

## Internet of Medical Things (IoMT): Acquiring and Transforming Data into HL7 FHIR through 5G Network Slicing

Argyro Mavrogiorgou <sup>a,\*</sup>, Athanasios Kiourtis <sup>a</sup>, Marios Touloupou <sup>a</sup>,  
Evgenia Kapassa <sup>a</sup>, Dimosthenis Kyriazis <sup>a</sup>

<sup>a</sup> Department of Digital Systems, University of Piraeus, Piraeus, 18534, Greece

### Abstract

The Healthcare 4.0 era is surrounded by challenges varying from the Internet of Medical Things (IoMT) devices' data collection, integration and interpretation. Several techniques have been developed that however do not propose solutions that can be applied to different scenarios or domains. When dealing with healthcare data, based on the severity and the application of their results, they should be provided almost in real-time, without any errors, inconsistencies or misunderstandings. Henceforth, in this manuscript a platform is proposed for efficiently managing healthcare data, by taking advantage of the latest techniques in Data Acquisition, 5G Network Slicing and Data Interoperability. In this platform, IoMT devices' data and network specifications can be acquired and segmented in different 5G network slices according to the severity and the computation requirements of different medical scenarios. In sequel, transformations are performed on the data of each network slice to address data heterogeneity issues, and provide the data of the same network slices into HL7 FHIR-compliant format, for further analysis.

### Keywords:

IoMT; HL7 FHIR;  
Network Slices;  
5G Networks;  
Healthcare Interoperability;  
Data Acquisition.

### Article History:

**Received:** 17 January 2019  
**Accepted:** 01 April 2019

## 1- Introduction

Most of the current researches are dealing with how medical and health-monitoring devices, clinical wearables, and remote sensors can contribute to better health for patients, and a more efficient healthcare that can drive better systems, population, and patient outcomes [1]. For the past years, medical device manufacturers and suppliers have been facing more challenges in product development to bring their devices faster and easier to the market, facing increasing complexities mainly in how to achieve and gather quicker, more reliable, interoperable and efficient measurements. Current healthcare – known as Healthcare 4.0 [2], is considered as one of the fastest industries to adopt the Internet of Things (IoT) technologies, aiming to help in personalized services, reduce operating costs, and improve patient care and quality of life. Such thing has led to the marriage of the IoT and medicine, generating the concept of Internet of Medical Things (IoMT) [3]. In this domain, IoMT is promising changes in the way that the medical industry delivers healthcare, as almost all of the IoMT devices are linked to cloud platforms, while they are equipped with cellular or wireless communication, thus allowing the machine-to-machine communication.

Nevertheless, for most patients and care providers, the promises of the IoT and IoMT have not yet introduced dramatic changes in how these stakeholders are experiencing healthcare. What is required is higher connection speeds that will be transforming the relationship of the patients and the healthcare providers, and integrating electronic communications into medical care, which can be achieved through the arrival of the 5G networks [4]. According to [5], 5G will enable more than \$1 trillion in devices and services for the global healthcare sector. However, most of these devices are facing high degrees of heterogeneity, since they are having different capabilities, or functionalities. In this context, the challenge remains on how medical data can be extracted out of these heterogeneous devices, and how this extracted information

\* **CONTACT:** Margy@unipi.gr

**DOI:** <http://dx.doi.org/10.28991/esj-2019-01170>

© This is an open access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>).

will be integrated, by also managing the abstractions and the different nature of these devices. Current data collection and integration techniques are facing difficulties in adapting and adopting to these changes and requirements, since they are mostly designed for specific areas and domains. However, the problem does not only remain to the difficulty of data acquisition among different devices, but also to the ingested data heterogeneity [6]. The latter is considered as a challenge for many decades in the healthcare domain, as tons of heterogeneous data are produced every second. By having data in different formats and terminologies is leading to non-consistent ways in the identification of patients, difficulties for sending, receiving and managing information between healthcare systems, and blocking knowledge that usually results into higher rates of mortality. Henceforth, data interoperability techniques require healthcare data to be restructured into a common format and standard terminologies, while being linked to other data sources.

To go beyond the hype and the use of the existing solutions, in this manuscript we are going to targeted address these challenges. In more detail, an end-to-end platform is introduced where IoMT devices that can serve different application scenarios (e.g. vital signs tracking, home monitoring) are connected through Wi-Fi. Through the proposed Data Acquisition mechanism, the IoMT devices' data are being acquired, as well as their network specifications, in order to be segmented in different 5G network slices through the Slicing Management mechanism, each one serving the computation requirements of the different medical scenarios. In sequel, a Data Interoperability mechanism is performed on each different network slice to address data heterogeneity, and provide the data that belong to the same slices into a common and interpretable format for further analysis.

The rest of this paper is organized as follows. Section 2 presents the background, and the challenges that led us to this research, with regards to the Healthcare 4.0 era, 5G networks, data acquisition and heterogeneity. Section 3 describes the materials and methods that were used in to order to develop the proposed platform for achieving interoperable data collection and transmission of high availability, while Section 4 states the evaluation of the platform, through an end-to-end scenario. Finally, in Section 5 the derived results are discussed, analysing our conclusions and future goals

## **2- Background & Challenges**

### ***2-1- Healthcare 4.0***

Healthcare is considered to be a complex industry including several stakeholders involved such as patients, caregivers, healthcare providers, hospitals and pharmaceutical companies, containing several restrictions, rules, and regulations which several times lead to less than adequate care accorded to patients. Most of these stakeholders are being currently affected by the promises and integration of new technologies that support the Healthcare domain, being impacted by the introduction of data enabled technologies, which has led to the generation of the Healthcare 4.0. Healthcare 4.0 envisions the creation of a system where everyone is connected through wearable devices, while every data point of patients is getting recorded without being affected by their current location. That is why according to BCC Research [7] it is estimated that the healthcare market will grow from \$521.2 billion in 2017 to \$674.5 billion by 2022, making companies to strive towards new markets and innovative products in order to stay competitive. Among the biggest challenges that the early adoption of Healthcare 4.0 is facing has to do with the lack of coherence in digital enablement across different stakeholders. For instance, many hospitals may not be able to implement a complete digital transformation, due to financial issues. What is more, the growing number of medical devices with millions of product variations, presents an additional challenge for the healthcare supply chain. With over 6 million medical devices on the market, information about them is often difficult to be managed, since all of this require manual entry, leading to high risks of faulty, inconsistent, and incomplete data. Therefore, the healthcare industry demands not only a high level of technological transformation but also efficient, reliable and fast results so as to fully produce reliable decisions that aim in improving personalized health. Currently, multiple fragmentations in the healthcare domain do not allow a seamless experience for patients, making them to interact with multiple parties and struggle with storing, analyzing and interpreting their medical records. All the challenges that are involved in the industry need to be confronted first, in order for the Healthcare 4.0 to become a reality.

### ***2-2- 5G in Healthcare 4.0***

While many things on the road to 5G are uncertain, it is easy to envision the emergence of new and innovative applications. 5G technologies allow a significantly higher data capacity and extremely fast response time, opening up completely new potential applications for a fully connected society. Especially in the Healthcare domain, this constitutes a prerequisite, as faster and more accurate results are needed for patients' remote monitoring, and examination, along with sophisticated equipment imaging, which can lead to additional pressure on healthcare network operators. This often increases congestion and slows network speeds, especially for healthcare providers who could interact with dozens of patients per day, leading to delayed patient care, which could hurt the results in the long run. Taking into consideration that the usage of IoT continues to grow, constantly increasing the volume of the data for networks, 5G technologies come to resolve the emerging challenges. In this context, 5G promises quick transmission of large medical imaging files that often needed to be exchanged between patients and caregivers, avoiding low bandwidth networks' performance,

and thus time-consuming data transmission [8]. Apart from this, 5G promises to cope with the reliability and the connectivity of the emerging devices. More particularly, Gartner states [9] that around 8.4 billion IoT devices were in use in 2017, up 31% from 2016, and this will likely reach 20.4 billion by 2020, the exact point of history when standardized 5G networks will start to hit the market and deliver unprecedented levels of connectivity. In addition, several healthcare functions begin using Artificial Intelligence (AI) to identify possible diseases and decide on the best treatment plan for a particular patient. At the same time, network operators are facing tremendous traffic increases with the introduction of IoMT devices, especially due to content rich multimedia and health applications. Thus, by switching to 5G networks, Healthcare 4.0 can use the AI tools needed providing the best care possible and patient experience, while reducing healthcare costs.

### ***2-3-Heterogeneous Data Acquisition***

In the last decade, there has been a transition from a data-poor to a data-rich world with the promise of unparalleled intelligence. Much of this unprecedented increase in data generation can be attributed to the abundance of thousands of mobile devices, wearables, and sensors that fall under the era of IoT. According to Machina Research [10], the total number of connected devices is expected to be 27 billion by 2024, while according to Cisco [11], there will belong nearly 1.5 mobile devices per capita by 2020, and more than 601 million of wearable devices will be in use. These numbers only confirm that the IoT remains one of the biggest trends in 2018. Such thing results into a myriad of heterogeneous devices that will be connected to the IoT world in the near future, producing data of different types that may have been collected at different rates and time scales, by different devices. Most of these devices are typically characterized by a high degree of heterogeneity, in terms of having different specifications, capabilities, or network specifications, needing to be easily manageable by single IoT platforms and systems. Therefore, devices' integration is one of the most challenging research topics in the IoT era, revealing the need of the existence of a global interface that will be able to manage all these heterogeneous devices. Especially for the IoMT devices that are widely adopted and used in Healthcare 4.0, the need for integration of these devices is an issue of vital importance. For example, patients' telemonitoring at home, requires a wide variety of sensors and devices that usually present a set of very specific features and characteristics, thus being difficult to be integrated into different platforms in order to offer each patient's medical data. Without the effective integration of the heterogeneous medical devices' data, no patients' data will be collected, and thus no further actions will be possible to be performed. Without the effective integration of the medical devices' data, no patients' data will be collected, and thus no further actions will be possible to be performed, failing in finally achieving integrated care.

### ***2-4-Data Heterogeneity***

Vast amounts of heterogeneous medical data are becoming available in various healthcare organizations and sensors, thus completely reshaping the Healthcare 4.0. Hundreds of hospitals and healthcare systems deal everyday with challenges in extracting data from silos, affecting both patient care and medical research. Currently, the rapidly increasing availability of health records is pushing towards the adoption of data-driven approaches, bringing the opportunities to automate healthcare related tasks, providing better disease detection, more accurate prognosis, faster clinical research advance and better fit for patient management. Nevertheless, the healthcare organizations are still facing many difficulties in implementing, maintaining and upgrading their healthcare systems, including many challenges in the technical, security and human interaction fields. Heterogeneity, timeliness, complexity, noise and incompleteness are among the most common challenges of creating value from data. Among others, what is missing is an integrated data exchange system that according to [12] could reduce Medicare spending by more than \$3 billion each year. In order to exchange data with as many stakeholders as possible, interoperability is the only way for letting systems interact with each other, in order to take the complete image of a patient, being considered a necessity in electronic healthcare systems for resolving data heterogeneity issues. A recent study [13] estimated that the global healthcare middleware market is expected to reach \$3.07 billion by 2023 from \$1.90 billion in 2018, in order to overcome healthcare interoperability issues. Among the different existing researches and standards, the Health Level Seven International (HL7) organization provides the development and the framework of standards that are widely used in the medical market and research, with HL7 Fast Healthcare Interoperability Resources (FHIR) [14] being the latest standard created by this organization, for the exchange of clinical information.

### ***2-5-Key Contributions of the Proposed Approach***

It is undeniable that the Healthcare 4.0 era is surrounded by challenges that can be allocated in different domains. Such challenges can vary from the data integration to the data manipulation, or from the IoMT devices' requirements to their data interpretation. It is undeniable that when dealing with healthcare related data, based on the severity and the application of their results, they should be provided almost in real-time, without any errors, inconsistencies or misunderstandings to the stakeholders, for further analysis. Such thing can be considered of vital importance since based on [13], in 2016 hospitals and healthcare providers' offices nationwide might have avoided nearly 2000 patient deaths,

if medical staff and patients communicated better. Based on these challenges and requirements, a trustworthy and efficient solution is provided in this manuscript for efficiently managing healthcare related data, by taking advantage of the latest technologies and techniques that have been currently developed under the context of Healthcare 4.0, 5G networking, Data Acquisition and Data Interoperability. These State of the Art techniques are thoroughly described in Section 3, addressing all the aforementioned challenges through an end-to-end application that can gather healthcare data out of 5G IoMT devices, through allocating the device's requirements and making its provided data interoperable, by using the concept of network slicing according to specific applications' criteria and requirements.

### 3- Materials and Methods

The developed platform deals with heterogeneous IoMT devices that are designed to serve different application scenarios and provide their data in different frequencies among them. Henceforth, this platform initially addresses the data collection part, where a Data Acquisition mechanism is implemented for acquiring the Wi-Fi connected devices' data being independent of the nature of the IoMT devices. At the same time, this mechanism gathers the network specifications of each IoMT device, in order to firstly provide the latter to the Slicing Management mechanism that is responsible for segmenting the aforementioned in different network slices, according to the different computational requirements of specific medical scenarios, and secondly the acquired healthcare data to the Data Interoperability mechanism. As soon as the slices are constructed, the Data Interoperability mechanism takes place, being performed on each different network slice in order to address data heterogeneity issues, and deliver data into a common and interpretable format, through identifying both syntactic and semantic similarities of the provided data. Henceforth, interoperable data can be given for further analysis to different stakeholders, in different periods based on the needs and the severity of each different medical scenario. Figure 1 depicts the overall architecture of the developed platform, while all the aforementioned steps are explained in the following sub-sections.

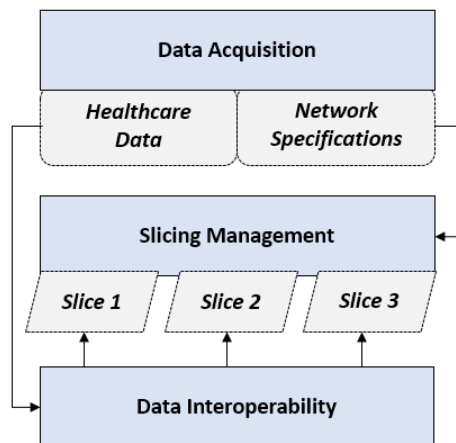


Figure 1. Overall Architecture

It should be noted that three medical scenarios are being considered for allocating the IoMT devices into three different slices, containing different computation requirements: a) Massive IoMT scenario including requirements for instantaneously capturing, collecting and receiving patient's medical data, physical activity and vital signs, b) Enhanced Mobile Broadband scenario demanding enhanced data rates and high connectivity for remote doctor examination, and c) Mission-Critical Services scenario requiring boosted network reliability for healthcare providers deliver healthcare in critical cases.

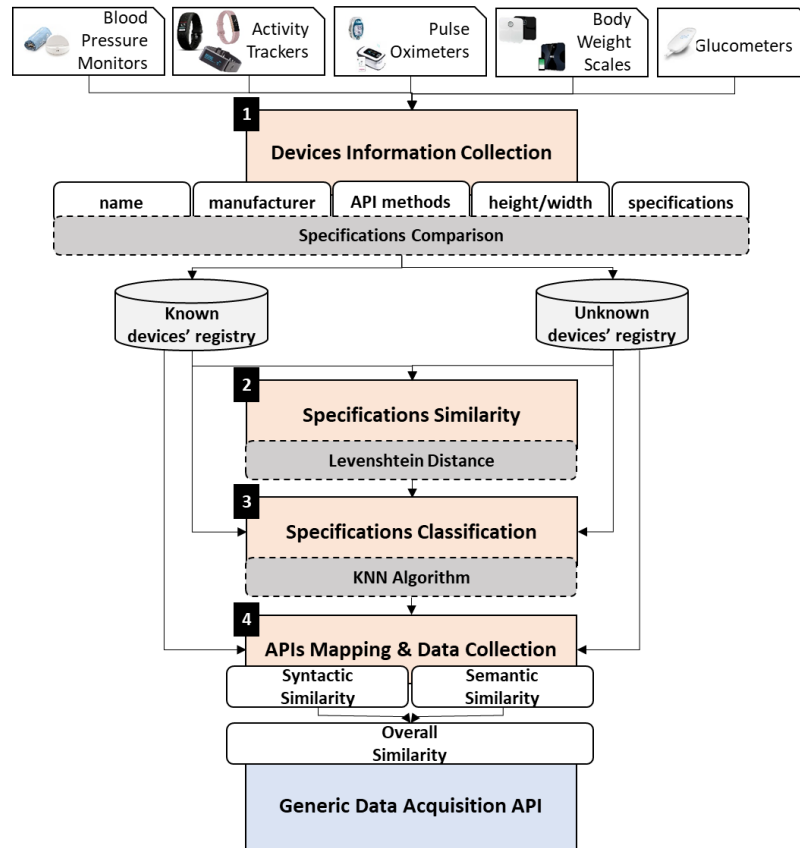
#### 3-1- Data Acquisition

In the first stage, as soon as all the devices have been connected to the platform through a Wi-Fi Gateway, the collection of these devices' specifications occurs, accompanied by their data collection. In more detail, the Data Acquisition mechanism is responsible for collecting all the specifications, APIs, and network specifications by the different connected devices, to finally gather their data. In order to achieve that, the proposed mechanism implements the following four steps that are depicted in Figure 2.

##### 3-1-1- Devices Information Collection

The objective of the Devices Information Collection is the collection of the connected devices' specifications, network specifications, and APIs. In more detail, through the Wi-Fi Gateway the mechanism initially gathers information about the connected devices' network specifications (i.e. reliability, availability, scalability, bandwidth, capacity, cache, end-to-end delay, and latency), in combination with their specifications (i.e. names, and Internet Protocol (IP) addresses). As soon as the IP address of each device is collected, through the latter it becomes feasible to

identify each device's Media Access Control (MAC) and as a result the corresponding Organizationally Unique Identifier (OUI) (i.e. manufacturer), using the MAC Vendors API [15]. However, apart from the manufacturer and the name of each device, the mechanism gathers information about each device's height and width, using the image-processing library of the OpenCV [16], so as to be used and combined with the manufacturer and the name of the corresponding devices in the next steps.



**Figure 2. Data Acquisition Mechanism**

As soon as all these devices' specifications are captured, the latter are queried and programatically compared with the corresponding specifications of the devices that already exist in a registry that includes known IoMT devices (i.e. known devices' registry). Depending on whether the connected devices already exist in this registry or not, they are characterized as either known or unknown correspondingly. More specifically, in the case that the specifications of the connected devices match exactly with the specifications of the devices of the known devices' registry, then the connected devices are automatically considered to be known (i.e. devices of known type), whereas if there is no resemblance or even a partial resemblance among them, then the connected devices are considered to be unknown (i.e. devices of unknown type). After capturing the specifications of the different devices and categorizing them, their information storage occurs, where the specifications of the unknown devices are stored into a new registry (i.e. unknown devices' registry), whilst the specifications of the known devices are not stored, as they already exist in the known devices' registry.

However, apart from this information, the mechanism requires knowledge about the devices' APIs. Therefore, in sequel, since we have knowledge for all the devices' manufacturers and nature (i.e. either known or unknown devices), we are able to retrieve from each manufacturer's website the corresponding online development documentations, gathering information about each device's (i) API Uniform Resource Locator (URL) paths for accessing the different methods that are included into the API, and (ii) API methods' descriptions. To this end it should be noted that since a manufacturer may produce more than one different types of devices, its online documentation will contain methods for all the different types of devices. Henceforth, we store for each different known device in the known devices' registry the corresponding API methods' information about each device's specific device type, whereas for the unknown devices we store in the unknown devices' registry all the APIs methods' information, as we do not know the type of the devices yet, in order to store the corresponding methods.

### 3-1-2- Specifications Similarity

The objective of the Specifications Similarity is to identify the syntactic similarity that exists between the devices of the known devices' registry and each connected device's specifications, to categorize the unknown connected devices



to the device type of the known ones, depending on the syntactic similarities that they will have among their names and manufacturers. It is worth mentioning that this step takes place only for the connected devices that are of unknown device type, as the known ones already exist in the known devices' registry and their device type is already known. Henceforth, the names and the manufacturers of the unknown connected devices are taken as an input, and are syntactically compared with the names and manufacturers of the devices of the known devices' registry, using the Levenshtein string metric [17]. In order to achieve the aforementioned, the mechanism proposed in [18] is implemented.

Based on the results of this process, the device of the known devices' registry with higher probability of similarity with each unknown connected device is considered that it is identical or almost identical with the corresponding unknown device, and as a result, the device type of each unknown connected device becomes known. To this end, it should be noted that in order to consider the results as trustful and reliable, these must exceed the set threshold of 85%. Henceforth, each unknown device that transcends this threshold is considered as known, being finally stored with its specifications in the known devices' registry.

### ***3-1-3- Specifications Classification***

There might be cases where the unknown connected devices do not exceed the set threshold, indicating that neither the name nor the manufacturer of them match with any corresponding specifications of the devices of the known devices' registry. For that reason, the objective of the Specifications Classification is to classify each unknown device's specifications based upon the K-Nearest Neighbors (KNN) algorithm [19], to group all the unknown connected devices with the existing known devices according to their specifications. In order to achieve this process, the classification process proposed in [20] is implemented.

Based on the results of this process, each unknown connected device may have different degrees of resemblance with one or more different known devices. Henceforth, in order to conclude into efficient results, we assume that each unknown device will be classified to the type of category of the device with which it has greater values of similarity (i.e. 90%), and thus it will be finally treated as a device of known type. As a result, each one of these connected recognized devices in combination with their specifications are finally stored in the known devices' registry.

### ***3-1-4- APIs Mapping & Data Collection***

Even if so far we have achieved to discover and automatically identify heterogeneous devices of unknown device type, their API methods' nature is still remaining unknown. As our primary goal is to recognize the API methods that are responsible for collecting the devices' data, the objective of the APIs Mapping & Data Collection is to carry out the mapping of the different devices' API methods, by syntactically and semantically comparing all the API methods of the devices of the same type, thus understanding the nature and the meaning of the different API methods that are used by the different devices, in order to finally collect their data. As described in the Devices Information Collection step, through the manufacturer of each device, we are able to gather information about the devices' (i) API URL paths for accessing the different methods, and (ii) API methods' descriptions. However, the API URL paths are not providing internal details for each method in terms of its source code and its functionality purpose, but they are only providing descriptions on what does each method is supposed to do, when it should be used and how it should be called. Therefore, through this step, it becomes feasible to identify the API methods of a prior-unknown connected device, for which we only have access to the names of its API methods. In order to achieve that, the mechanism proposed in [21] is implemented.

Based on the results of this process, as soon as all the API methods for gathering data have been identified, the implementation of the generic data acquisition API occurs. More particularly, the latter constitutes a unified API that merges into a single data method all the different devices' API methods that are responsible for collecting data. Thus, this API is responsible for finally collecting the data from all the connected devices.

## ***3-2- Slicing Management***

In the second stage, the collected data are promoted among with the devices' network specifications for further analysis to the Slicing Management mechanism. In more detail, the proposed mechanism's design relies on 5G network slicing, which is able to let 5G Core (5GC) operators to offer parts of their networks for specific medical scenarios. Since each one of the medical scenarios impose their own set of network specifications (e.g. low latency, high bandwidth), a network slice instance is created by a group of network functions and logical resources, enabling the preparation of a completely logical network infrastructure capable to accommodate the computation requirements for each corresponding case [22]. To address all the aforementioned, the proposed Slicing Management mechanism, develops the following three steps that are depicted in Figure 3:

### 3-2-1- OS Management & Orchestration

The Slicing Management mechanism interacts with the Operations Support System (OSS) [23] of the 5GC which in our case is an established infrastructure management and orchestration framework, using the Open Source (OS) Management and Orchestration (OSM - MANO) framework [24] that exposes REST APIs for controlling slicing, including management and orchestration of infrastructure resources.

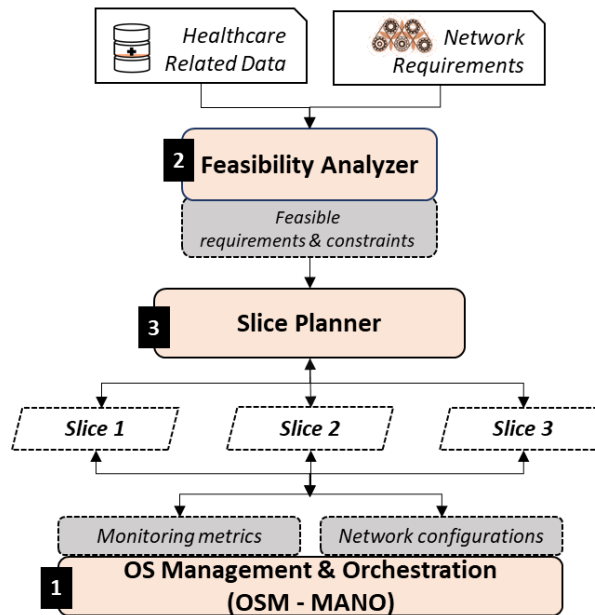


Figure 3. Slicing Management Mechanism

### 3-2-2- Feasibility Analyzer

The objective of the Feasibility Analyzer is to determine whether the infrastructures have available resources, as well as whether it can afford the request of a new slice in terms of cost, instantiation and successful deployment of the requested slice. Therefore, this process is triggered after the collection of the connected devices' network specifications, in the Data Acquisition stage. To this end, it should be noted that these specifications are considered to be expressed in the form of policies, while they are used so as to formulate each requested slice's expectations and targets, thus develop the required services, processes and capabilities, inside the 5GC.

### 3-2-3- Slice Planner

Sequentially, as soon as the network specifications have been successfully analyzed, the Slice Planner takes place, in order to adopt the network-sharing concept in the form of separate building blocks, on behalf of the network operator. For that purpose, in order to create the needed network slices, the MANO gathers the medical scenarios' computational demands, and according to their computation requirements, it deploys a network slice that responds to their requests. To achieve this, the Slice Planner supports the provisioning of network slices using the feasible network specifications per medical scenario, obtained from the connected IoMT devices. Based on a high-level description of the medical scenario and the network specifications extracted from the Data Acquisition mechanism, we deploy an end-to-end slices deployment planner. To achieve that, the planner uses the dynamic discovery of unfractured resources, as well as the monitoring data coming from MANO. To this end, it should be pointed out that a slice is composed of a set of Virtual Network Functions (VNFs) in order to create a chained Network Service for the Data Interoperability mechanism, implemented on top of the infrastructure, whereas each one of them is isolated from the other slices, having a dedicated network proceeding. Henceforth, based on the computation requirements of each medical scenario, three different slices are deployed, in which the IoMT devices' data are provided in order to be given as an input to the Data Interoperability mechanism. To successfully construct these slices, we consider that a) the Massive IoMT scenario demands IoMT devices with high levels of reliability, scalability, low end-to-end delay and high bandwidth, b) the Enhanced Mobile Broadband scenario demands high capacity, cache, low end-to-end delay, and high bandwidth, whilst c) the Mission-Critical Services scenario demands low latency, high reliability, low end-to-end delay and high bandwidth.

### 3-3- Data Interoperability

The last stage includes a mechanism that is capable of constructing healthcare ontologies from the different datasets that have been acquired from the previous layer, and then identifying similarities and common links between these ontologies and the ontologies that represent the HL7 FHIR Resources, for finally translating the healthcare datasets into

the HL7 FHIR standard. The Data Interoperability mechanism is deployed in the 5GC in form of a chained Network Service, considering the three different medical scenarios that are allocated into the three network slices accordingly. The Data Interoperability mechanism behaves in the same way for each network slice, with differences in the execution speed according to the medical scenario's computation requirements, implementing the following four steps that are depicted in Figure 4:

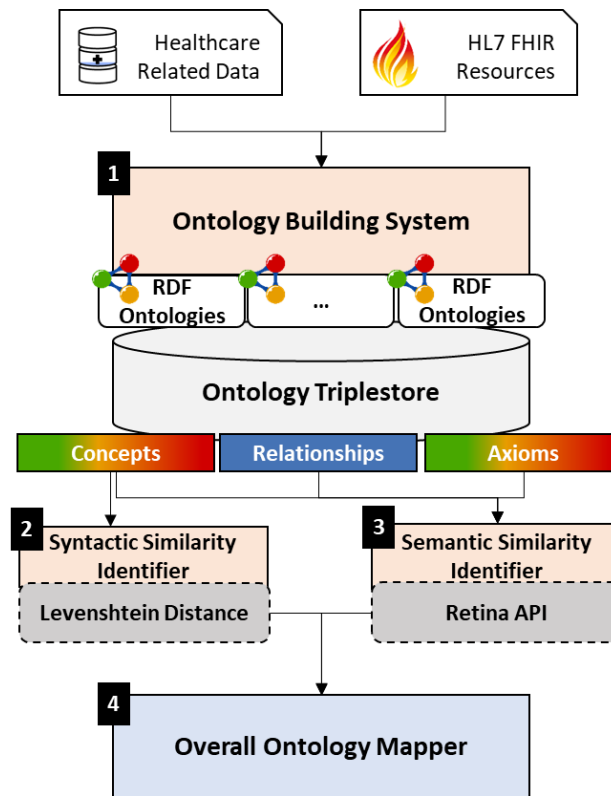


Figure 4. Data Interoperability Mechanism

### 3-3-1- Ontology Building System

Initially, the Ontology Building System takes place, being responsible to gather the eXtensible Markup Language (XML) schema of the healthcare datasets – despite their primary format, and the HL7 FHIR resources, and build the ontologies. The generation of the ontologies is initially achieved through the transformation of the healthcare dataset into XML format, by identifying the different elements of the document, as well as the parent and the child nodes, along with their attributes. Sequentially, the transformation of the created XML document into XML schema using the Trang API [25] occurs, followed by the analysis of the XML schema using the XML Schema Object Model (XSOM) [26], where the output of the XSOM is used as an input for the Java Universal Network/Graph framework (JUNG) [27] that is used for graph-based manipulations. The latter generates a XML Schema Graph (XSG) [28] which is used as an input to the Apache Jena [29], in order to generate RDF entities (i.e. ontologies), emerging from complex types, element group declarations, and attribute-group declarations according to specific matching rules. It should be noted that the concepts and the properties that were identified from the created ontologies are stored through an RDF triplestore [30] that can store and retrieve the different: (i) concepts, (ii) relationships, and (iii) axioms of the derived ontologies.

### 3-3-2- Syntactic Similarity Identifier

Afterwards, the Syntactic Similarity Identifier takes place, in order to provide a way for identifying the syntactic similarity between the healthcare datasets and the HL7 FHIR Resources ontologies, and provide the probability that a specific ontology is syntactically same with another ontology. In our case, in order to measure the syntactic similarity between two ontologies, firstly the values of the different ontological concepts are gathered, and are then compared one-by-one, in order to calculate their syntactic similarity. For that reason, the Levenshtein distance measure [17] is implemented, as a measure of similarity between the two concept names. By the time that all the different combinations of comparisons have occurred, different tables of the syntactic similarity between the ontological concept names are created. Among the different degrees of similarities, this mechanism finally stores and syntactically maps the concept names and the HL7 FHIR Resources that syntactically match better.



### 3-3-3- Semantic Similarity Identifier

In sequel, the Semantic Similarity Identifier occurs, being responsible to provide the means for aligning and matching the different healthcare datasets and the HL7 FHIR Resources ontologies, according to their semantic meaning. In our case, in order to measure this semantic similarity, firstly the different concepts, along with their relationships and axioms of the stored ontologies are gathered. As soon as the different triples of information are gathered from the ontologies, these are compared one-by-one, in order to calculate their semantic similarity. For that reason, the Retina API [31] is implemented, where each instance of the triples of information is translated into semantic fingerprints, in order to compare their meanings by overlaying their semantic fingerprints and calculating their distance. By the time that all the different combinations of comparisons have occurred, different tables of the semantic similarity between the different triples of arrays are created. The purpose of the Semantic Similarity Identifier is to identify the triples of arrays of the healthcare dataset that have greater degree of resemblance with the HL7 FHIR Resources. Consequently, among the different degrees of similarities, this mechanism finally stores and semantically maps the triples of arrays and the HL7 FHIR Resources that semantically match better.

### 3-3-4- Overall Ontology Mapper

Finally, the Overall Ontology Mapper takes place, in order to provide a process being able to aggregate and merge the results that have derived from the previous mechanisms of the Syntactic and the Semantic Similarity Identifiers, in order to finally map the healthcare datasets to the proper HL7 FHIR Resources. After the calculation of the mean of the Syntactic and Semantic Similarity Identifiers, the HL7 FHIR Resource attribute with higher probability of similarity with each compared ontological concept is automatically considered that it represents the specific dataset's attribute. The same mechanism iterates in order to finally assign the ontologies of the healthcare data into their HL7 FHIR format

## 4- Evaluation

### 4-1- Working Environment

The overall platform was developed in Java SE using the NetBeans IDE v8.0.2. Current study used a processing environment with 16GB RAM, Intel i7-4790 @ 3.60 GHz x 8 CPU Cores, 2TB Storage, and Windows 10 operating system. The 5GC testbed was implemented on top of a real 5G environment, currently under deployment on behalf of the 5GTANGO EU project [32]. The provided European Telecommunications Standards Institute (ETSI) compliant 5GTANGO Service Platform is a modular Network Function Virtualization (NFV) platform that manages network services throughout their end-to-end lifecycle, from instantiation to termination. The platform combines a policy mechanism with a customizable MANO Framework to allow the operator and the service developer to join forces in managing their lifecycles to fulfil specific network slices. It should be noted that concerning the source code availability, it is not currently in fully open-source format, since it has not been finalized.

### 4-2- Experimental Results

*Data Acquisition:* In order to perform a complete testing and evaluation of the platform, 30 IoMT devices of various types and open APIs were chosen, being able to connect through a Wi-Fi Gateway with our platform. In more detail, following the Devices Information Collection step, 29 of them were recognized as of known device type (these devices can be seen in Table 4, Table 5 and Table 6), whereas 1 of them was of unknown device type. Therefore, for the 29 known devices, we had prior knowledge about seven of their specifications (i.e. name, IP address, MAC address, manufacturer, height, width, and type) outlining the type of the device that each one of them belongs to (i.e. activity tracker, blood pressure monitor, pulse oximeter, body weight scale, toothbrush, glucometer, and thermometer). Regarding the 1 unknown device, we had only prior knowledge about two of its specifications (i.e. name, and IP address), whereas its manufacturer, height, width, and type were initially unknown. Table 1 depicts a snapshot of the specifications of a known IoMT device, and a snapshot of the unknown device's specifications.

**Table 1. Fitbit Flex Specifications**

MT Specifications	Fitbit Flex	Misfit Vapor
<b>IP address</b>	104.16.65.26	104.229.25.19
<b>Mac address</b>	1800DB6C572B	Unknown
<b>Manufacturer</b>	Fitbit	Unknown
<b>Height (cm)</b>	1.5	Unknown
<b>Width (cm)</b>	1.01	Unknown
<b>Type</b>	Activity Tracker	Unknown

Following the steps of the Devices Information Collection, as soon as the IP address of the unknown device was collected, we identified the device's MAC address, and manufacturer, which was the Misfit, while we gathered information about the device's height and width, which were 3.5 cm and 3.5 cm correspondingly. Apart from these, since we acquired knowledge for the device's manufacturer (i.e. Misfit), we retrieved from Misfit website its online development documentations, gathering information about the device's API URL paths, and API methods' descriptions, following exactly the same process for the 29 known devices.

Sequentially, following the steps of the Specifications Similarity, for the unknown device's name, the syntactic similarity was calculated, between the unknown device's name and each different known device's name. The same process was followed for the unknown device's manufacturer, where the syntactic similarity was calculated again, between the unknown device's manufacturer and each different known device's manufacturer. In our case, we identified (Table 2) that Misfit Vapor had greater degree of similarity with Fitbit Aria (47.7% of overall similarity). However, since this result did not exceed the set threshold (i.e. 85%), it was not efficient enough to conclude to the type of the device, as it was not possible to map Misfit Vapor to any of the devices of the known devices' registry.

**Table 2. Overall Similarity between Misfit Vapor and Fitbit Aria**

Similarity Type	Value
Name Similarity	45.5%
Manufacturer Similarity	50%
Overall Similarity	47.7%

For that reason, the stage of the Specifications Classification had to be implemented for grouping the 1 unknown connected device with the 29 existing known devices based upon their specifications (i.e. name, manufacturer, height, width, and device type), and thus predicting the unknown devices' type. KNN was performed and predicted that the 1 unknown connected device belonged to the group of the activity trackers, as its classification's output illustrated that the unknown instance was classified correctly in a percentage of 95% into the group of the activity trackers, having approximately the same specifications with the other pre-existing 12 activity trackers.

Having recognized the connected device's type, the final stage of the mechanism was implemented in order to identify the connected device's API methods' nature, thus distinguishing the ones that are responsible for collecting data from the device. For the 1 connected device, as mentioned in the Devices Information Collection, we retrieved from Misfit's online documentation all its API methods, gathering information about how the device's API methods are called (i.e. URL path), and what they offer as functionality (i.e. description). Table 3 depicts a snapshot of an unknown API method of the unknown device.

**Table 3. Snapshot of an unknown API method**

API Information	Value
Method Identifier	Unknown1
URL Path	activity/goals?start_date=X&end_date=Y
Description	Goals of a user in the period

Concerning the description of the 1 connected device's API methods, even though we had the syntactic description of it, we did not know its semantic meaning, in contrast to the descriptions of the 12 other activity trackers API methods' descriptions. After performing the steps of the APIs Mapping & Data Collection, the API methods that were used for collecting data from the activity trackers were recognized, including the ones that were used into the Misfit Vapor in order to gather its data. Finally, having recognized all the API methods that were responsible for gathering data from the IoMT devices, their data was acquired along with the network specifications of each IoMT device, in order to be provided to the Slicing Management mechanism for splitting the IoMT devices to the three different medical scenarios.

Slicing Management: Based on the computation needs of the three medical scenarios, the Slicing Management mechanism assigned the IoMT devices into the slices of Table 4, Table 5 and Table 6.

**Table 4. Massive IoMT scenario**

Slice 1	
IoMT Devices	Fitbit Aria, iHealth Sense, Withings Pulse O2, iHealth Track, Mi Electric Toothbrush, iHealth View, Garmin Vivofit, Mi Body Composition Scale, Fitbit Alta HR, Withings Steel HR, iHealth Core, Garmin Forerunner 630, Misfit Vapor.
Requirements	High Reliability, High Availability, High Scalability, Low End-to-end delay, High Bandwidth.

**Table 5. Enhanced Mobile Broadband scenario**

Slice 2	
<b>IoMT Devices</b>	Fitbit Flex, Withings BPM, iHealth Ease, iHealth Air, Withings Body Cardio, Garmin Vivomove HR Sport, iHealth Lite, iHealth Smart, Garmin Vivosport, Fitbit Charge 2, iHealth Wave, Withings Body+.
<b>Requirements</b>	High Capacity, Low Cache, Low End-to-end delay, High Bandwidth.

**Table 6. Mission-Critical Services scenario.**

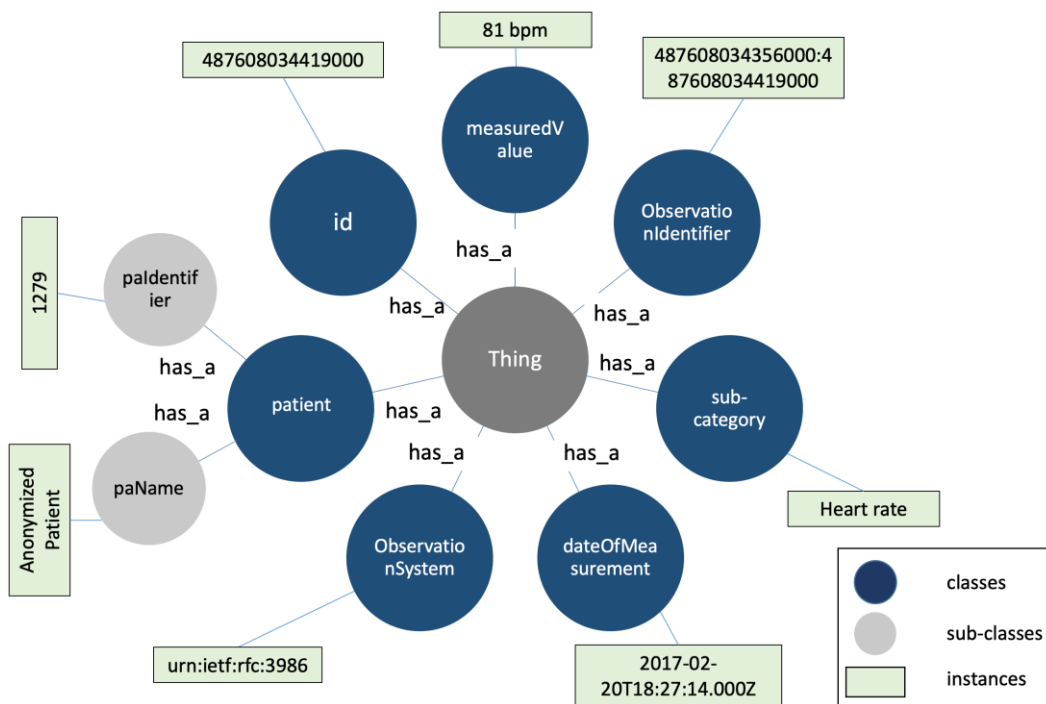
Slice 3	
<b>IoMT Devices</b>	iHealth PT3, Garmin Vivosmart 4, Withings Steel, Fitbit Zip, Withings Thermo.
<b>Requirements</b>	Low Latency, High Reliability, Low End-to-end delay, High Bandwidth.

*Data Interoperability:* For each different constructed slice, the ontological hierarchical tree of the IoMT provided dataset (Figure 5a) was created representing its concepts, relationships, and axioms (Figure 5b). Shortly, Figure 5a depicts an instance of the IoMT device dataset that represents the heart rate measurement of an anonymized patient, on the 2017-02-20T18:27:14.000Z, that was of value 81 bmp measured by the previously unknown, Misfit Vapor device (Slice 1). Furthermore, the HL7 FHIR Resources were also constructed as ontologies, taking as an input the XML representation of each one of the different HL7 FHIR Resources.

```

31 "observation": [
32   {
33     "ObservationIdentifier": "487608034356000:487608034419000",
34     "patient": {
35       "name": "Anonymized Patient",
36       "identifier": 1279
37     },
38     "id": 487608034419000,
39     "dateOfMeasurement": "2017-02-20T18:27:14.000Z",
40     "ObservationSystem": "urn:ietf:rfc:3986",
41     "measuredValue": "81 bpm",
42     "sub-category": "Heart rate",
43   },

```



**Figure 5. (a) Instance and (b) Ontological Hierarchical Tree of the IoMT device dataset**

Afterwards, the Syntactic Similarity and the Semantic Similarity Identifier were performed for each IoMT device's data in each slice, providing the HL7 FHIR Resources syntactic and semantic similarity degrees accordingly. A snapshot of the Syntactic and Semantic Similarity Identifier higher similarity degrees is depicted in Table 7, with regards to the instance of Figure 5.

Finally, the Overall Ontology Mapper dealt with the identification of the largest calculated pairs of means of both the Syntactic and the Semantic Similarity Identifiers' results, providing finally the attributes of the IoMT device dataset with higher probabilities of resemblance with the identified HL7 FHIR Resources attributes. It should be noted that this mechanism iterated for all the different slices and IoMT devices' data, transforming them into a common format, being able to be provided to different stakeholders for further analysis or usage.

**Table 7. HL7 FHIR Resources with higher Syntactic and Semantic Similarity Degrees with the Misfit Vapor dataset**

IoMT Device Dataset Attribute	FHIR Resource
<b>ObservationIdentifier</b>	observation.id (94%   76%)
<b>patient</b>	observation.subject (79%   31%)
<b>paName</b>	observation.subject.patient.name (87%   76%)
<b>paIdentifier</b>	subject.patient.identifier (83%   72%)
<b>id</b>	observation.id (100%   93%)
<b>dateOfMeasurement</b>	observation.effectiveDateTime (23%   78%)
<b>ObservationSystem</b>	observation.code (91%   26%)
<b>measuredValue</b>	observation.value (87%   81%)
<b>sub-category</b>	observation.category (69%   77%)

## 5- Evaluation

Based on the captured results, concerning the Data Acquisition mechanism, we resulted that the device type of an unknown connected device is feasible to become known. It is worth mentioning that there might be some cases where the unknown device may not have any similarity with any other known device, as in the case of the Misfit Vapor in our experimental results. Moreover, in the Data Acquisition mechanism it is clear that there might be cases that an ontology may have syntactical similarity with a specific API method, but it may not have any semantic similarity with this specific API method. Consequently, we are not able to create patterns and rules mentioning that in the case that an ontology matches syntactically or semantically with a specific ontology, it will have always an exact match with it. With regards to the Slicing Management mechanism, we can conclude that the dedicated slices were split into a correct form, since the provided slices were also calculated manually and compared with the aforementioned results. Concerning the Data Interoperability mechanism, the specific IoMT device dataset was translated manually in HL7 FHIR, in order to compare the results of the developed mechanism, with the actual outcomes. This was the main reason that a small data sample was chosen for the platform's evaluation. Through that, we resulted that the developed mechanism provided results of 100% accuracy.

To conclude, through the developed platform we were able to effectively identify a device of unknown type and collect its data, along with its network specifications, for assigning it into network slices along with other IoMT devices, and making their data interoperable through their translation into HL7 FHIR with speeds and performance which are adapted to the severity and the computational needs of the medical scenario that each IoMT device has been assigned to. This approach is extremely innovative, considering that all the researches that have been proposed in the literature so far, lack of sufficient flexibility and adaptability to heterogeneous incoming devices. What is more, in the context of healthcare interoperability, multiple attempts have also been provided to standardize and cover most of the healthcare field, which still result into clinical misinterpretations. However, through the provided platform, the known and unknown heterogeneous IoMT devices' parallel connection and data collection is becoming a reality, towards their translation into a globally recognized medical standard, with speeds and performance which are adapted to the severity of the medical scenario that each IoMT device has been assigned to.

Our future work includes the evaluation of the provided platform by testing it with additional heterogeneous IoMT devices of different types, as well as with datasets of different sizes, standards and formats, respecting privacy and

human-computer interaction issues [33]. Apart from this, we aim to extend our mechanism by supporting a general interface in which all the devices will be able to be connected regardless of their communication protocol. Finally, we aim to implement a visualization module providing abstractions over the available devices through a Web interface that will enable users to observe and manage their connected devices, to monitor the state and the location of their devices, and to visualize the acquired data, the network slices and the overall data transformation process.

## 6- Funding

The research work was supported by the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT), under the HFRI PhD Fellowship grants (2468, and 1792). Moreover, part of this work has been partially supported by the 5GTANGO project, funded by the European Commission under Grant number H2020ICT-2016-2 761493 through the Horizon 2020 and 5G-PPP programs (<http://5gtango.eu>).

## 7- Conflict of Interest

The authors declare no conflict of interest.

## 8- References

- [1] Population health outcomes, Available from: <http://www.healthcatalyst.com/population-health-outcomes-3-keys-to-drive-improvement>.
- [2] Healthcare 4.0, Available from: <http://www.vph-institute.org/news/healthcare-4-0-a-new-way-of-life.html>.
- [3] Why the IoMT will start to transform healthcare, Available from: <https://www.forbes.com/sites/bernardmarr/2018/01/25/why-the-internet-of-medical-things-iomt-will-start-to-transform-healthcare-in-2018/#40cd2e804a3c>.
- [4] Chen, Min, Jun Yang, Yixue Hao, Shiwen Mao, and Kai Hwang. "A 5G Cognitive System for Healthcare." *Big Data and Cognitive Computing* 1, no. 1 (March 30, 2017): 2. doi:10.3390/bdcc1010002.
- [5] How will 5G impact different industries?, Available from: <http://prescouter.com/2018/01/5g-impact-different-industries>
- [6] Primary data heterogeneity, Available from: <https://healthcare.ai/primary-data-heterogeneity/>.
- [7] Healthy Growth Forecast for Medical Devices Global Market, Available from: <https://globenewswire.com/news-release/2018/03/19/1441876/0/en/Healthy-Growth-Forecast-for-Medical-Devices-Global-Market.html>
- [8] Prasad, Ramjee. "5G Outlook- Innovations and Applications." *5G Outlook- Innovations and Applications* (2016): 1–260. doi:10.13052/rp-9788793379787.
- [9] Gartner says 8.4 billion connected things, Available from: <https://www.gartner.com/newsroom/id/3598917>.
- [10] Global M2M market, Available from: <https://machinaresearch.com/news/global-m2m-market-to-grow-to-27-billion-devices-generating-usd16-trillion-revenue-in-2024/>.
- [11] Cisco, Cisco visual networking index: Global mobile data traffic forecast update, White Paper (2015).
- [12] Health data exchange could save Medicare, Available from: <https://healthitanalytics.com/news/health-data-exchange-could-save-medicare-3.12-billion-a-year>.
- [13] Growing Investments to overcome Healthcare Interoperability Issues, Available from: <https://www.prnewswire.com/news-releases/global-healthcare-middleware-market-to-2023-growing-investments-to-overcome-healthcare-interoperability-issues-300659900.html>.
- [14] HL7 FHIR, Available from: <https://www.hl7.org/fhir/>.
- [15] Find MAC Address Vendors Now, Available from: <https://macvendors.com/>.
- [16] OpenCV, Available from: <https://opencv.org/>.
- [17] Levenshtein Distance measure, Available from: <https://people.cs.pitt.edu/~kirk/cs1501/Pruhs/Spring2006/assignments/editdistance/Levenshtein%20Distance.htm>
- [18] Mavrogiorgou, Argyro, Kiourtis Athanasios, Touloupou Marios, and Kyriazis Dimosthenis. "Identification of Bluetooth-Enabled IoT Devices Through Syntactic Similarity Techniques". The 11th International Conference on Ubiquitous and Future Networks (under review), (2019).
- [19] Zhang, Shichao, Debo Cheng, Zhenyun Deng, Ming Zong, and Xuelian Deng. "A Novel k NN Algorithm with Data-Driven k Parameter Computation." *Pattern Recognition Letters* 109 (July 2018): 44–54. doi:10.1016/j.patrec.2017.09.036.
- [20] Mavrogiorgou, Argyro, Athanasios Kiourtis, and Dimosthenis Kyriazis. "A Comparative Study of Classification Techniques for Managing IoT Devices of Common Specifications." *Lecture Notes in Computer Science* (2017): 67–77. doi:10.1007/978-3-319-68066-8\_6.



- [21] Mavrogiorgou, Argyro, Kiourtis, Athanasios, and Kyriazis, Dimosthenis. "Identification of medical IoT devices APIs through ontology mapping techniques." 5th EAI International Conference on IoT Technologies for HealthCare (in press), (2018).
- [22] Martínez, Ricardo, Ricard Vilalta, Ramon Casellas, Raul Muñoz, Li Fei, Pengcheng Tang, and Victor López. "Network Slicing Resource Allocation and Monitoring over Multiple Clouds and Networks." Optical Fiber Communication Conference (2018). doi:10.1364/ofc.2018.tu3d.11.
- [23] "Operations Support Systems." Network Convergence (n.d.): 385–406. doi:10.1002/9780470024430.ch9.
- [24] Bolivar, Luis Tomas, Christos Tselios, Daniel Mellado Area, and George Tsolis. "On the Deployment of an Open-Source, 5G-Aware Evaluation Testbed." 2018 6th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud) (March 2018). doi:10.1109/mobilecloud.2018.00016.
- [25] Trang, Available from: <http://www.thaiopensource.com/relaxng/trang.html>
- [26] XSOM, Available from: <https://docs.microsoft.com/en-us/dotnet/standard/data/xml/xml-schema-object-model-som>
- [27] Jung, Available from: [http://jung.sourceforge.net/presentations/JUNG\\_M2K.pdf](http://jung.sourceforge.net/presentations/JUNG_M2K.pdf)
- [28] Møller, Anders, and Michael Schwartzbach. "XML Graphs in Program Analysis." Science of Computer Programming 76, no. 6 (June 2011): 492–515. doi:10.1016/j.scico.2009.11.007.
- [29] Apache JENA, Available from: <https://jena.apache.org>
- [30] RDF triplestore, Available from: [ontotext.com/knowledgehub/fundamentals/what-is-rdf-triplestore/](http://ontotext.com/knowledgehub/fundamentals/what-is-rdf-triplestore/)
- [31] Cortical.io, Available from: <http://www.cortical.io/api.html>
- [32] 5GTANGO Project: D5.1 First open-source release of the Service Platform, Available from: <https://5gtango.eu/project-outcomes.html>
- [33] Kim, Gerard Jounghyun. "Human–Computer Interaction" (March 20, 2015). doi:10.1201/b18071.