International Standard

ISO/IEC 23093-1

Information technology — Internet of media things —

Part 1: Architecture

Technologies de l'information — Internet des objets media — Partie 1: Architecture

Third edition
2025-11

Reference number
ISO/IEC 23093-1:2025(en)

© ISO/IEC 2025



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2025

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 ..... **v**

Introduction ..... **vi**

1 Scope **1**

2 Normative references **1**

3 Terms and definitions **1**
3.1 Internet of media things terms 1
3.2 Internet of things terms 2

4 Architecture **4**

5 Use cases **5**
5.1 General 5
5.2 Smart spaces: Monitoring and control with network of audio-video cameras 8
5.2.1 General 8
5.2.2 Human tracking with multiple network cameras 8
5.2.3 Dangerous region surveillance system 8
5.2.4 Intelligent firefighting with IP surveillance cameras 9
5.2.5 Automatic security alert and title generation system using, time, GPS and visual information 9
5.2.6 Pedestrian-car accident detection in video using prediction result description 10
5.2.7 Networked digital signs for customized advertisement 10
5.2.8 Digital signage and second screen use 10
5.2.9 Self-adaptive quality of experience for multimedia applications 11
5.2.10 Ultra-wide viewing video composition 11
5.2.11 Face recognition to evoke sensorial actuations 12
5.2.12 Automatic video clip generation by detecting event information 12
5.2.13 Temporal synchronization of multiple videos for creating 360° or multiple view video 12
5.2.14 Intelligent similar content recommendations using information from IoMT devices 13
5.2.15 Understand and explain events in video by instance segmentation 13
5.2.16 Indoor/outdoor acoustic event detection 13
5.2.17 Safety equipment detection on construction sites 13
5.3 Smart spaces: Multi-modal guided navigation 14
5.3.1 General 14
5.3.2 Blind person assistant system 14
5.3.3 Elderly people assistance with consecutive vibration haptic devices 14
5.3.4 Personalized navigation by visual communication 15
5.3.5 Personalized tourist navigation with natural language functionalities 15
5.3.6 Smart identifier: Face recognition on smart glasses 16
5.3.7 Smart advertisement: QR code recognition on smart glasses 17
5.4 Smart audio/video environments in smart cities 17
5.4.1 General 17
5.4.2 Smart factory: Car maintenance assistance A/V system using smart glasses 17
5.4.3 Smart museum: Augmented visit using smart glasses 18
5.4.4 Smart house: enhanced perception modes 19
5.4.5 Smart house: control of home appliance devices 20
5.4.6 Smart car: Head-light adjustment and speed monitoring to provide automatic volume control 20
5.5 Smart audio/video environments in smart rural areas 21
5.5.1 General 21
5.5.2 Crop smart farming 21
5.5.3 Smart crop growth monitoring 21
5.5.4 Livestock smart farming 22
5.6 Smart multi-modal collaborative health 23

© ISO/IEC 2025 – All rights reserved

5.6.1 General 23
5.6.2 Increasing patient autonomy by remote control of left-ventricular assisted devices 23
5.6.3 Diabetic coma prevention 23
5.6.4 Enhanced physical activity with smart fabrics networks 24
5.6.5 Medical assistance with smart glasses 24
5.6.6 Managing healthcare information for smart glasses 25
5.6.7 Emergency health event detection with infrared camera 26
5.6.8 Personalized detection of health danger by multimodal data sensing and processing 26
5.6.9 Multimodal question answer with blood pressure data 27
5.6.10 Indoor air quality prediction 28
5.7 Blockchain usage for IoMT transactions authentication and monetizing 28
5.7.1 General 28
5.7.2 Reward function in IoMT people counting by using blockchains 28
5.7.3 Content authentication with blockchains 29
5.8 Metaverse usage of IoMT technologies 29
5.8.1 General 29
5.8.2 Human pose estimation for avatar animation 29
5.8.3 Facial landmark detection for human avatar animation 30
Bibliography **31**

© ISO/IEC 2025 – All rights reserved

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).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

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 third edition cancels and replaces the second edition (ISO/IEC 23093-2:2022), which has been technically revised.

The main changes are as follows:

— addition of complementary use cases;

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

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.

© ISO/IEC 2025 – All rights reserved

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.

© ISO/IEC 2025 – All rights reserved

Information technology — Internet of media things —

Part 1:
Architecture

1 Scope

This document describes the architecture of systems for the internet of media things. It also includes a comprehensive set of use cases that can be deployed on such an architecture.

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:

— 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

© ISO/IEC 2025 – All rights reserved

3.1.7
internet of media things
IoMT
special subset of *IoT* (3.2.8) whose main functionalities are related to media processing

3.1.8
IoMT device
IoT (3.2.8) device that contains more than one *MThing* (3.1.11)

3.1.9
IoMT system
MSystem
IoT (3.2.8) system whose main functionality is related to media processing

3.1.10
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.

3.1.11
media thing
MThing
thing (3.2.18) capable of sensing, acquiring, actuating, or processing of media or metadata

3.1.12
media token
virtual token for accessing functionalities, resources and data of media things

3.1.13
microphone
entity capable of capture and transform acoustic waves into changes in electric currents or voltage, used in recording or transmitting sound

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

Note 1 to entry: A component can expose or use interfaces[1] (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.

3.2.3
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.

3.2.4
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.

© ISO/IEC 2025 – All rights reserved

3.2.5
entity
anything (physical or non-physical) having a distinct existence

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

3.2.7
identity
characteristics determining who or what a person or thing is

3.2.8
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

3.2.9
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.

3.2.10
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 Reference [2].

3.2.11
network
entity that connects endpoints, sources to destinations, and may itself act as a value-added element in the IoT system or services

3.2.12
process
procedure to carry out operations on data

3.2.13
physical entity
thing (3.2.18) that is discrete, identifiable and observable, and that has material existence in real world

3.2.14
resource
any element of a data processing system needed to perform required operations

Note 1 to entry: See ISO/IEC 2382.

3.2.15
sensor
device that observes and measures a physical property of a natural phenomenon or man-made process and converts that measurement into a signal

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

3.2.16
service
distinct part of the functionality that is provided by an entity through interfaces

© ISO/IEC 2025 – All rights reserved

3.2.17
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

3.2.18
thing
any entity that can communicate with other entities

3.2.19
user
human or any digital entity that is interested in interacting with a particular physical object

3.2.20
visual
any object perceptible by the sense of sight

4 Architecture

The documents on MPEG IoMT developed by JTC 1/SC 29 (formally known the ISO/IEC 23093 series) are composed of 6 parts.

The first four (namely Parts 1, 2, 3 and 4) encompass the definition of the architectural framework, the specification of application programming interfaces (APIs) and compressed representation of data flowing among media things, as well as reference software resources. 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 to access valuable functionalities, resources, and data from media things. This way, the media things can be represented in industrial applications as abstract entities, providing their functionalities in a standard way.

Additionally, in Part 5, the MPEG IoMT further deal with an incremental industrial need, namely the possibility of defining autonomous services over media things. In this respect, the mission to be performed by various media things that collaboratively achieve a high-level task is described in a standard way that can also be exchanged with external processing resources (like blockchains).

Finally, Part 6 provides standard representations for the input/output data of AI powering IoMT Analysers.

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;

— sensed data (raw or processed data) 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;

— MThing mission commands and control information, with reference to interface 4 and 5;

— MThing media data formats and API for distributed AI processing, with reference to interface 6.

© ISO/IEC 2025 – All rights reserved