5G Network programmability enabling Industry 4.0 transformation

George Makropoulos[1,2], Dimitrios Fragkos[1,3], Harilaos Koumaras[1], Nancy Alonistioti[2], Alexandros Kaloxylos[3], Vaios Koumaras[4], Theoni Dounia[4], Christos Sakkas[4], Dimitris Tsolkas[5].

[1]NCSR Demokritos, Greece, [2]Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Greece, [3]Department of Informatics and Telecommunications, University of Peloponnese, Greece, [4]Infolysis P.C., Greece, [5]Fogus Innovations & Services P.C., Greece

EXECUTIVE SUMMARY

The 5G system aims at enabling innovative services of vertical industries by utilizing network programmability to its full extension. The 3rd Generation Partnership Project (3GPP) has already established the foundations to provide third parties with 5G Core's capabilities by introducing the Common API Framework (CAPIF) and Service Enabler Layer Architecture (SEAL) on top of the services offered by the Network Exposure Function (NEF). However, a scalable, robust and secure ecosystem should be set for the verticals that exploit those capabilities. In this context, cloud native approach is becoming a key enabler for that ecosystem, taking advantage of the inherent cloud-native characteristics of the 5G Service Based Architecture (SBA). The chapter presents the capabilities that 5G provides to verticals as well as the ecosystem that is built around the exploitation of those capabilities. As a matter of better justification and exemplification, Industry 4.0 vertical is targeted while developments related to Vertical Application Enablers (VAEs) for the factory of the future are provided.

Keywords: 5G, programmability, Application Programmable Interface, Industry 4.0, Factory of Future, Vertical Application Enabler.

INTRODUCTION

We’re at the dawn of the next industrial revolution, commonly known as Industry 4.0, which will deliver greater operational efficiencies and flexibility at lower costs. The transition towards Industry 4.0 will offer advances in every aspect, from remote monitoring to advanced analytics and maintenance. However, the key factor for Industry 4.0 is connectivity, so manufacturers are able to use data to gain insight about their assets, be informed and make decisions on how to optimise their processes.

The intense research work on 5G experimentation globally (Díaz-Zayas et al., 2020), has reached the point where the 5G capabilities and evolvements, are appealing to be the ideal enablers aiming to shape a new and dynamic ecosystem in mobile networks from both the technology and marketing perspectives (Kostakis et al., 2021). 5G networks are envisioned to achieve a wider variety of objectives in terms of higher multi Gbps data speeds, ultra-low latency, advanced reliability, increased network capacity and availability, as well as greater bandwidth and throughput (Koumaras et al., 2021). These characteristics are essential for effectively leveraging the various services that are taking place across the entire lifecycle of the operations and processes within the verticals related to Industry 4.0. Among these unique characteristics another crucial functionality that 5G networks provide and offers high business potential, is the network exposure, which can in turn enable new levels of programmability within core networks. The programmability provided by 5G networks will unveil a wide list of network capabilities and services to third-party developers allowing them to enhance existing use cases or even create new ones.

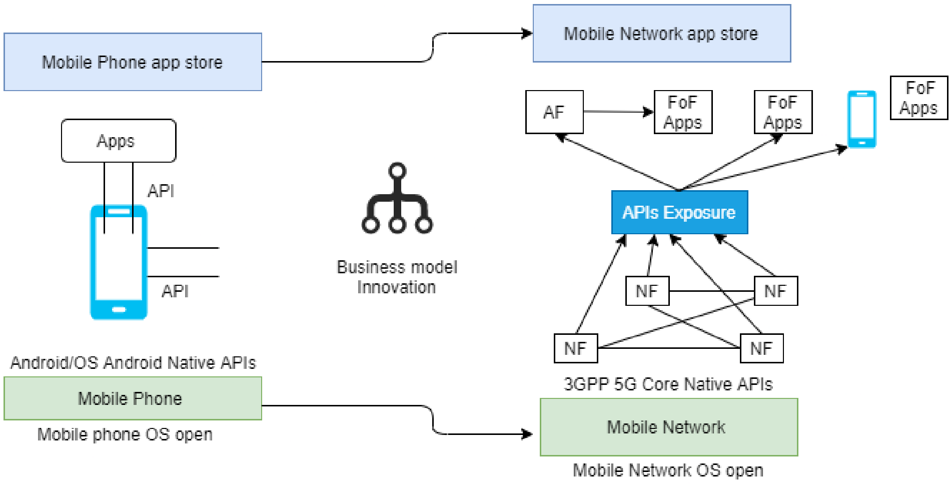
In the light of the above, the goal of this chapter is to advocate the optimal exploitation of the 5G technology towards the enhancement of the Industry 4.0 applications, and to present the new network architecture including the “open” core capabilities, as the new frontier for business innovation in industrial applications. Towards this direction, the first 3 sections of the chapter describe the capabilities that the 5G system will provide to the verticals along with the key concepts and services that are related to the 5G openness, whereas section 4 introduces the Vertical Application Enabler concept related to the deployment of vertical applications. Section 5 introduces the realisation of the Non-Public Network (NPN) Infrastructure in the Industry 4.0 and section 6 presents a Factory of the Future (FoF) use case in order to showcase the utilisation of 5G programmability and network exposure.

1. 5G CORE NETWORK CAPABILITIES TO VERTICAL INDUSTRIES

The enormous growth in connectivity, the high volume of traffic data and the broad range of business models nowadays, impose the need to move towards highly flexible infrastructures that are characterized by consistency in terms of performance and Quality of Service (QoS) provision. Moreover, given the fact that 5G networks will be the infrastructure leveraging a variety of verticals, the set of services per vertical industry, is mandatory to be able to meet a broad range of requirements. These requirements are related to the provisioning of enhanced capabilities in terms of programmability as well as efficient management of infrastructure resources (5GPPP, 2016).

In 2007, the telecom industry was upended by the launch of a new smartphone Operating System (OS) platform that disrupted the market by its unique openness features of programmability via the offered Application Programmable Interfaces (APIs). Exploiting the opportunity for innovation, third-party developers embraced the open Mobile phone OS and provided easy-to-use, programmable Software Development Kits (SDKs) and APIs to develop new apps and novel services (Ericsson, 2019). Due to the plurality of the developed apps, the App Store (Marketplace) launched very soon and the mobile app ecosystem provided a platform for programmers to reach potentially millions of customers thus becoming a successful business. Building on this paradigm, as depicted in Figure 1, the concept of network programmability through 5GCore creates opportunities that open up new avenues for growth and innovation beyond simply accelerating connectivity and data transfer. In particular, it offers even more disruption in applications' programmability, by combining the untapped capacity of multiple simultaneous network features and promising a new generation of applications that deliver an unprecedented user experience via the openness of the network. The business potential that 5G openness provides is high, taking into account that exposing the OS of a mobile phone to external developers impacted the mobile market, then the potential by opening up a whole mobile network is enormous and is expected to disrupt Industry 4.0.

Figure 1. Mobile OS opened up vs Mobile network opened up (Ericsson, 2019)

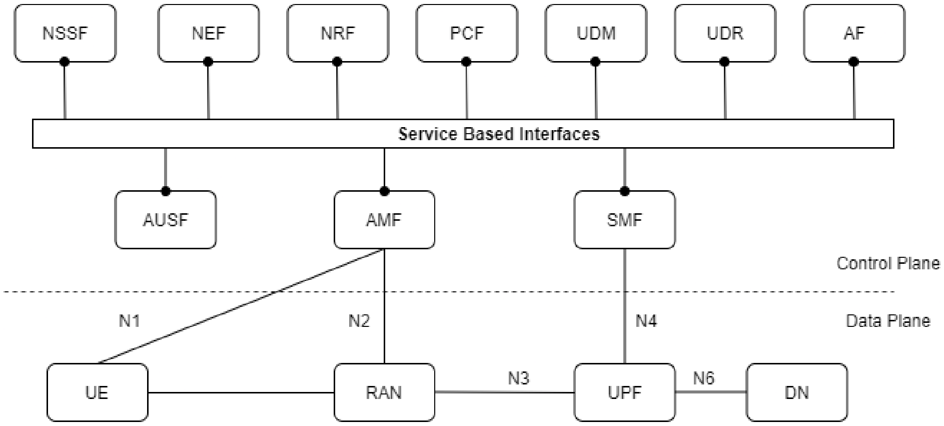


1.1. NEF and API Exposure

The above-mentioned capabilities related to the programmability of the network, are materialised through the Service Based Architecture (SBA), adopted by the 5GC network, as illustrated in Figure 2. This new shifted approach towards the SBA, allows for the elimination of resource inefficiency and performance degradation associated with virtual machines and hypervisors, thereby improving the network in terms of flexibility, speed, and automation. The 5GC control plane Network Functions (NFs) communicate through API-calls that define the related Service Based Interfaces (SBIs). In this context, the Network Repository Function (NRF) allows other NFs to register their services, which can then be discovered by other NFs.

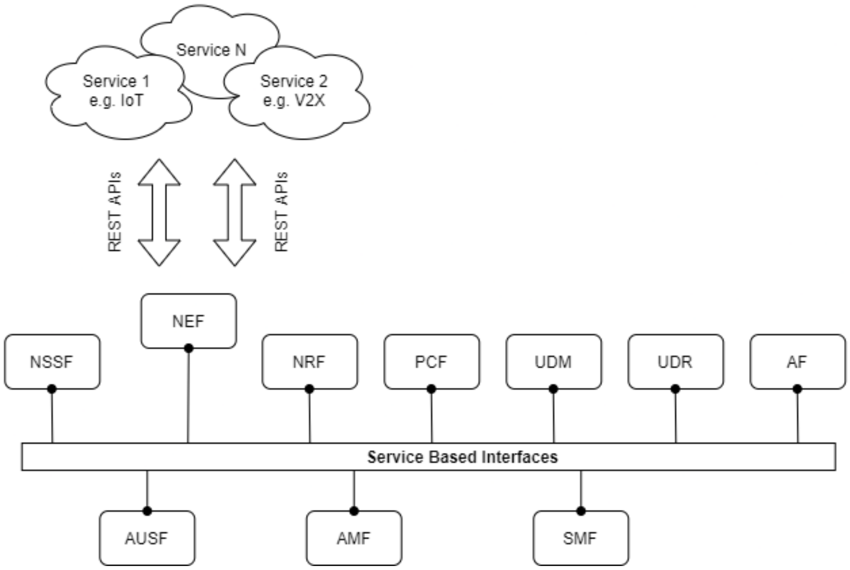
This allows for a versatile implementation, in which each NF allows other approved NFs to access the available resources of the infrastructure.

Figure 2. 5G System Architecture



In addition, the Network Exposure Function (NEF), provides a set of northbound APIs for exposing network data and receiving management commands. More precisely, NEF provides adaptors for connecting the southbound interfaces with the SBA to an exposure layer with northbound interfaces offered to third-party developers (Fragkos et al., 2021). The overall approach is illustrated in Figure 3. In this way, NEF facilitates the secure disclosure of network resources to 3rd parties, such as network slicing, edge computing, and machine learning utilizing the 5G system, fully compliant with the innovative paradigms that underpin a wide range of services (Tsolkas & Koumaras, 2022). The overall approach follows the concept of service producer-service consumer paradigm which is already established for cloud native services and now is adopted by the cellular network infrastructure.

Figure 3. RESTful APIs for the Service Based Interfaces and Northbound communication

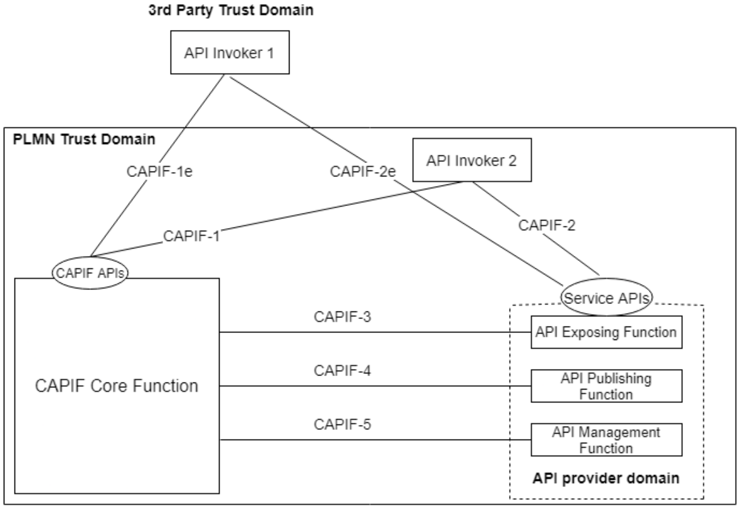


3GPP has already established the foundations, and progressively performing work in order to provide 5GC Network capabilities to vertical industries. The key concepts that have emerged are the Common API Framework (CAPIF) and the Service Enabler Architecture Layer (SEAL) together with NEF. The following sections provide a brief synopsis of these two core concepts.

2. CAPIF ARCHITECTURE

CAPIF was introduced in 3GPP Rel. 15 (3GPP, 2021a), to enable a unified approach between 5GC's northbound APIs framework and vertical apps. The key concept is the standardization and development of the common supporting capabilities (e.g., authentication, service discovery, charging policies) that are applicable to northbound APIs in order to facilitate the development of vertical apps. CAPIF consists of the CAPIF Core Function (CCF), API Invokers and API provider domain which comprises API Exposing Function (AEF), API Publishing Function (APF) and API Management Function (AMF) as described by Tangudu et al., (2020).

Figure 4. Simplified CAPIF Architecture (Source 3GPP)



The architectural model adapted from (3GPP, 2021a) is presented in Figure 4 and the functional entities are briefly described as follows:

- CCF, acts as an orchestrator that manages the interaction between service consumers (vertical apps) and service providers (e.g., NEF, SEAL). The main responsibilities of CCF are authentication of the API invoker, authorization of the API invoker to access the available service APIs, monitoring the service API invocations.
- API Invoker, represents the vertical app which consumes the service APIs utilizing CAPIF. API Invoker provides to the CCF the required information for authentication, discovers and then invokes the available service APIs.
- AEF, is responsible for the exposure of the service APIs. Assuming that API Invokers are authorized by the CCF, AEF validates the authorization and subsequently provides the direct communication entry points to the service APIs. AEF may also authorize API invokers and record the invocations in log files.
- APF, is responsible for the publication of the service APIs to CCF in order to enable the discovery capability to the API Invokers.

- AMF supplies the API provider domain with administrative capabilities. Some of these capabilities include auditing the service API invocation logs received from the CCF, on-boarding/off-boarding new API invokers and monitoring the status of the service APIs.

3GPP considers two main architectural deployment models, centralized, when the CCF and API Provider domain functions are co-located, and distributed, when CCF and API Provider domain functions are not co-located and they are interacting through CAPIF-3/4/5 interfaces. Therefore, multiple CCFs can be deployed in the same PLMN trust domain (3GPP, 2021a). CAPIF is located within the PLMN operator network. Thus, there are two functional options for API Invokers; usually 3rd party applications, which have service agreement with PLMN operator, represent API invokers (i.e., API Invoker 1) but they may be co-located within the same PLMN trust domain (i.e., API Invoker 2). Whether third parties have business relationships with PLMN, they can provide their own service APIs to CCF through CAPIF-3e/4e/5e interfaces, but they need to act in accordance with the functionalities of the API provider domain. In order to be compliant with the overall architecture NEF and SEAL (i.e., SEAL server) supports the CAPIF's API provider domain capabilities, as specified in 3GPP (2021b, 2021c).

2.1. CAPIF Services

The available CAPIF services and their respective APIs according to 3GPP (2021a) are listed hereby. Services are divided into four categories, common, security, management and internal connectivity services:

2.1.1 Common Services

Common services are fundamental to bridge the communication between services consumers (i.e., vertical app) and service providers (i.e., NEF, VAE, SEAL etc.). In the first place, service providers have to publish their services to the CCF. After the CCF becomes aware of the available services, a service provider initiates the procedure to discover the services. The common services are described as:

- **Discover (CAPIF_Discover_Service_API):** This service enables API Invokers to retrieve the available services that have been registered in CCF.
- **Publish (CAPIF_Publish_Service_API):** APF consumes this service to publish/unpublish a service API to the CCF. The publication includes details about the specific service API. APF can also update already published services.

  Retrieve (CAPIF_Publish_Service_API): APF requests from CCF information related with previous published services. When a publication occurs CAPIF registers all the related information in a repository (i.e., API registry).

2.1.2 Management Services

Considering that Invokers are onboarded and use the common services, management services assess the communication between service consumers and service providers, thus management services include:

- **Logging (CAPIF_Logging_API_Invocation):** Upon invocations (i.e., from API Invokers), CCF may store valuable information such as API invoker's ID, IP address, service API name etc. AEF utilizes this service to log API Invoker's service invocations onto CCF and potentially to access log files that have been stored previously in CCF.
- **Auditing (CAPIF_Auditing):** This service can be used to control CAPIF interactions with API Invokers (e.g., invocation events, onboarding events, authentication), which are stored in CCF. AMF initiates a request to fetch the respective log files.
- **Charging:** AEF can use this service to retrieve charging related information flows from the CCF.

- **Monitoring events (CAPIF_Monitoring):** Monitoring event service is used by AMF in order to get notified whether an event occurs in the CCF. Some of the events are the availability of service APIs (e.g., active, inactive), changes in service APIs (e.g., after an update), service API invocations, API invoker status (e.g., onboarded, offboarded) and performance related events (e.g., load conditions).

2.1.3 Internal Connectivity

The services described below accommodate the various deployment scenarios of the CAPIF Architecture. As mentioned, CAPIF supports distributed deployments where multiple CCF can be realized. Moreover, Invokers can be part of the PLMN trust domain or third-party trust domains and they may access the available services through different CCFs. The internal connectivity services that fulfill the described aspects of the CAPIF are listed below:

- **CCF interconnection (CAPIF_Discover_Service_API / CAPIF_Publish_Service_API):** This service enables the interconnection between multiple CAPIF providers. Each CAPIF provider has a CCF which utilizes publish and discover services in order to interchange its APIs.

- **Topology hiding (CAPIF_Routing_Info):** This service enables hiding the topology in the functional scenario where CAPIF includes PLMN trust domains, third party domains and API invokers access the service APIs from outside both the PLMN and third-party trust domains. In this case, API invokers access an AEF which acts as an entry point. Thus, the information for the entry AEF is shared with API Invoker in the discovery service. Then, subsequently, AEF resolves the actual destination address of the requested service API and forwards the initial request.

The aforementioned services need to fulfill the authentication and authorization prerequisites. The capabilities of the services are presented under the assumption that API provider domain functions (i.e., AEF, APF, AMF) and API Invokers are already authorized by the CCF and they are active. The details on the security aspects are out of scope. However, more information can be found in 3GPP (2020).

3. SEAL ARCHITECTURE

SEAL was introduced in Rel. 16 to support easier and faster development and deployment of vertical apps (3GPP, 2019). While the demand to develop vertical app standards for different types of industries was continuously increasing, it became obvious that many auxiliary services, such as location management, are needed across multiple vertical apps. As a result, capturing these commonly used auxiliary services and offering them to verticals as a common service layer, will benefit both verticals, allowing them to focus only on the core features and functionality of the vertical app, and operators, saving them from enormous efforts and time to develop the corresponding services for each vertical. The afore described concept became reality with the standardization of SEAL architecture (3GPP, 2021c). SEAL architecture enables these common services to be consumed by vertical apps over 3GPP, CAPIF compliant, northbound APIs. SEAL architecture supports two functional models: on-network (i.e., SEAL-Uu), when the UE connects to the 3GPP network system to consume the service, and off-network (i.e., SEAL-PC5), when UEs connect to each other directly. The functional architecture is depicted in Figure 4. For simplification, we consider only the on-network model.

The main functional entities of SEAL architecture are the following:

- **Vertical Application Layer Client (VAL client):** This entity provides the client-side functionalities of the corresponding vertical app (e.g., Vehicle to Everything (V2X) client).

- **Vertical Application Layer Server (VAL server):** This entity provides the server-side functionalities of the corresponding vertical app (e.g., V2X application server). If CAPIF is supported, VAL server acts as an AEF to provide the service APIs to the Vertical Application Server (VAS) or another VAE server. It can also act like an API Invoker to consume the service APIs, whether they are provided by another VAL server.

- **SEAL Client:** This entity provides the client-side functionalities corresponding to a specific SEAL service (e.g., Location Management client)
- **SEAL Server:** This entity provides the server-side functionalities corresponding to a specific SEAL service (e.g., Location Management server). It can act as CAPIF's API exposing function.

Various deployment scenarios have been proposed in SEAL architecture, concerning the domain in which SEAL servers are deployed. The SEAL servers can be deployed: a) in a single PLMN operator domain (centralized deployment), b) in multiple PLMN operator domains, as distributed function, with or without interconnection between the SEAL servers, c) in the VAL service provider domain or d) in a separate SEAL provider domain (3GPP, 2021c).

3.1. SEAL Services

The following section describes the common set of SEAL services designed to be used by vertical apps:

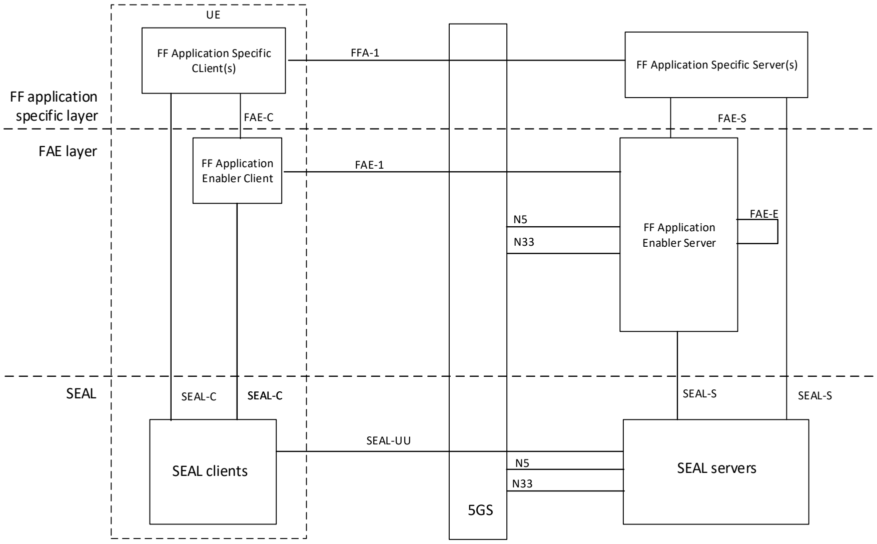
- **Location Management:** Enables the vertical app to have access to network location information of its corresponding UEs. More specifically, this service can send reports on-demand to a VAS about the location of its UEs, subscribe the VAS so as to receive notification when location information of UEs changes, share UE location information etc.
- **Group Management:** Allows vertical apps to group UEs, thus enabling group management operations, such as enforcing group policies, edit group configurations etc. The service also allows the vertical app to subscribe for and receive notifications when group information or status is modified.
- **Configuration Management:** Enables the vertical app to create and manage configuration on its UEs (provide initial configuration, edit configuration, notify server when configuration changes etc.)
- **Identity Management:** This service is responsible for the authentication and authorization procedures of a vertical app user.
- **Key Management:** Enables a vertical app to support secure transfer of data by providing and storing encryption keys.
- **Network Resource Management:** Allows a vertical app to manage network resources by managing (create, modify, delete) unicast and/or multicast bearers.

4. VERTICAL APPLICATION ENABLERS

Within this frame of reference, 3GPP has introduced the concept of Vertical Application Enablers (VAEs) in Rel. 16 (3GPP, 2021), enabling the efficient use and deployment of vertical apps over 3GPP systems (3GPP, 2019). The specifications and the architecture are based on the notion of the VAE layer that interfaces with one or more Vertical applications. VAEs communicate via network-based interfaces that are well defined and version-controlled. The main focus of VAEs is to provide key capabilities, such as message distribution, service continuity, application resource management, dynamic group management and vertical application server APIs utilising the 5G system (5GPPP, 2021).

VAE layer acts as a support layer between SEAL layer and a specific vertical application layer (e.g., V2X application client and server). This intermediate layer enables the deployment of the actual vertical app, by utilizing SEAL/NEF APIs and translating all the underlying network data to vertical application specific. In the case of Industry 4.0 applications the VAE layer can be described as Factory of Future Application Enabler layer (FAE). The functional model of the VAE/FAE layer is depicted in Figure 5. Similarly, to SEAL architecture, VAE supports both on-network and off-network models. Note that, both VAE Client and VAE Server are mutually-exclusive with VAL Client (SEAL) and VAL Server (SEAL), respectively.

Figure 5. FAE/VAE-SEAL Functional Model (3GPP, 2021d)


UE
FF Application Specific CLient(s)
FFA-1
FF Application Specific Server(s)
FF application specific layer
FAE-C
FAE-S
FAE layer
FF Application Enabler Client
FAE-1
N5
N33
FF Application Enabler Server
FAE-E
SEAL
SEAL-C
SEAL-C
SEAL-S
SEAL-S
SEAL-UU
SEAL clients
N5
N33
SEAL servers
5GS

The main entities that comprise the VAE architecture are the following:

- **Vertical application specific client:** Provides client-side functionalities corresponding to a specific vertical app (e.g., a platooning client in V2X use case).
- **Vertical application specific server:** Provides server-side functionalities corresponding to a specific vertical app (e.g., a platooning server in V2X use case). As mentioned, vertical apps can act as API invokers, if CAPIF is adapted. Specifically, the vertical app's server side represents the invoker (3GPP, 2021).
- **VAE client:** Provides the client-side support functions for a specific vertical app. Some of the functions include delivering application messages to vertical app clients, receiving monitoring reports from VAE server, and providing location information to VAE server.
- **VAE server:** Provides the server-side support functions for a specific vertical app (e.g., communicate with the underlying network, provide service discovery, support resource adaptation etc.).

VAE servers can be deployed either in a centralized manner, in which one VAE server supports one or more vertical app specific servers, or in a distributed manner, in which one or more VAE servers (with or without interconnection between them) support one vertical application specific server. Furthermore, the VAE server can be deployed either in a PLMN operator domain or in a vertical service provider domain. As mentioned, the VAE layer utilizes the capabilities of the underlying SEAL, thus it provides additional vertical specific capabilities to enable the applications. 3GPP has already specified the VAE architecture for Vehicle to Everything (V2X) services and Unmanned Aerial Systems (UAS). Work and studies are ongoing also for the FoF services (3GPP, 2021d), which is the sector that the specific chapter is focusing on. The above-mentioned frameworks comprise 3GPP's SA6 initiatives occupying a vital role between the

5GS and the vertical service providers. With CAPIF enabling a unified and harmonized Northbound API framework, improvements to the existing platforms such as SEAL which serves as a common service layer, and VAE enabling the creation of the vertical specific application, can be applied. Thus, new frameworks may arise in the future to enhance and support the ever-increasing requirements of the vertical applications.

5. NPN INFRASTRUCTURE

At the onset of the Industry 4.0 era, factories and manufacturing companies around the globe are focusing on reducing the operational costs and increasing their production numbers and effectiveness, due to shorter business and product life cycles. To that end, the progress towards a new generation of more efficient and novel ICT technologies, is deemed necessary in order to introduce sustainable and connected industrial systems, allowing the FoF concept to emerge. Indeed, several leading manufacturers are already engaged in the digital transformation era that will leverage connectivity, intelligence and allow for flexible automation.

The envisioned digital transformation to be applied to manufacturing processes will provide several advantages such as advanced analytics (based on measuring device specific parameters like vibration, temperature or noise levels), situational awareness and prediction of maintenance needs. In addition, the interconnection of the several machineries extends to using the most prominent communication technology, which can enable new features, including, for instance, flexibility of the topology and load balancing that controls the quality of the communication traffic. Moreover, within the industrial environments, the cooperation of robots and humans is intensified due to massive utilization of wirelessly connected sensors within the factory. Compared to common use cases from other vertical sectors (e.g., energy, media, smart cities), the use cases that are related to Industry 4.0 will induce strict requirements in terms of latency, reliability and high-accuracy positioning.

All these challenges related to the envisioned digital transformation, rely on efficient connectivity, high throughput and low latency. Through the years different wireless technologies have been proposed to support the industrial domain. The wireless technologies that have been proposed through the last years are generally classified into short range (i.e., ZigBee, Wi-Fi), long range (i.e., LoRa, Sigfox) and cellular networks (Sikimić et al., 2020). The technologies when used for Industry 4.0 do not seem to be able to provide the necessary requirements. Both short-range and long-range networks draw upon the unlicensed spectrum, but they cannot offer the critical capabilities, latency, and Quality of Service that are needed, with the exception of Wi-Fi. The alternative option of wired connections is limited to stationary objects and is impractical when it comes to connecting the large number of devices used in Industry 4.0 factories.

In this context, the concept of 5G Non-Public Networks (NPN) has emerged in order to support communication applications with heterogeneous and demanding requirements within the Industrial sector (Ordonez-Lucena et al., 2019). According to 3GPP, next generation mobile networks can be classified into public and Non-Public (3GPP, 2020a). The former refers to Public Land Mobile Networks (PLMN) which are typically led by Mobile Network Operators (MNOs) that provide their services publicly and they primarily operate in national scope. On the other hand, a NPN enables the deployment of a 5G System (5GS) that restricts its operability to private organizations, typically an industry vertical and offers private network services to end users acting within organization's premises.

As Ordonez et al. (2019) describes, NPNs can be divided into two main categories namely Stand alone and Public Network Integrated (PNI). A stand-alone NPN is an isolated private network that has zero interaction with PLMN. Within the environment of Industry 4.0, the utilization of a NPN permits a vertical to make use of an in-premise 5G network, and as a consequence the traffic that is related to this network will be limited within the premises' boundaries, without spanning to the public domain. This aspect offers the following advantages:

- Provision of QoS utilizing 5G network functions and service applications as close as possible to the devices and making use of advanced technologies like Time Sensitive Networking.

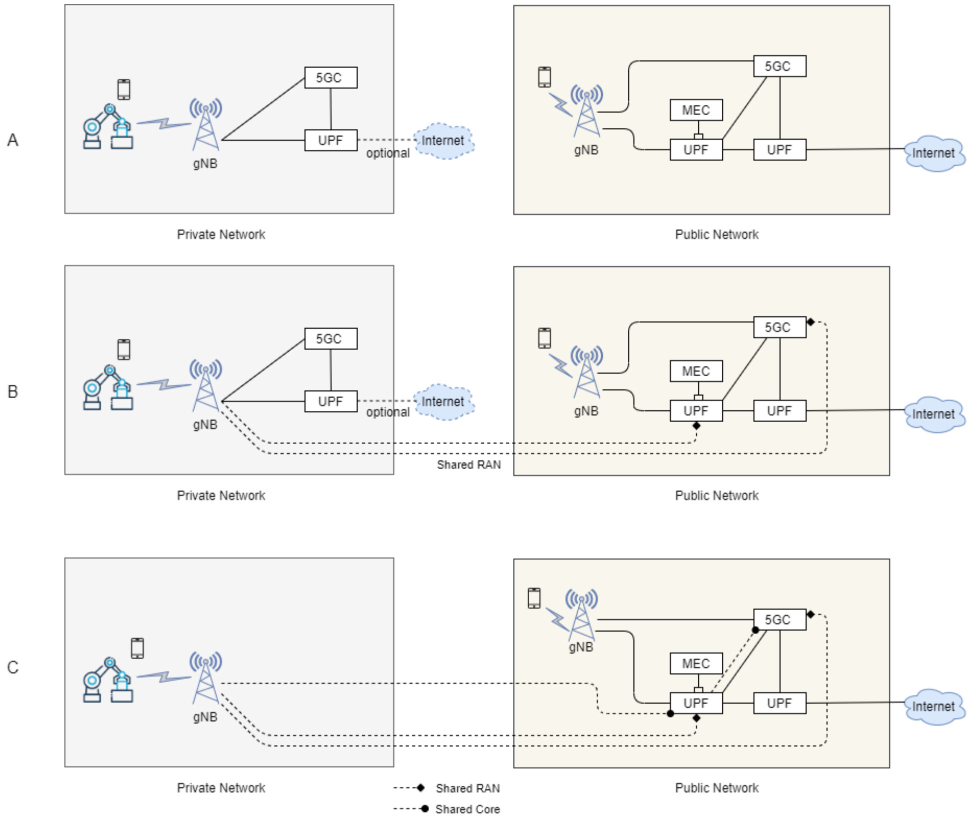
- Isolation for the public domain, which enables advanced security towards the NPN
- Efficient authorization and authentication of the devices within NPN, through the isolated network operations

On the contrary, PNI-NPNs rely on PLMN, meaning that a part or the whole of the functionality is provided by the MNO. The autonomy of a NPN in relation to a PLMN can be described based on the following characteristics:

- Use of a unique identifier for the NPN.
- Allocation of private spectrum to the NPN.
- Full deployment of a 5G system, meaning both Radio Access Network and Core Network, within the perimeter of an industrial environment.

PNI-NPN category supports a variety of deployment modes and architectural options, which in turn facilitate the heterogeneous use cases with different requirements within the Industry 4.0 ecosystem. The 3GPP specifications describe a variety of deployment modes related to 5G NPN. One approach is to set up a standalone network independently of the public network. In such case all network functionalities, including both the radio access network (RAN) and the control plane, are placed within the facility. Moreover, dedicated spectrum (licensed or unlicensed) shall be obtained from a mobile network operator for the standalone deployment case.

Figure 6. Deployments models of NPN



In a second deployment scenario, the NPN and service provider share the same radio-access network (RAN), with control plane elements and other network operations located at the NPN site, as illustrated in Figure 6B. This type of configuration allows to the traffic of the network to be routed locally within the NPN's physical premises, while other type of data is routed to the service provider's network if need be. A third major deployment is when the NPN is directly connected to a public network, as depicted in Figure 6C The traffic stemming from both the public and private network is located off-site. By using the slicing concept, the network services and the resources between the public network and the NPN are virtualised and can be kept fully distinct and isolated. In this sense, each slice created on top of the physical infrastructure stands for a complete logical network consisting of network capabilities as well as associated resources which can provide specific end-to-end enhanced service capabilities (Kourtis et al., 2020).

6. FOF USE CASE

In the era of Industry 4.0, factories and manufacturers are pressured to increase their production and effectiveness by including new technologies and equipment while at the same time focusing on interconnectivity, automation, machine learning, and real-time data. To that end and as more and more manufacturing companies incline to that direction, thus forcing the Factory of Future (FoF) concept to continue its growth, there is an increasing need for new ideas that will further support that concept and bring fundamental changes to the core of manufacturing. While industry workers are directly affected by these changes and actively encouraged to interact with machines and collaborate using digital systems on an everyday level, human-computer interaction (HCI) is being placed at the heart of industry 4.0. Supporting the concept, chatbots are one of the most important applications of industry 4.0, combining artificial intelligence and HCI, being the perfect candidate for supporting employees' everyday work. At the same time, the emergence of the 5G system comes to support this evolution by promising a new generation of applications able to realize the innovative ideas towards Industry 4.0.

In the light of the above, the following sections describe a use case scenario focusing on how 5GC exposure of standard APIs and the concept of Vertical Application Enablers (VAE) can be utilised so as to build innovative applications that will be based on top of various vertical apps within the Industry 4.0 ecosystem.

6.1. The Opportunity

Currently, the traditional way that maintenance processes are being addressed mainly relies on custom internal procedures, which include the reporting of an issue and the assignment of the reported issues, in second time, to the appropriate personnel. Although this process might be working so far, it is limited in terms of automation, quick response time and workplace safety. For example, not all workers can undertake dealing with any maintenance issue. Thus, their access to the respective documentation should be denied and the proper person should be notified. On the other hand, some malfunction could be in high priority and thus require as quick a response time as possible. Additionally, given that a worker starts dealing with a reported issue, he would most probably have to rely on paper documentation. This could potentially be time consuming since the worker will have to fetch the corresponding documentation and then address the issue. As industry 4.0 aims at replacing cumbersome and time-consuming paper documentation with digital alternatives such as AI-driven autonomous assistance systems, chatbots can act as easy-to-use conversational agents that will support the engineers and technicians during daily workflows by using data collected from sensors and databases. However, such assistance systems require the appropriate network infrastructure within the factory in order to work and benefit both the factory and its workers to the maximum possible level.

More specifically, for the described scenario, chatbots need to locate users to perform efficiently and provide to the worker the most relevant information for the machinery that is located in close proximity with the identified problem. This additional information will boost the performance of the chatbot and the efficiency of the assistance that is to be provided. Such requirements are currently not addressed in traditional factories but can be provided by the 5G System (5GS).

The described scenario will most definitely have many benefits among which the following are expected:

- To make the maintenance process more efficient and safer, as only the authorized worker can proceed with the maintenance and have on his phone all the documentation and user manuals that he needs to resolve the incident.
- To optimize maintenance time and effectiveness, as only authorized employees will be allowed to address the reported issue.
- To introduce new ways of security monitoring and alerting of the factory environment.
- To enrich the value chain with a new means of communication and interactions which provide accuracy of information, ease of use, minimum delay of operations and versatile accessibility, leading to new business models and opportunities.

6.2. The use case scenario

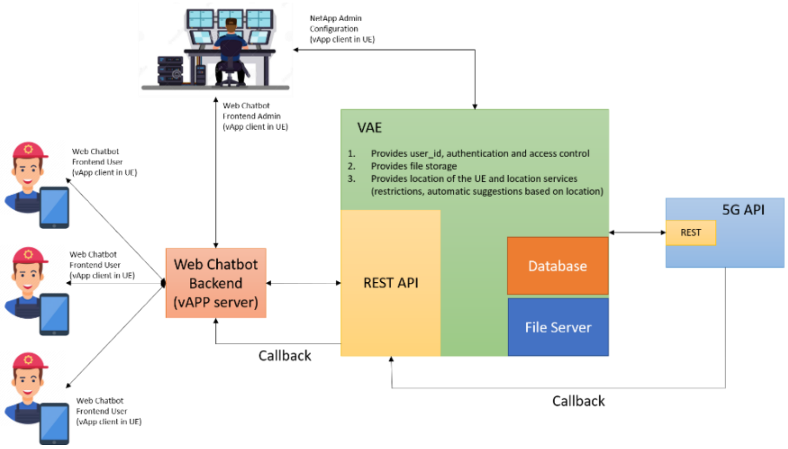
Leveraging the concept of the VAE and the APIs exposed by the 5GS, a use case scenario that can tackle the described scenario by proposing an innovative solution, using the advantages and capabilities of a chatbot, has been identified. The main idea is to establish a dedicated series of actions that will take place in a factory environment in order to realize the handling of maintenance scenarios via a chatbot platform. This use case targets the ways that maintenance scenarios are taking place within a factory. The goal is to introduce the chatbot application for reporting malfunctions and facilitating the course of actions needed to address and fix the reported issue. Additionally, a VAE will also need to be defined for the specification and realization of the use case.

Given a 5G non-public network that can be installed in a factory, a VAE will enable the use of a chatbot to help identify and solve possible malfunctions in a shorter time frame using a more user-friendly solution. Also, the 5G network can provide an Identification (ID) for all the connected workers of the factory and their relative location. It is also important to mention that the workers can be connected from any device of their choice (mobile phone or tablet device). At any point, a worker might encounter faulty equipment, and then he/she can use the chatbot app, that is installed and configured to their device, to report the issue. The overall procedure is as follows. Firstly, the VAE gets the location of the worker, for example the worker is located in Area A. After checking whether the specific area is safe, and thus not initiating an evacuation procedure, a second check takes place regarding the access status of the worker in “Area A”. In case the person should not be in the specific area he/she is prompted to leave. If a worker is cleared a request for the manuals of the area can be made and further instructions to fix the damaged machinery. Finally, if the problem is not resolved the closest specialized technician will be notified to aid. In order to control accessibility to the areas and provide easy assistance with a friendly automated way are some of the benefits of this use case.

6.3. The Vertical Application Enabler

For the realization and specifications of the use case described in the previous section, there is a need for defining a VAE that will reside between the vApp and the 5G System in order to advance the functionalities that are offered to the worker, following the architecture described in Figure 7.

Figure 7. Service Architecture



More specifically, the VAE, by using a REST API callback or making API calls to the 5G exposure interfaces, and receiving data from the internal database and files from the embedded file server, provides the following functionalities:

- Retrieves the User Equipment (UE) location via the 5GS. The VAE, based on the user ID, will request from the 5G system to retrieve the location info of the user.
- Correlates the UE-location provided by the 5GS to specific Factory areas. The VAE will have been properly configured in order to be aware of the factory areas and be able to map the retrieved 5G UE location to each of them. This functionality will allow the worker to easily choose only between the machines that are under this area while it also ensures the workers' safety since this area could be marked as dangerous.
- Prevents unauthorized personnel to perform actions that are not approved by the factory policy. Based on the factory area that has been spotted according to the UE location, the VAE performs an authorization process of the specific worker to proceed with the maintenance at the specific area. If the worker has not granted the appropriate level of authorization, then the worker in close proximity with the appropriate authorization level should be notified to fix the problem. Otherwise, clearance is given to the worker.
- Provides to authorized personnel, access to Maintenance documentation and Service manuals. The VAE, upon successful authorization and clearance of the worker to fix the issue, will provide access to its data storage in order to store and retrieve maintenance documentation and service manuals. Considering that, according to the use case scenario, storing input information and displaying documentation files, combined with the local deployment of the system, the VAE is also important to provide access to data storage.

The VAE as described above, is a standalone application that can be properly configured to connect with any compatible vertical application. Combining the VAE with a chatbot application is an effective solution that can follow the developments in Industry 4.0, taking advantage of the functionalities of the 5G System. The fact that chatbots are easy to learn and use only support that claim since it is easier to break into the traditional manufacturing processes and support the workers to their everyday tasks.

CONCLUSION

In the journey towards fully exploitable 5G infrastructures, we have reached the point where performance gains need to be made accessible to 3rd party innovators and SMEs. With the service composition provided by the new 5G Core architecture, it has become easier to create new services as well as to simplify the internal complexity of the networks This kind of accessibility can be performed through the development of a software layer that interacts with the control plane of a mobile network by consuming exposed APIs, in a standardized and trusted way. In this chapter we presented the capabilities provided by the exposure of 5G APIs via the software layer-VAE, the main services related to these APIs, as well as the specific characteristics that 5G NPN infrastructure can offer. Moreover, we described in details the realization of the aforementioned exposure of APIs through a use case scenario, namely a chatbot application that handles the maintenance activities in an Industry 4.0 environment.

ACKNOWLEDGEMENT

Part of the work in this chapter has been funded by H2020 EU-funded projects EVOLVED-5G (GA No. 101016608).

Key terms and definition

Industry 4.0 refers to the fourth industrial revolution, which is characterized by the integration of advanced technologies and digitalization into manufacturing and other industrial processes. It builds on the previous revolutions, which introduced mechanization, mass production, and automation to industry. Industry 4.0 is driven by the increasing availability of digital data, connectivity, and computing power, which enable the use of advanced technologies, such as the Internet of Things (IoT), artificial intelligence (AI), machine learning (ML), and robotics, to optimize and automate industrial processes.

Non-Public Network (NPN) Infrastructure refers to a private network infrastructure that is designed to provide secure and reliable communication services for industrial applications within Industry 4.0. NPN Infrastructure enables the deployment of vertical applications and the integration of various Industrial Internet of Things (IIoT) devices and sensors, which can be used to collect data and provide valuable insights for industrial processes. This infrastructure is isolated from public networks and provides dedicated connectivity, security, and quality of service, which is essential for the reliable and efficient operation of industrial applications.

Service-Based Architecture (SBA) of 5G is the network architecture that underpins the 5G mobile networks. It is a new and innovative approach to designing and implementing network services, where network functions are developed as independent services that can be easily combined and configured to create customized services for specific use cases. SBA of 5G provides a flexible and modular framework for deploying and managing network services that is designed to meet the requirements of diverse applications and industries, including Industry 4.0.

Network Functions (NFs) are software-based entities that provide various network-related services in a network infrastructure. These services include routing, switching, firewall, load balancing, and other network-related functions that are necessary for the proper operation of the network. NFs are typically deployed on standard computing platforms and virtualized environments to provide network services, which can be scaled up or down based on the network traffic demands.

Application Programming Interface (API) is a set of rules, protocols, and tools for building software applications. APIs define how different software components should interact with each other, allowing for the exchange of data and functionality between different applications or systems. They provide a layer of abstraction between software components, making it easier for developers to create software that can work seamlessly with other applications or services.

Network Exposure Function (NEF) is a key component of the Service Based Architecture (SBA) adopted by the 5G Core (5GC) network. It provides a set of northbound APIs for exposing network data and receiving management commands. NEF facilitates the secure disclosure of network resources to third-party developers, such as network slicing, edge computing, and machine learning utilizing the 5G system.

CAPIF is a common API framework and its functionality is considered as a cornerstone in the realization of 5G openness, since it allows secure exposure of 5G core APIs to third party domains, and also, enables third parties to define and expose their own APIs.

Vertical Application Enablers (VAEs) are a set of capabilities that enable the development and deployment of specialized applications for specific industries or use cases, such as smart cities, autonomous vehicles, and industrial automation. VAEs provide an abstraction layer between the underlying 5G network infrastructure and the application layer, allowing developers to access and utilize network resources and services in a simplified and standardized way.

Factory of the Future (FoF) is a concept that refers to the integration of advanced technologies such as artificial intelligence, the Internet of Things (IoT), robotics, and automation into manufacturing processes. The aim is to create a more efficient, agile, and responsive factory that can quickly adapt to changing customer demands and market conditions.

Vertical Application Enabler (VAE) is a software component or set of components that provides a set of functions to support the development of specific applications in a particular industry or vertical market.

REFERENCES

Díaz Zayas, A., Caso, G., Alay, Ö., Merino, P., Brunstrom, A., Tsolkas, D., & Koumaras, H. (2020). *A modular experimentation methodology for 5G deployments: The 5GENESIS approach*. Sensors, 20 (22), 6652.

Ericsson (2019). *Network Programmability, a new frontier in 5G* [White paper] https://www.ericsson.com/en/blog/2019/1/network-programmability---in-5g-an-invisible-goldmine-for-service-providers-and-industry

5GPPP (2016) 5G Vision, *The 5G Infrastructure Public Private Partnership: the next generation of communication networks and services* [Whitepaper]. https://espas.secure.europarl.europa.eu/orbis/document/5g-vision-5g-infrastructure-public-private-partnership-next-generation-communication

5GPPP (2021) Architecture Working Group, "*View on 5G Architecture*", [White paper]. https://5gppp.eu/wp-content/uploads/2020/02/5G-PPP-5G-Architecture-White-Paper_final.pdf

Fragkos, D., Makropoulos, G., Sarantos, P., Koumaras, H., Charismiadis, A. S., & Tsolkas, D. (2021, September). *5G Vertical Application Enablers Implementation Challenges and Perspectives*. In 2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom) (pp. 117-122). IEEE.

Kostakis, P., Charismiadis, A. S., Tsolkas, D., & Koumaras, H. (2021, May). *An Experimentation Platform for Automated Assessment of Multimedia Services over Mobile Networks*. In IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) (pp. 1-2). IEEE.

Koumaras, H., Makropoulos, G., Batistatos, M., Kolometsos, S., Gogos, A., Xilouris, G., Sarlas, A., & Kourtis, M. A. (2021). *5G-enabled UAVs with command-and-control software component at the edge for supporting energy efficient opportunistic networks*. Energies, 14(5), 1480.

Kourtis, M. A., Anagnostopoulos, T., Kukliński, S., Wierzbicki, M., Oikonomakis, A., Xilouris, G., Chochliouros, I., Yi, N., Kostopoulos, A., Tomaszewski, L., Sarlas, T., & Koumaras, H, (2020, November). *5G network slicing enabling edge services.* In 2020 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN) (pp. 155-160). IEEE.

Ordonez-Lucena, J., Chavarria, J. F., Contreras, L. M., & Pastor, A. (2019, October). *The use of 5G Non-Public Networks to support Industry 4.0 scenarios*. In 2019 IEEE Conference on Standards for Communications and Networking (CSCN) (pp. 1-7). IEEE.

Sikimić, M., Amović, M., Vujović, V., Suknović, B., & Manjak, D. (2020, March). *An overview of wireless technologies for IoT network.* In 2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH) (pp. 1-6). IEEE.

Tangudu, N. D., Gupta, N., Shah, S. P., Pattan, B. J., & Chitturi, S. (2020, September*). Common framework for 5G northbound APIs*. In 2020 IEEE 3rd 5G World Forum (5GWF) (pp. 275-280). IEEE.

3GPP TS 23.434 (2019), *"Service Enabler Architecture Layer for Verticals (SEAL); Functional architecture and information flows".* Release 16, v1.1.0.

3GPP TS 33.122 (2020), *"Security aspects of Common API Framework (CAPIF) for 3GPP northbound APIs"*, Release 16, v16.3.0.

3GPP TR 21.916 (2020a) *"*Services and System Aspects*"*, Release 16, v0.5.0

3GPP TS 23.286 (2021), "*Application layer support for Vehicle-to-Everything (V2X) services*", Release 17, v17.1.0.

3GPP TS 23.222, (2021a) "*Common API Framework for 3GPP Northbound APIs*", Release 17, v17.4.0.

3GPP TS 23.501 (2021b), *"System architecture for the 5G System (5GS)",* Release 17, v17.0.0.

3GPP TS 23.434 (2021c), *"Service Enabler Architecture Layer for Verticals (SEAL)",* Release 17, v17.1.0.

3GPP TR 23.745 (2021d), *"Study on application layer support for Factories of the Future in 5G network",* Release 17, v17.0.0.

Tsolkas, D., & Koumaras, H. (2022). *On the development and provisioning of vertical applications in the beyond 5G era*. IEEE Networking Letters, 4(1), 43-47.

World Economic Forum (2019). *Fourth Industrial Revolution Beacons of Technology and Innovation in Manufacturing* [White paper]. https://www3.weforum.org/docs/WEF_4IR_Beacons_of_Technology_and_Innovation_in_Manufacturing_report_2019.pdf