**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2016/N18879**

**October 2019, Geneva, CH**

|  |  |
| --- | --- |
| **Source** | **IoMT / Communications** |
| **Status** | **Output document** |
| **Title** | **IoMT White Paper** |
| **Author** | M. Mitrea |

# ISO23093: Internet of Media Things

ISO/IEC 23093 series provide architecture and specify APIs and compressed representation of data flowing between Media Things. This standard is structured into four parts:

* Part 1: ***Architecture***
* Part 2: ***Discovery and communication API***
* Part 3: ***Media data formats and API***
* Part 4: ***Reference software and conformance***

# Rationale for the IoMT standard

Internet of Things (IoT) is a paradigm for making objects we deal within our everyday life more informative, interactive and intelligent. Cellular network operators as the main providers of wireless connectivity are leveraging on the opportunity by standardizing and deploying Machine-to-Machine (M2M) services, thus enabling surge of IoT-based applications and increasing their user base to tens or hundreds of billions of IoT-enabled devices. M2M service platforms enable massive scale communication among machines, providing support for а plethora of new services (connected cities, industrial application, smart homes, intelligent transportation, remote health-care, etc.).

Within the large field of IoTs, the media-centric IoT applications and services are challenging and appealing to users, as they offer provision, interpretation, representation and analysis of multi-media content collected from media devices (cameras, microphones). Video and audio are the preferred informative medium and the current explosion of traffic over mobile cellular networks is due to their delivery.

MPEG has introduced the notion of Media Thing (MThing), which is defined as a Thing capable of sensing, acquiring, actuating, or processing media or metadata. The notion of IoMT, an extension of the IoT paradigm for visual and audio devices (cameras and microphones), has also been introduced by MPEG.

However, there is a lack of standards enabling communication, storage, analysis, interpretation, and retrieval of media big data emerging from massive-scale IoMT devices. Such standards can provide the basis for a novel machine-based visual/audio content service, and fill the gap in M2M services and applications. These standards make it possible to realize large-scale interoperable IoMT applications, such as massive-scale automated surveillance, video-enabled intelligent transportation systems, audio-video-based environmental monitoring, visual-content based smart city applications, video-crowd-sourcing platforms, etc. Figure 1 illustrates possible scenarios for IoMT applications.

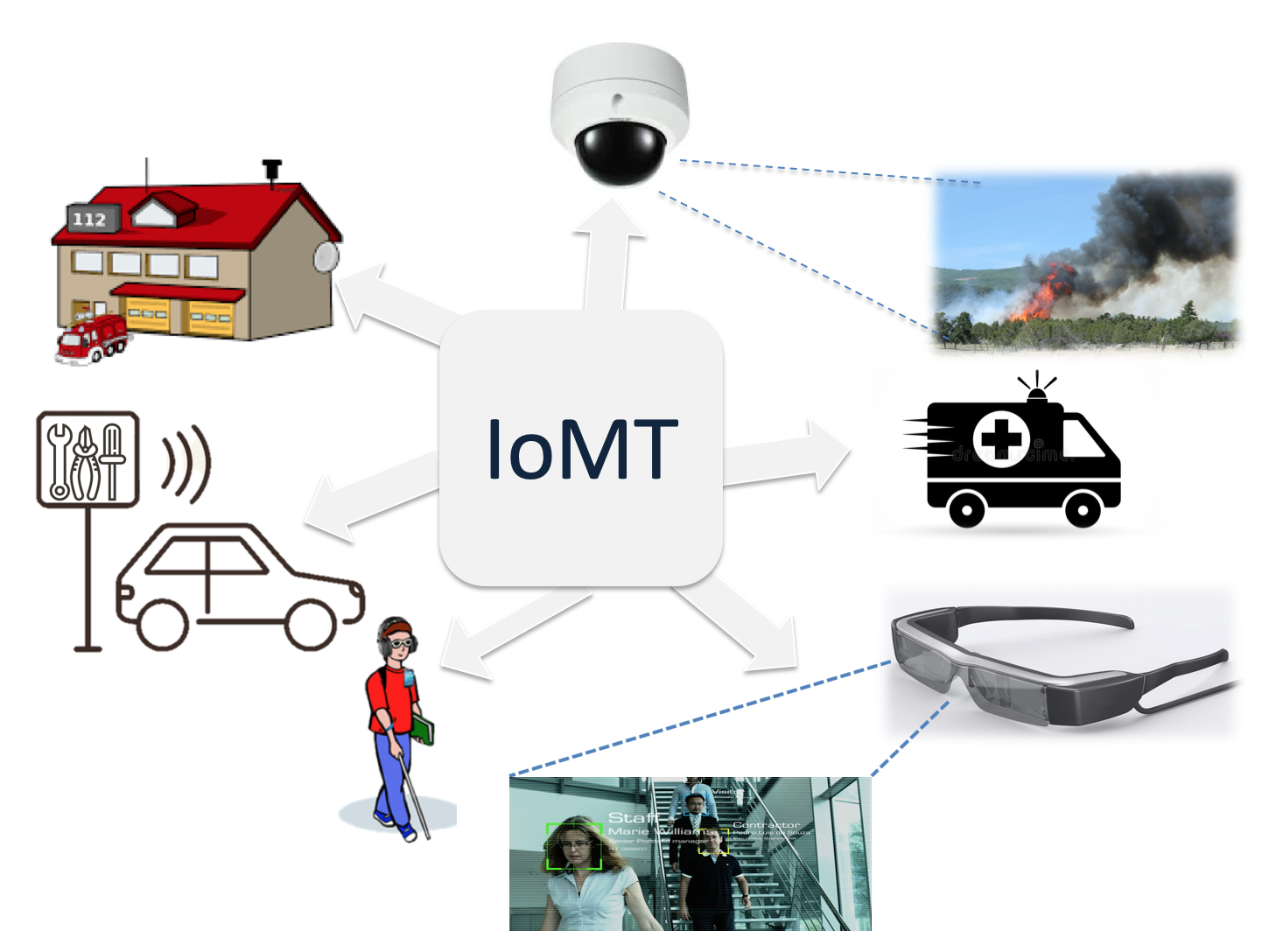


Figure 1: Examples of IoMT applications

IoMT applications and services can be designed and implemented by using a large variety of technologies available such as information-centric networking, media analytics, cloud technologies, big data for media content, content streaming and caching… However, there are still demands for orchestration and synchronization of these components. This can be done in a sustainable manner by using open standards and IoMT fills this gap.

# Scope of the IoMT standard

The IoMT architecture, generic with respect to the targeted application, specifies three types of interfaces among and between the various entities involved in any IoMT related application.

The APIs for the Media Things facilitate the discovery, connection and efficient data exchange among Media Things in the network. The APIs also provide the 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. ISO/IEC 23093 specifies data formats of input and output for media sensors, media actuators, media storages, media analyzers, etc. as well as the reference software and conformance bit streams for data formats.

Media analyzers can process data, and be cascaded in order to extract semantic information. The standard does not specify how the process of sensing and analyzing is carried out but only the interfaces among Media Things.

## IoMT architecture

The global IoMT architecture is presented in Figure 2, which identifies a set of interfaces, protocols and associated media-related information representations:

* User commands (setup information) between a system manager and an MThing, *cf*. Interface 1.
* User commands (setup information) forwarded by an MThing to another MThing, possibly in a modified form (*e.g.*, subset of 1), *cf*. Interface 1’.
* Sensed data (raw or processed data, compressed or semantic extraction) and actuation information, cf. Interface 2.
* Transformation for individual data exchange channels (*e.g*. for transmission), *cf*. Interface 2’.
* MThing characteristics, discovery, *cf*. Interface 3.

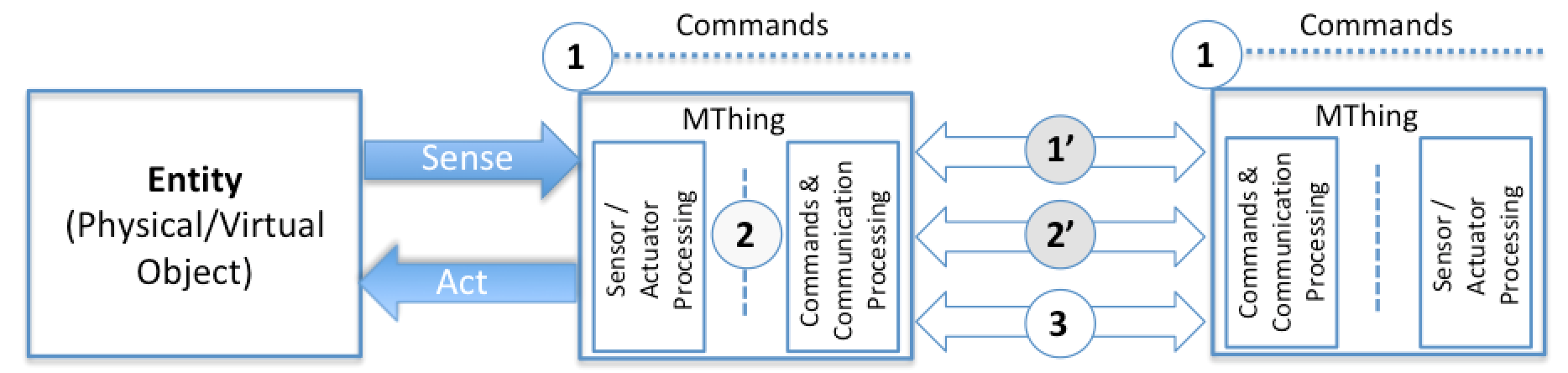


Figure 2: IoMT Architecture

## Discovery and communication API

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 suite specifies data formats of input and output for media sensors, actuators, storages, analyzers, etc. Media analyzers produce analyzed data from data sensed by media sensor and can be cascaded in order to extract semantic information.

The standard specifies the APIs to discover and communicate between Media Things in the network, and another APIs to facilitate transactions between Media Things, which are implemented as follows;

* discover other MThing(s) in the network;
* connect/disconnect MThing(s);
* transactions (e.g., payments) using media tokens between MThings.

## Media data formats and API

While the normative APIs are specified in the Part 2 of the standard, the Part 3 contains the tools to describe data exchanged between media things (e.g. media sensors, media actuators, media analyzers, media storages) for their APIs, which is illustrated in Figure 3.

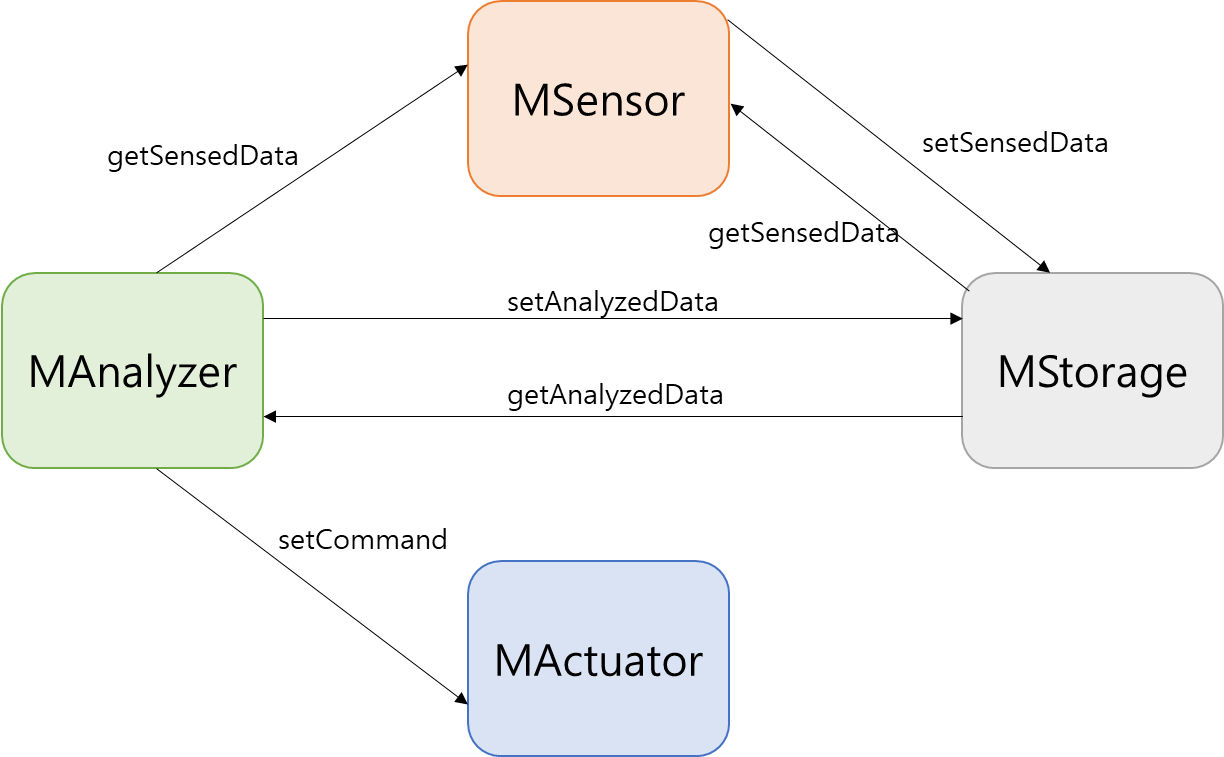


Figure 3 An example of function invocation between MThings

## Reference software and conformance

The Part 4 of the standard specifies the conformance and reference software implementing the normative clauses in Part 3. This part also provides means for conformance testing, i.e. bit-streams.

# Usage of IoMT

MPEG identified 26 use-cases for IoMT, which are classified in five main categories such as: *Smart spaces: Monitoring and control with network of audio-video cameras*, *Smart spaces: Multi-modal guided navigation*, *Smart audio/video environments in smart cities*, *Smart multi-modal collaborative health services* and *Blockchain usage for IoMT transactions authentication and monetizing*.

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

The large variety of sensors, actuators, displays, and computational elements are used in our day-by-day professional and private space. The better and easier access to services leads to 11 IoMT use cases, mainly related to the processing of the video information.

* + Human tracking with multiple network cameras
  + Automatic title generation
  + Intelligent firefighting with IP surveillance cameras
  + 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 contents recommendation using information from IoMT devices

## Smart spaces: Multi-modal guided navigation

Multimodal information can be processed and fused inside IoMT systems in order to provide the user with enhanced navigation experience as being examplified through 5 IoMT use cases:

* + Blind person assistant system
  + 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

## Smart audio/video environments in smart cities

A city becomes smart when its traditional infrastructures are combined with innovative technologies to improve life of its citizen and business activities in a sustainable way. The realization of smart cities involves aggregating operation of different subsystems (smart spaces) that will cause to retain their primary private function but must interact with each other in order to fulfill more global objectives.

Homogeneous or heterogeneous subsystems, operating inside the same area or in spatially correlated areas, can be aggregated.

The IoMT standard considers four related yet different aspects related to smart environments:

* + 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

## Smart multi-modal collaborative health services

It is currently accepted that the cost of healthcare can be reduced while improving the quality of life of patients, by integrating professional and user created data. To illustrate this trend, 5 use cases are considered in the IoMT standard:

* + 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 smartglass

## Blockchain usage for IoMT transactions authentication and monetizing

The authentication and monetization of the transactions related to MThings can be realized by blockchain solution by using the IoMT standard, as illustrated by two use cases:

* + Reward function in IoMT by using blockchains
  + Content authentication with blockchains