FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/IEC
FDIS
23093-1**

ISO/IEC JTC **1**/SC **29**

Secretariat: **JISC**

Voting begins on:
2021-11-16

Voting terminates on:
2022-01-11

Information technology — Internet of media things —

Part 1:
Architecture

Technologies de l'information — Internet des objets media —

Partie 1: L'architecture IoMT

Reference number
ISO/IEC FDIS 23093-1:2021(E)

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

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 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 29, *Coding of audio, picture, multimedia and hypermedia information*.

This second edition cancels and replaces the first edition (ISO/IEC 23093-1:2020), which has been technically revised.

The main changes are as follows:

— use case description and the underlying technology

A list of all parts in the ISO/IEC 23093 series can be found on the ISO and IEC websites.

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

The ISO/IEC 23093 series provides an architecture and specifies application programming interfaces (APIs) and compressed representation of data flowing between media things.

The APIs for the media things facilitate discovering other media things in the network, connecting and efficiently exchanging data between media things. The APIs also provide means for supporting transaction tokens in order to access valuable functionalities, resources, and data from media things.

Media things related information consists of characteristics and discovery data, setup information from a system designer, raw and processed sensed data, and actuation information. The ISO/IEC 23093 series specifies data formats of input and output for media sensors, media actuators, media storages, media analysers, etc. Sensed data from media sensors can be processed by media analysers to produce analysed data, and the media analysers can be cascaded in order to extract semantic information.

This document does not specify how the process of sensing and analysing is carried out but specifies the interfaces between the media things. This document describes the architecture of systems for the internet of media things.

This second edition cancels and replaces the first edition (ISO/IEC 23093-1:2019), which has been technically revised. The main changes compared to the previous edition are as follows:

— Complementary use cases

— Sequence diagrams and mission state diagrams for the use-case description in order to enhance the readability of the document

Information technology — Internet of media things —

Part 1: Architecture

1 Scope

This document describes the architecture of systems for the internet of media things.

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:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1 Internet of media things terms

3.1.1
audio
anything related to sound in terms of receiving, transmitting or reproducing it or of its specific frequency

3.1.2
camera
special form of an *image capture device* (3.1.6) that senses and captures photo-optical signals

3.1.3
display
visual representation of the output of an electronic device or the portion of an electronic device that shows this representation, as a screen, lens or reticle

3.1.4
gesture
movement or position of the hand, arm, body, head or face that is expressive of an idea, opinion, emotion, etc.

3.1.5
haptics
input or output device that senses or actuates the body's movements by means of physical contact with the user

3.1.6
image capture device
device which is capable of sensing and capturing acoustic, electrical or photo-optical signals of a physical entity that can be converted into an image

3.1.7
internet of media things
IoMT
special subset of *IoT* (3.2.9) whose main functionalities are related to media processing

3.1.8
IoMT device
IoT (3.2.9) device that contains more than one *MThing* (3.1.12)

3.1.9
IoMT system
MSystem
IoT (3.2.9) system whose main functionality is related to media processing

3.1.10
loudspeaker
electroacoustic device, connected as a component in an audio system, generating audible acoustic waves

3.1.11
media
data that can be rendered, including audio, video, text, graphics, images, haptic and tactile information

Note 1 to entry: These data can be timed or non-timed.

3.1.12
media thing
MThing
thing (3.2.20) capable of sensing, acquiring, actuating, or processing of media or metadata

3.1.13
media token
virtual token for accessing functionalities, resources and data of media things

3.1.14
microphone
entity capable of capture and transform acoustic waves into changes in electric currents or voltage, used in recording or transmitting sound

3.1.15
media wearable
MWearable
MThing (3.1.12) intended to be located near, on or in an organism

3.1.16
motion
action or process of changing place or position

3.1.17
natural user interface
NUI
system for human-computer interaction that the user operates through intuitive actions related to natural, everyday human behaviour

3.1.18
presentation
act of producing human recognizable output of rendered media

3.2 Internet of things terms

3.2.1
actuator
component which conveys digital information to effect a change of some property of a physical entity

3.2.2
capability
characteristic or property of an entity that can be used to describe its state, appearance or other aspects

EXAMPLE An entity type, address information, telephone number, a privilege, a MAC address, a domain name are possible attributes, see ISO/IEC 24760-1.

3.2.3
component
modular, deployable and replaceable part of a system that encapsulates implementations

Note 1 to entry: A component may expose or use interfaces (local or on a network) to interact with other entities, see ISO 19104. A component which exposes or uses network interfaces is called an endpoint.

3.2.4
digital entity
any computational or data element of an IT-based system

Note 1 to entry: It may exist as a service based in a data centre or cloud, or a network element or a gateway.

3.2.5
discovery
service to find unknown resources/entities/services based on a rough specification of the desired result

Note 1 to entry: It may be utilized by a human or another service; credentials for authorization are considered when executing the discovery, see ISO/IEC 30141.

3.2.6
entity
anything (physical or non-physical) having a distinct existence

3.2.7
identifier
information that unambiguously distinguishes one *entity* (3.2.6) from another one in a given identity context

3.2.8
identity
characteristics determining who or what a person or thing is

3.2.9
internet of things
IoT
infrastructure of interconnected objects, people, systems and information resources together with intelligent services to allow them to process information of the physical and the virtual world and to react

3.2.10
interface
shared boundary between two functional components, defined by various characteristics pertaining to the functions, physical interconnections, signal exchanges, and other characteristics, as appropriate

Note 1 to entry: See ISO/IEC 13066-1.

3.2.11
IoT system
system that is comprised of functions that provide the system the capabilities for identification, sensing, actuation, communication and management, and applications and services to a user

Note 1 to entry: See Bahga and Madisetti [1].

3.2.12
network
entity that connects endpoints, sources to destinations, and may itself act as a value-added element in the IoT system or services

3.2.13
process
procedure to carry out operations on data

3.2.14
physical entity
thing (3.2.20) that is discrete, identifiable and observable, and that has material existence in real world

3.2.15
reference architecture
description of common features, common vocabulary, guidelines, interrelations and interactions among the entities, and a template for an IoT architecture

3.2.16
resource
any element of a data processing system needed to perform required operations

Note 1 to entry: See ISO/IEC 2382.

3.2.17
sensor
device that observes and measures a physical property of a natural phenomenon or a human induced process and converts that measurement into a signal

Note 1 to entry: A signal can be electrical, chemical, etc., see ISO/IEC 29182-2.

3.2.18
service
distinct part of the functionality that is provided by an entity through interfaces

3.2.19
storage
capacity of a digital entity to store information subject to recall or the components of a digital entity in which such information is stored

3.2.20
thing
any entity that can communicate with other entities

3.2.21
user
human or any digital entity that is interested in interacting with a particular physical object

3.2.23
visual
any object perceptible by the sense of sight

4 Architecture

The global IoMT architecture is presented in Figure 1, which identifies a set of interfaces, protocols and associated media-related information representations related to:

— user commands (setup information) between a system manager and an MThing, with reference to interface 1.

— user commands (setup information) forwarded by an MThing to another MThing, possibly in a modified form (e.g., subset of 1), with reference to interface 1'.

— sensed data (raw or processed data) (compressed or semantic extraction) and actuation information, with reference to Interface 2.

— wrapped interface 2 (e.g., for transmission), with reference to interface 2'.

— MThing characteristics, discovery, with reference to interface 3.

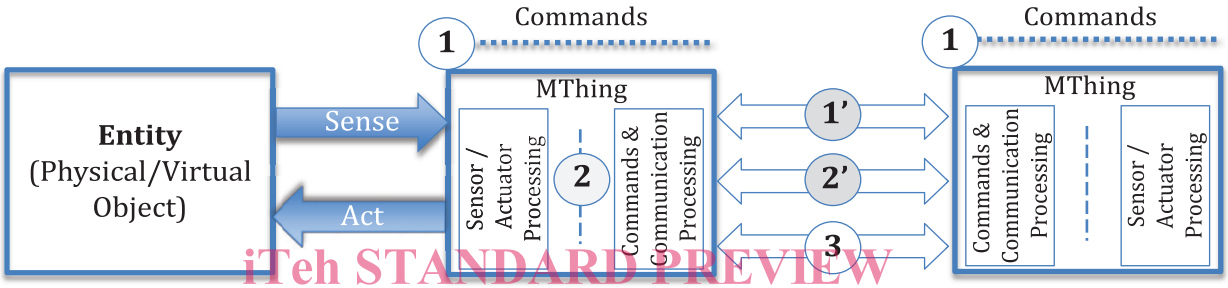

Figure 1 — IoMT architecture

This IoMT architecture can be mapped to the IoT reference architecture, see ISO/IEC 30141, as shown in Annex A.

5 Use cases

5.1 General

MPEG identified 31 use-cases for IoMT; they are structured in the following five main categories:

a) **Smart spaces: Monitoring and control with network of audio-video cameras (see 5.2)**

- human tracking with multiple network cameras
- dangerous region surveillance system
- intelligent firefighting with IP surveillance cameras
- automatic security alert generation system using, time, GPS and visual information
- networked digital signs for customized advertisement
- digital signage and second screen use
- self-adaptive quality of experience for multimedia applications
- ultra-wide viewing video composition
- face recognition to evoke sensorial actuations
- automatic video clip generation by detecting event information

- temporal synchronization of multiple videos for creating 360° or multiple view video
- intelligent similar content recommendations using information from IoMT devices
- safety equipment detection in construction sites

b) **Smart spaces: Multi-modal guided navigation (see 5.3)**

- blind person assistant system
- elderly people assistance with consecutive vibration haptic devices
- personalized navigation by visual communication
- personalized tourist navigation with natural language functionalities
- smart identifier: face recognition on smart glasses
- smart advertisement: QR code recognition on smart glasses

c) **Smart audio/video environments in smart cities (see 5.4)**

- smart factory: car maintenance assistance A/V system using smart glasses
- smart museum: augmented visit museum using smart glasses
- smart house: light control, vibrating subtitle, olfaction media content consumption
- smart car: head-light adjustment and speed monitoring to provide automatic volume control

d) **Smart multi-modal collaborative health (see 5.5)**

- increasing patient autonomy by remote control of left-ventricular assisted devices

- diabetic coma prevention by monitoring networks of in-body/near body sensors
- enhanced physical activity with smart fabrics networks
- medical assistance with smart glasses
- managing healthcare information for smart glass
- indoor air quality prediction

e) **Blockchain usage for IoMT transactions authentication and monetizing (see 5.6)**

- reward function in IoMT by using blockchains
- content authentication with blockchains

5.2 Smart spaces: Monitoring and control with network of audio-video cameras

5.2.1 General

The large variety of sensors, actuators, displays and computational elements acting in our day-by-day professional and private space in order to provide us with better and easier accessible services lead to 13 use cases of interest for IoMT, mainly related to the processing of video information.

5.2.2 Human tracking with multiple network cameras

As urban growth is today accompanied by an increase in crimes rate (e.g., theft, vandalism), many local authorities consider surveillance systems as a possible tool to fight this phenomenon. A city video