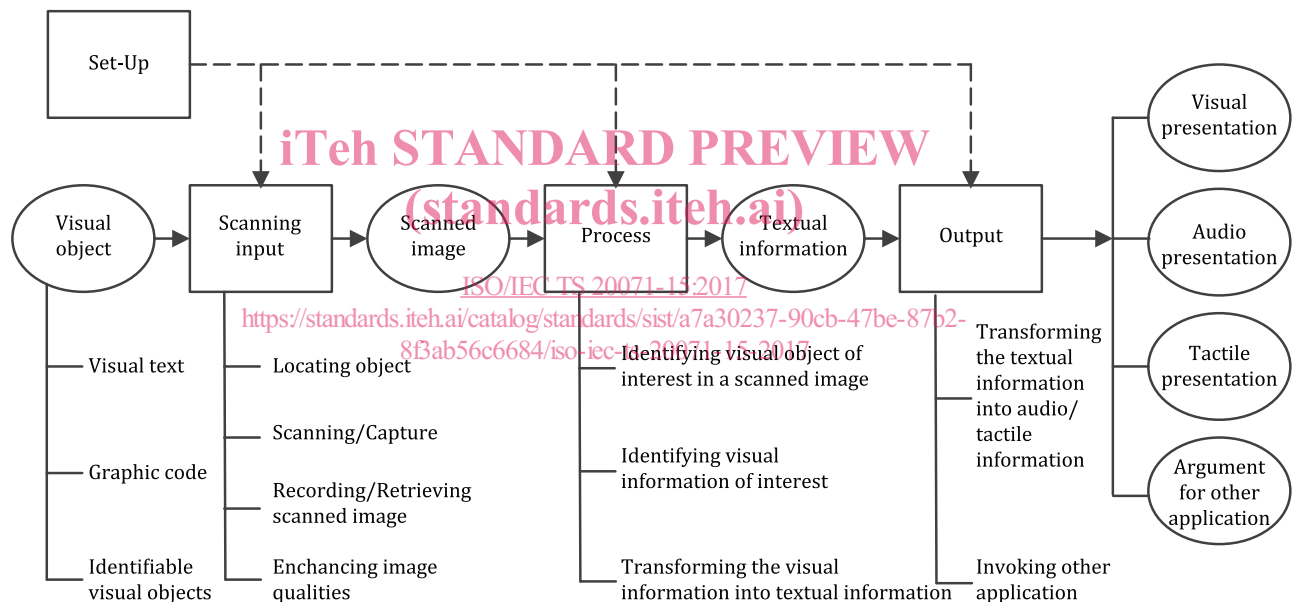


Figure 1 — Procedure of scanning visual information for presentation as text in various modalities

5.4 Types of devices

Typically, a scanning application consists of three components for (scanning) input, processing (scanned input), and outputting (results). Depending on the type of components, a scanning application can be implemented as an all-in-one device or as individual devices that are connected with or without wire connection. Scanning applications can be implemented on:

- specialized scanning devices that are dedicated to scanning and processing visual information;
- mobile devices (such as smartphones and tablets);
- general purpose computers with camera capabilities;
- office machines with scanning function.

NOTE 1 This distinction is not intended to limit the range of usage. There might be other type of devices or system that could be configured.

Specialized devices are those that are dedicated for one or more types of visual information in which the hardware and software are integrated within. Scanning applications can also be implemented on a range of mobile devices or general purpose computers containing cameras with processing capabilities as well as output capabilities. A wider variety of processing can be obtained by connecting to other applications that might be on the device or remotely connected.

The scanning input component consists of the scanning input device, such as a camera or a scanner that is used to obtain scanned images of visual objects.

NOTE 2 Scanners tend to provide their own light for viewing an object and thus are typically used at short distances from the object being scanned.

NOTE 3 Cameras can be used at longer distances from objects and can make use of existing lighting or might provide additional lighting, when needed.

NOTE 4 Regardless of whether a scanner or a camera was used, a scanned image that has been obtained will be referred to as “scanned image” within this document.

NOTE 5 The scanning input component can include a step for increasing the quality of the scanned image. This step can be performed automatically or manually in accordance to the user's instruction.

Scanning components can be fixed (in their location), hand-held or wearable.

EXAMPLE 1 Wall-mounted fingerprint scanners and cameras on laptop computers are some examples of scanning input devices that are typically fixed in a single location where it is easier to take the object to the scanner than the scanner to the object.

EXAMPLE 2 Barcode scanners, OCR pen scanners, smartphone cameras, tablet cameras are some examples of hand-held scanning input devices.

EXAMPLE 3 Cameras mounted on glasses, smart watch cameras, wearable ring cameras are some examples of wearable scanning input devices. The processing component is used to store the scanned information, to recognize objects of interest from the scanned image, and to process recognizable information within the scanned image to generate appropriate information. Data that are used for analysing a scanned image could be stored either in the scanning device itself or in storage, which is connected to the processing component via the Internet or cloud. Likewise, the processing could take place in the scanning device, via remote connection or by a combination of locations.

The processing component includes the scanning application and the hardware on which the application is running. The scanning application of the processing component identifies all the visual objects of interest from the scanned image, received from the scanning input component and then the component recognizes the visual information from these objects.

The processing component is also responsible for providing the user with the ability to configure the scanning application and set user preferences (including identifying the types of information currently of interest to the user).

Various output components (that are either part of the scanning device or that could be connected to the scanning device) can present the information generated by the scanning application.

When an assistive technology (AT) device is installed in the scanning system, the output component provides the information to the user through the AT device.

5.5 Type of software

In everyday life, people often encounter software that uses scanning functionality. Some software itself constitutes a scanning application. Some applications are designed to include scanning functionalities or to interact with an independent scanning application.