
**Information technology — Affective
computing user interface (AUI) —**

**Part 1:
Model**

*Technologies de l'information — interface pour la reconnaissance et
la simulation des émotions —*

Partie 1: Modèle



IECNORM.COM : Click to view the full PDF of ISO/IEC 30150-1:2022



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2022

All rights reserved. Unless otherwise specified, or required in the context of its implementation, 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
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Model	2
4.1 Affective characteristics	2
4.2 General model	2
4.3 AUI-based interaction model	3
4.4 Affect representation	4
4.4.1 General	4
4.4.2 Affective characteristic categories	4
4.4.3 Dimensional affective space	5
4.5 Affective characteristic data	5
4.5.1 General	5
4.5.2 Data type	5
4.5.3 Modalities and media	5
4.6 Recognition	7
4.6.1 General	7
4.6.2 Modalities of affective characteristic recognition	7
4.6.3 Multi-modal fusion	8
4.6.4 Affective characteristic recognition	8
4.7 Strategy	8
4.7.1 General	8
4.7.2 Context of use	9
4.7.3 Inference	9
4.7.4 Decision-making	9
4.8 Presentation	10
4.8.1 General	10
4.8.2 Modality of presentation	10
Annex A (informative) Example of AUI application	11
Annex B (informative) Topics for AUI standardization	14
Bibliography	16

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.

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 document 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 or www.iec.ch/members_experts/refdocs).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC 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) or the IEC list of patent declarations received (see <https://patents.iec.ch>).

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

For an explanation of 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 www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

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

A list of all parts in the ISO/IEC 30150 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Affective computing individualizes user experience based on user needs and characteristics to achieve better outcomes, including: increased effectiveness, efficiency, satisfaction, and user experience.

It is important to consider affective characteristics of humans in the design and presentation of human-computer interactions. Affective computing builds a harmonious human-computer environment by enabling computing-based systems to recognize, interpret and simulate human affects. Affective applications promise new insights into what people are feeling and can better serve their needs.

Limitations on affective computing include diverse affective characteristics currently used and the way to interpret and reply to these affective characteristics.

A general, standardized and systematic model is needed to facilitate applying affective computing within human-computer interaction regarding usability and accessibility.

This document presents a systematically defined model for affective computing user interfaces (AUI) and topics for AUI standardization. This can be important to establish the core and foundation of AUI and their applications.

This document is intended to be utilized as a reference (for guidance) for developers of systems that want to meet the needs of diverse users.

Future parts of the ISO/IEC 30150 series will provide additional guidance on implementation of AUI.

IECNORM.COM : Click to view the full PDF of ISO/IEC 30150-1:2022

Information technology — Affective computing user interface (AUI) —

Part 1: Model

1 Scope

This document establishes a model for affective computing user interfaces (AUI). [Annex B](#) also suggests topics for AUI standardization.

This document does not specify the implementation of affective computing.

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 terminology databases for use in standardization at the following addresses:

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

3.1 affective computing user interface AUI

user interface that includes the use of *affective characteristics* (3.3) to meet users' affective needs

Note 1 to entry: AUI can be implemented in software or hardware.

3.2 affective computing

collection, recognition, strategy and presentation of *affective characteristics* (3.3) of human-computer interactions

3.3 affective characteristic

particular type of affect that is believed to be useful

3.4 affect representation

general and computable model that can be used for *affective characteristic recognition* (3.6) and presentation within a computing-based system

3.5 dimensional affective space

combination of several *affective characteristics* (3.3) that can be represented as a set of dimensions

Note 1 to entry: Each affective characteristic can have a range of values (e.g. from strongly positive to strongly negative, or between two alternatives) that can form the basis of dimension in affective dimensional space.

3.6 affective characteristic recognition

process of identifying *affective characteristics* (3.3)

3.7 affective characteristic classification

process of classifying *affective characteristics* (3.3) into categories

3.8 context of use

combination of users, goals and tasks, resources and environment

Note 1 to entry: The “environment” in a context of use includes the technical, physical, social, cultural and organizational environments.

Note 2 to entry: The “users” in affective computing are represented in user profiles.

Note 3 to entry: The “resources” relevant to affective computing include collection devices, for example, cameras and microphones.

[SOURCE: ISO 9241-11:2018, 3.1.15, modified — Notes 2 and 3 to entry were added.]

4 Model

4.1 Affective characteristics

There are a range of affective characteristics (including culture and emotions) that determine a user's behaviours and needs when interacting with an information and communications technology (ICT) system. This document provides a model that works with all types of affective characteristics, without defining specific characteristics.

4.2 General model

Affective computing user interface (AUI) is a user interface that interacts with user's affective needs and characteristics. AUI processing involves affective characteristic data collection, recognition, strategy and presentation. A user adjusts his or her affects according to computing-based system's feedback and initiates further interactions with the system (shown in Figure 1). Affect representation provides a unified description of affects in AUI.

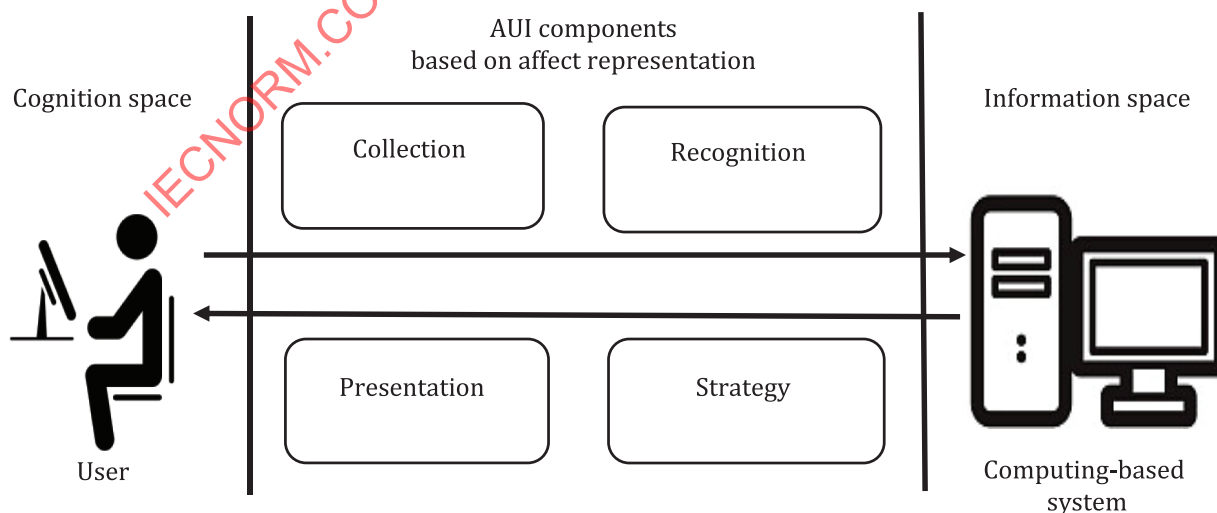


Figure 1 — AUI model

[The following serves as the text alternative for [Figure 1](#)]

The figure consists of three parts: cognition space, AUI components and information space.

The cognition space contains a user. The AUI components include four boxes demonstrating the four components of the process of the AUI model, which stand for collection of affective characteristic data, recognition, strategy and presentation. The information space contains a computing-based system.

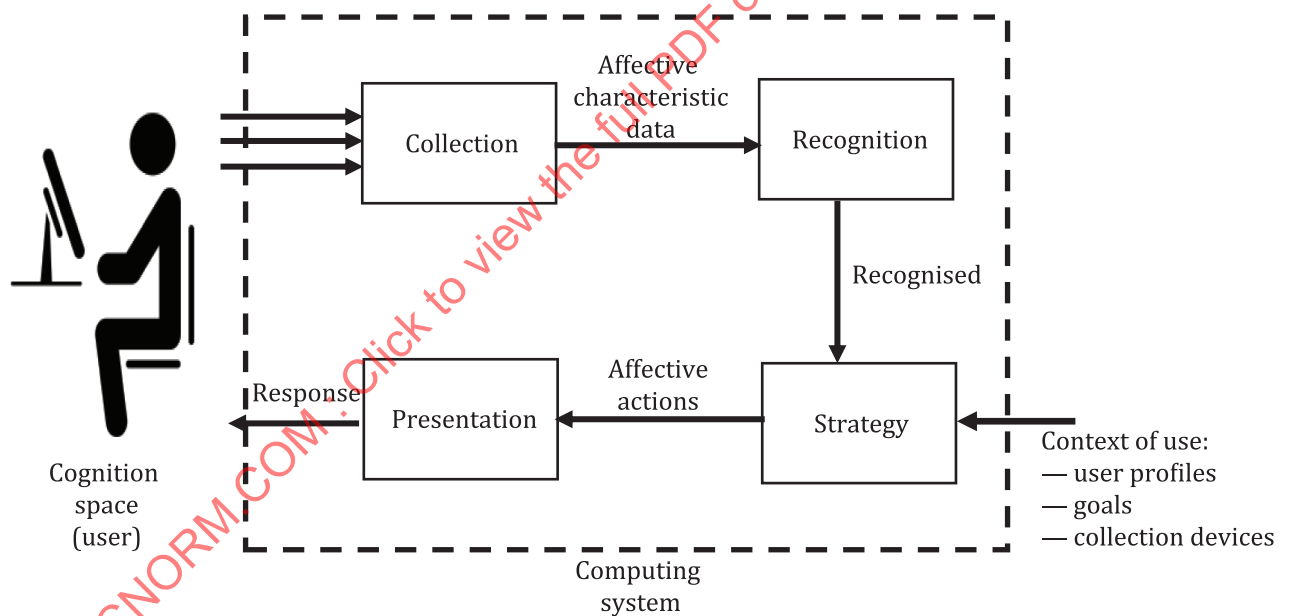
There are interactions between the cognition space and information space passing through the AUI components: one from user points to the computing-based system, indicating the input from user to the system, while the other one from computing-based system points to user, indicating the feedback to user.

Affect representation provides the foundation for the four components of the AUI model.

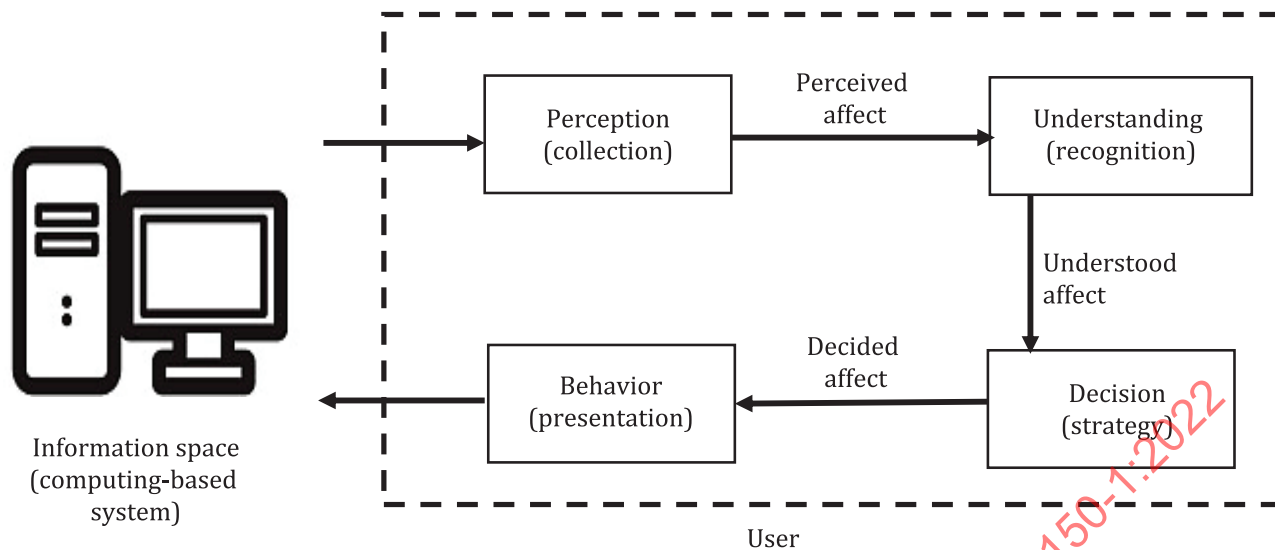
[end of the text alternative for [Figure 1](#)]

4.3 AUI-based interaction model

AUI can support interactions that implement affective computing. [Figure 2](#) provides a model of human-computer interaction using AUI involving two loops. Loop 1 is the user seen from the side of computing-based system and loop 2 is computing-based system seen from the user side. [Annex A](#) provides an example of an AUI application.



a) Loop 1: User seen from the side of computing-based system



b) Loop 2: Computing-based system seen from user side

Figure 2 — AUI-based interactions

[The following serves as the text alternative for [Figure 2](#)]

a) Loop 1: "User seen from the side of computing-based system" is illustrated in [Figure 2a](#)). It shows how human affects are seen and processed within information space. There is multi-type information from a user, such as speech, facial expression and gesture. The computing-based system first uses a collection component to gather affective characteristic data from a user. Then a recognition component identifies human affects, and a strategy component, which combines result of the recognition with externally provided context of use information and generates affective actions to be executed in the presentation component.

b) Loop 2: "Computing-based system seen from user side" is illustrated in [Figure 2b](#)).

It shows how the output of computing-based system is seen from user side. In cognition space, user processes the output of the system to perceive and understand affect. Then user decides the affect to perform behaviour as feedback to the computing-based system in order to initiate a new turn of AUI interactions. The process of perception, understanding, decision and behaviour correspond to the components of collection, recognition, strategy and presentation in loop 1, respectively.

[end of the text alternative for [Figure 2](#)]

NOTE 1 [Annex A](#) provides an application example of AUI-based interaction model.

NOTE 2 [Annex B](#) identifies possible standardization topics to implement this model.

4.4 Affect representation

4.4.1 General

Affects can be described via affect representation, which contains affective characteristic categories and dimensional affective space.

4.4.2 Affective characteristic categories

Categories of affective characteristics can be developed from any individual affective characteristic or combination of affective characteristics.

Different combinations of affective characteristics result in different user behaviours and needs.

4.4.3 Dimensional affective space

Dimensional affective space can be used to represent affective characteristics dimensionally.

NOTE 1 Different combinations of affective dimensions correspond to different dimensional affective spaces.

Recognized affect can be represented by a set of numeric (or ordinal) values, which indicate its location in the dimensional affective space.

NOTE 2 Represented recognized affect is sometimes referred to as an affective state.

4.5 Affective characteristic data

4.5.1 General

Affective characteristic data refers to data that is captured from users and is involved in affective computing.

4.5.2 Data type

Affective characteristic data can be static data or dynamic data.

EXAMPLE

- a) Static data can record user's affective characteristics at a certain moment, such as an image.
- b) Dynamic data can record user's continuous behaviours in a period, such as a video or a voice clip. Dynamic data stream reflects dynamic variance of affective characteristics.

4.5.3 Modalities and media

Modalities of affective characteristic data can be various, according to the interactions between users and AUI. Major modalities include visual, auditory, tactile and physiological signals. For each modality, diverse types of media can be used. Affective characteristic data can be of one or more than one modality. The modality-specific information and descriptions of media types are shown in [Table 1](#) and [Table 2](#), respectively.

Table 1 — Modalities and media types of input

Modality-specific information	
Modalities	{ALL} or one or more of {Visual, Auditory, Tactile, Neurophysiological}

Table 1 (continued)

Modality-specific information	
Media types	{ALL} or one or more of {Picture, Video, Gesture, TextWritten, TextSpoken, TextTactile, Facial expression, Music, Voice, Sound, Electrophysiological signal, Non-electrophysiological signal}

NOTE 1 Modalities and media types in [Table 1](#) are specified based on ISO/IEC 24756: 2009, Table 4.

Table 2 — Media types and descriptions

Media types	Description
Picture	A static image presented by the system or loaded into the system by a user (including images of user's facial expression, gesture, text).
Video	A dynamic image presented by the system or loaded into the system by a user.
Gesture	Movements of the user that can express an idea or meaning.
TextWritten	A language-based medium of words presented in a written symbolic script either statically or dynamically, typically by a system on a screen or by a user on a keyboard.
TextSpoken	A language-based medium of words spoken by the user or system.
TextTactile	A language-based medium of words presented in a tactile symbolic script either statically or dynamically, typically by a system on a tactile display or by a user on a specialized keyboard.
Facial expression	Observed facial movement that reflects one or more motions or positions of the muscles beneath the skin of the face.
Music	Sounds produced by the system or user arranged in time possessing a degree of melody, harmony or rhythm.
Voice	Sounds produced by human being using vocal folds.
Sound	Any media that can be heard by the system or the user but does not necessarily have an associated meaning.
Electrophysiological signal	Measurements of voltage changes or electric current or manipulations on biological cells, tissues or organs.
Non-electrophysiological signal	Neurophysiological signals except electrophysiological signals, usually include magnetic, mechanic, optic, acoustic, chemical and thermal bio-signals.

NOTE 2 Descriptions of some media types in [Table 2](#) are revised based on ISO/IEC 24756: 2009, Table 5.

4.6 Recognition

4.6.1 General

Recognition is the process of identifying affects through focusing on patterns and regularities in affective characteristic data. [Figure 3](#) shows a diagram of affective characteristic recognition.

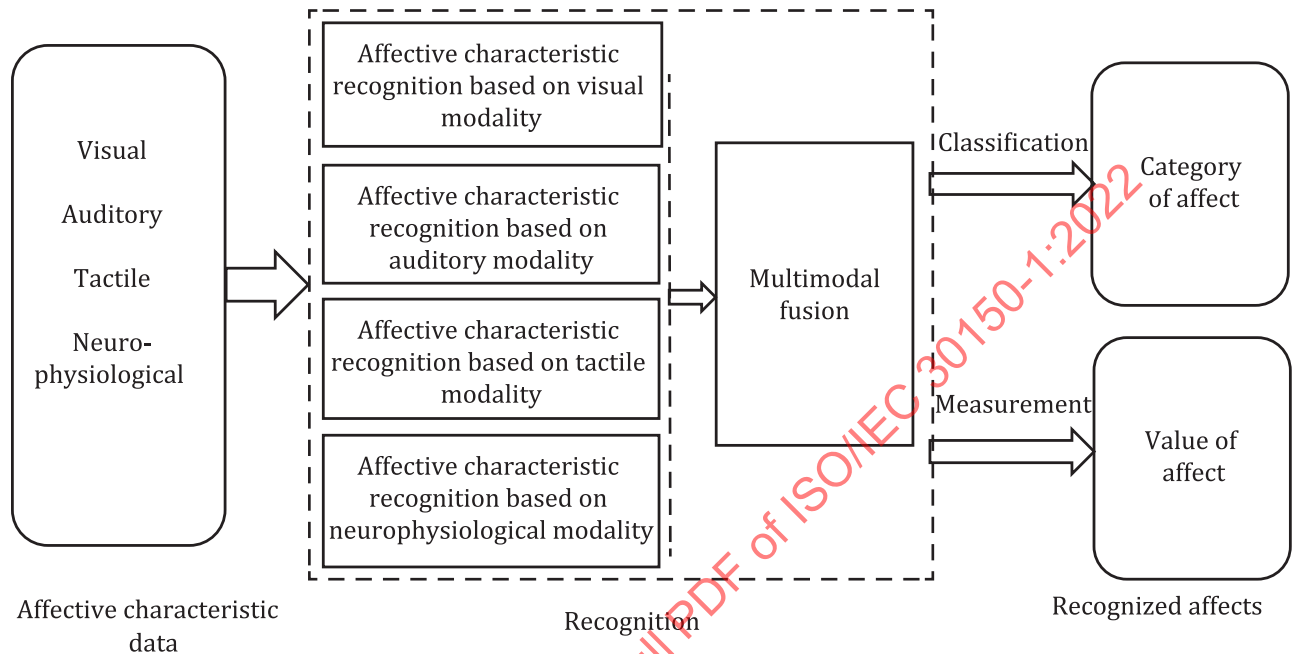


Figure 3 — Affective characteristic recognition

[The following serves as the text alternative for [Figure 3](#)]

The figure contains three parts: affective characteristic data, recognition and recognized affects.

The affective characteristic data can involve four data modalities: “Visual”, “Auditory”, “Tactile” and “Neurophysiological”.

The recognition component processes the affective characteristic data based on the data modality. Multi-modal fusion can process combinations of data modalities.

The recognized affects can be identified into affective characteristic categories through affective characteristic classification, or into dimensional affective space through measurement.

[end of the text alternative for [Figure 3](#)]

4.6.2 Modalities of affective characteristic recognition

User’s affective characteristics are recognized by affective characteristic recognition component, which processes affective characteristic data via different modalities.

NOTE For each modality, the affective characteristic recognition process can be different depending on the specific media type.

4.6.3 Multi-modal fusion

Multi-modal fusion is used to combine the result of affective characteristic recognition in one or more modalities.

EXAMPLE Multi-modal fusion can be used to combine affective characteristic recognition involving facial recognition and voice recognition.

4.6.4 Affective characteristic recognition

Affective characteristic recognition can involve:

- a) Classification of affective characteristic into categories;

NOTE 1 An input is any segment of affective characteristic data (see 4.5.3). The output of affective characteristic classification is one or more of a given set of affective characteristic categories.

EXAMPLE 1 Facial emotion recognition identifies that a person is smiling. This input can be classified as evidence that the person has joy, where “joy” is a category of affective characteristic.

- b) Measurement of a specific affective characteristic into values.

NOTE 2 An input is any segment of affective characteristic data (see 4.5.3). The output of affective characteristic measurement is a set of one or more values in dimensional affective space.

EXAMPLE 2 Voice emotion recognition identifies that a person is laughing. This input provides a measure of the joy with value (0.82, 0.43) of the person, where the rating of the value (0.82, 0.43) is based on continuous valence and arousal scales in a dimensional affective space to represent affective characteristics dimensionally.

Affective characteristic classification and measurement should be independently performed.

4.7 Strategy

4.7.1 General

Affective strategy aims to generate affective actions to be executed by computing-based systems for presentation. The strategy component accepts recognized affects and context of use as inputs. The strategy component consists of two sub-components, i.e. inference and decision-making (see Figure 4).

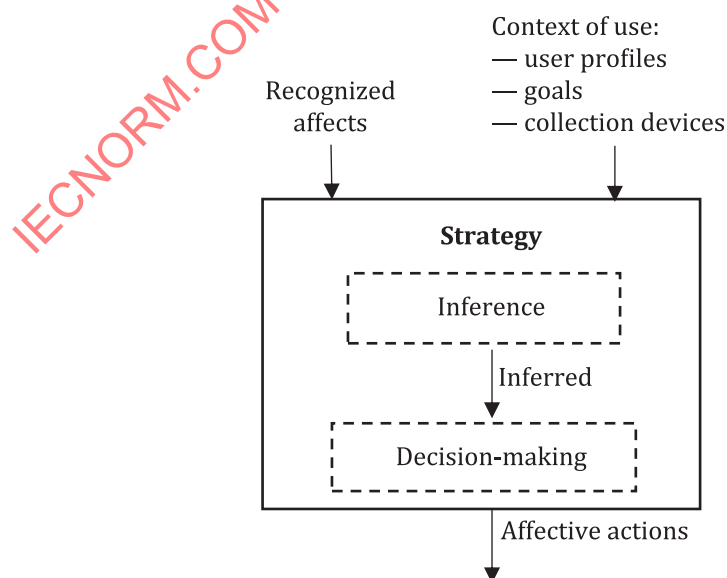


Figure 4 — Strategy module

[The following serves as the text alternative for [Figure 4](#)]

The input to the strategy component includes recognized affects and context of use. The elements of the context of use that are most used in AUI are: user profiles, goals, collection devices and environment.

The strategy component generates inferred intentions through an “inference” sub-component, and then identifies the activities undertaken through a “decision-making” sub-component based on the inferred intentions and context of use.

The output of the strategy component is the affective actions to be executed by computing-based systems for presentation.

[end of the text alternative for [Figure 4](#)]

4.7.2 Context of use

In AUI, externally provided context of use information focuses on user profiles, goals, collection devices and environment.

ISO 9241-11 provides a detailed discussion of the elements and applications of context of use. For the purpose of AUI, the main elements of context of use are:

- a) User profile: personal data associated with an individual user.
- b) Goal: intended outcome.
- c) Collection device: input device that captures affective characteristic data.
- d) Environment: technical, physical, social, cultural and organizational environments that influence usability.

4.7.3 Inference

4.7.3.1 General

The core of inference sub-component is to assign intentions through inferring the user's request according to the user's recognized affects and context of use.

4.7.3.2 Inferred intention

Inferred intention is a set of user's requests.

Inferred intention can be interaction intention or affective intention, or both.

- a) Interaction intention: a clear behavioural request of a user (such as “a user is asking a certain question” or “a user is asking for a certain service”).
- b) Affective intention: user's request for affective response or affective regulation (for example, if the user is asking a certain question anxiously, his or her affective intention about the expected affective response can be comforting).

4.7.4 Decision-making

Decision-making is a process of identifying and choosing activities undertaken in order to achieve a goal based on inferred intentions and context of use.

NOTE 1 Different combinations of activities can provide different ways to achieve the same goal.

Decision-making also aims to generate a set of affective actions to be executed by computing-based system for presentation.

Affective actions are generated to realize specific activities with expressible affects.

NOTE 2 It is important that expressible affects are generally accepted and understood.

4.8 Presentation

4.8.1 General

Presentation aims to convey and express affects to the user on output devices via affective actions.

The input of presentation is affective actions. And the output of presentation is the generated affects in various modalities and media types on output devices.

4.8.2 Modality of presentation

Major presentation modalities include visual, auditory or both. Data of each modality or media type can be presented via output devices, for example, a video or a picture can be shown on a screen and a voice can be heard from a speaker (see [Table 3](#)).

Table 3 — Presentation output

Modalities	Devices	Media types
Visual	Screen	Picture, Video, TextWritten, Gesture, Facial expression
Auditory	Speaker	Music, Sound, Voice, TextSpoken

NOTE 1 The selection of presentation modality and media type is scenario-dependent.

NOTE 2 Other modalities and media types can sometimes be used for presentation, for example, tactile modality and TextTactile media type.

Annex A (informative)

Example of AUI application

A.1 General

This annex presents a specific example of AUI application for the purpose of clarifying the AUI-based interactions shown in [Figure 2a](#)) by illustrating how all components of AUI work as a whole system and how AUI standards are defined and applied in one specific scenario – the virtual customer service in financial industry on XFINWIZ (an instant messaging platform as implemented on a smart phone). This example focuses on emotional affects.

A.2 User

The user is a customer who uses the virtual customer service on XFINWIZ to receive service with a smile from a financial company or organization. The user can interact with the system via visual, auditory, tactile and neurophysiological modalities. The media type can be, e.g. TextWritten, TextSpoken, voice, music, facial expression.

NOTE The modalities can vary with the application scenario. In this example, the modality and media types are limited by the XFINWIZ Platform supports.

A.3 Emotion representation

The emotion representation used in the virtual customer service on XFINWIZ can be emotion categories. In this application scenario, the emotion categories of particular interest include, e.g. “joy”, “sadness”, “anger”, “surprise”, “fear”.

The representation standards can regulate the emotion categories, and the mapping between emotion categories and positions in dimensional emotion space.

A.4 Dataset

Given an AUI system on XFINWIZ, a dataset is used to train and test the whole system before actual usage in order to ensure emotional alignment between a user and a computer, especially for emotion recognition and presentation.

For each modality supported by XFINWIZ platform, there is a corresponding dataset that is prescribed in terms of, e.g. the data collection, annotation, data structure.

Taking the dataset of facial images for example, the images can be collected by a camera with professional lighting and spatial resolution not lower than 250*250 pixel. There will be only one face in a single image, and the dataset will have gender balance and cover the primary age group of the users on XFINWIZ. The collected images are saved in JPEG format and labelled in certain label format based on the output of emotion recognition, e.g. emotion categories used in recognition component.

The datasets of diverse modalities will be defined and regulated by the dataset standard.

A.5 Collection of affective characteristic data

The AUI system is able to acquire the affective characteristic data of users through the modalities supported by XFINWIZ via, e.g. keypad (keyboard), microphone, camera.

The acquisition of affective characteristic data via different devices and the data structure will be regulated by the standard of collection.

A.6 Emotion recognition

The affective characteristic data is processed and analysed in recognition component to recognize the user's emotion. Suppose that the AUI system captures a voice clip from a nervous customer, after data processing and analysis, the recognition module identifies the user's emotion as "anger", which is of particular interest to financial customer service, and is represented as emotion category "ANGER".

NOTE The emotion categories of interest are scenario-dependent.

The standard of recognition will regulate the classification process in different modalities to identify the emotion categories, and the data structure for recognized emotions.

A.7 Strategy

Based on the recognized emotion of "ANGER" and the context of use, the inference sub-component can infer the user's interaction intentions (e.g. "the customer has lost his or her credit card and wants to report the loss" or "the customer forgets his or her password and wants to reset it"). It also infers the user's affective intention, e.g. "the angry customer may expect a comforting response".

The inferred intention allows choosing the response by selecting the suitable modality type, media type and output device for presentation and generates the affective action, such as "present the steps to reset the password" and "COMFORT" the customer via "text on the screen", which instructs the presentation to invoke related functions to achieve the customer's intended goals. The structure of affective actions can be normalized by the standard of strategy and decision-making.

A.8 Presentation

The presentation component accepts and executes the affective action, as shown in [Figure A.1](#). In the specific scenario of financial virtual customer service on XFINWIZ, the presentation through visual and auditory modalities, such as media types of facial expression, TextWritten, voice and music, can be defined and regulated by the presentation standard.

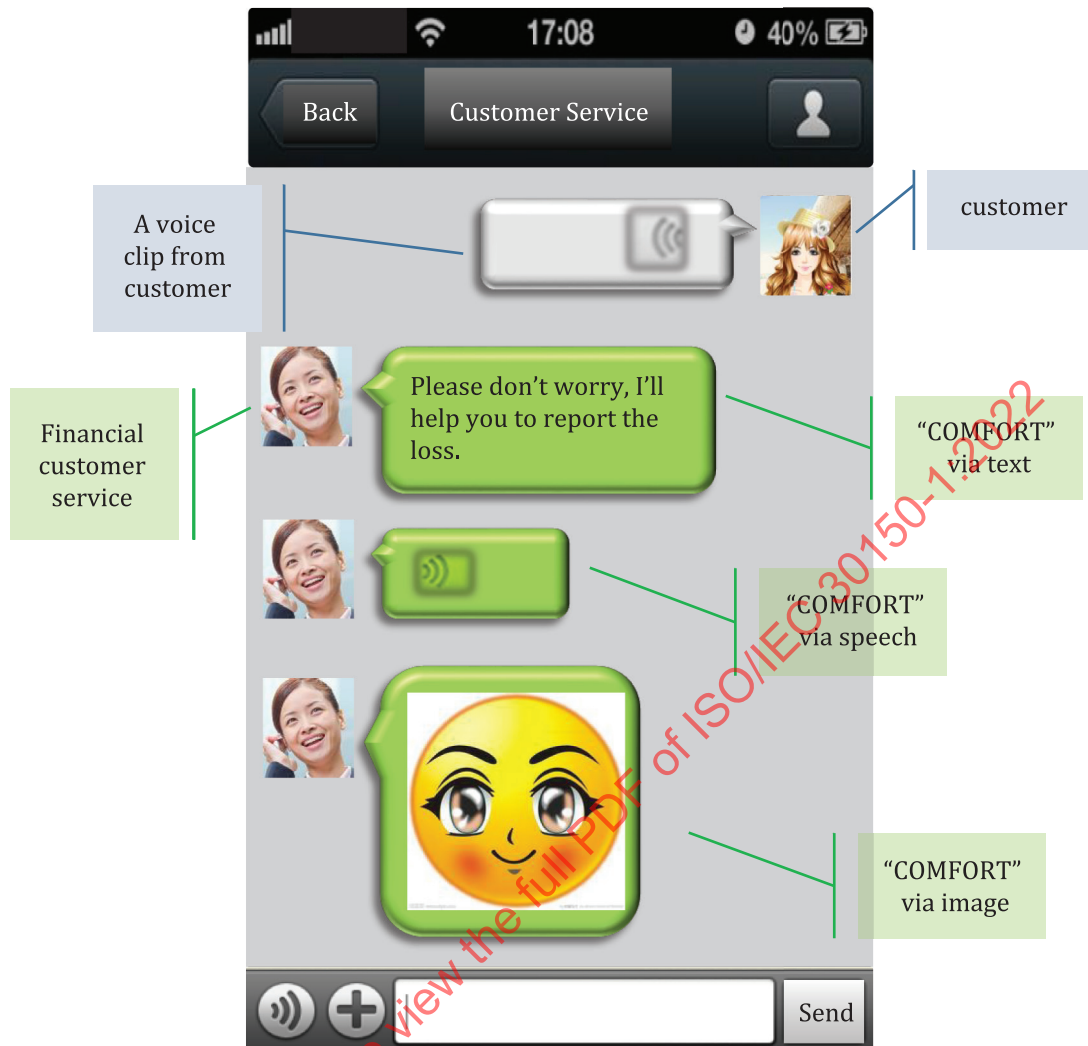


Figure A.1 — Example of AUI presentation in the scenario of financial customer service on XFINWIZ

[The following serves as the text alternative for [Figure A.1](#)]

The figure shows the interface of financial customer service on XFINWIZ. The interface mainly contains two parts: the title bar and dialogue area.

In the title bar, there is a “back” button, a “customer service” title and a user icon from the left side to the right side.

In the dialogue area, the “customer” sends a message of a voice clip. The financial customer service comforts the customer using three messages, which are a text message of “Don’t worry, I’ll help you to report the loss.”, a speech with the same text, and an image message of a smiling face.

[end of the text alternative for [Figure A.1](#)]

NOTE In the scenario of financial virtual customer service on XFINWIZ, modalities of presentation include visual and auditory. Media types can include TextWritten, voice, facial expression, symbol, or their combinations supported by the platform.