HORIZON 2020

ICT - Information and Communication Technologies

Empower White Paper

AI/ML in testing of mobile wireless networks

Project Acronym:	**EMPOWER**
Project Full Title:	**EMpowering transatlantic PlatfOrms for advanced WirEless Research**
Grant Agreement:	**824994**
Project Duration:	**42 months (Nov. 2018 - Apr. 2022)**
Publication Date:	**21 March, 2022**
Dissemination Level:	**Public**
Contact:	https://www.advancedwireless.eu/contact/

Executive summary

Artificial intelligence and Machine Learning (AI/ML) are gaining momentum and being used in various field and vertical industries such as health, transportation, energy, finance, etc. It is hence not surprising that AI/ML are used in the testing of mobile wireless network. This white paper identifies different aspects of this usage from two different angles:

- Changes in software testing:
    - Paradigm shift in software development and operation;
    - DevOps;
    - DevSecOps (Development Security Operations)/ DevQualOps (Development Quality Operations);
    - Test automation;
    - Low code/No code testing;
    - Cloud testing;
    - Mobile device testing;
    - IoT device testing;
    - Use of AI/ML in testing.
- Emerging mobile networks.

In addition, this paper will cover other research topics related to the use of AI/ML in testing of mobile wireless networks that are currently being studied by the scientific community.

The conclusion will present a summary of topics that are considered necessary for the inclusion of AI/ML in the testing of mobile wireless networks.

Table of Contents

Introduction

Artificial intelligence and Machine Learning (AI/ML) are gaining momentum and are considered for being used both in the core and access network as well as in various vertical applications such as health, transportation, energy, finance, etc. It is hence not surprising that AI/ML is also expected to be used in the testing of mobile wireless networks. The first part of the paper emphasizes the motivations for using AI/ML in testing both reasonable and feasible. We also develop a comprehensive explanation for these motivations. Then, we present the situation and current developments in this domain. Finally, the paper suggests important working items for future developments.

1. The need for AI/ML in the testing of mobile wireless networks

The need for AI/ML in the testing of mobile wireless networks can be generally explained by its vast usability and efficiency in an increasing number of fields and industries ranging from agriculture, commerce, health, energy to banking, finance and government. Indeed, AI/ML will contribute to the alleviation of the burden on the human, reduction of faults, improvement of performance and reduction of costs. However, a more comprehensive clarification of the need for AI/ML in testing of mobile wireless networks lies on the following reasons:

- Changes in software testing;
- Use of AI/ML in emerging mobile networks.

1.1 Changes in software testing

With the softwarisation and virtualization of mobile wireless networks materialized in the 5G mobile networks, software testing has become an essential component of the testing of mobile wireless networks. Consequently, advances in software testing including the use of AI/ML will also play a considerable role in the shaping of emerging testing tools for mobile wireless networks.

1.1.1 Paradigm shift in software development and operation

Software testing was born with the arrival of software in the 1950s. Focused in early days mostly on debugging i.e., to find fault, it evolved in 1980s to testing applications in real environment with a broader scope that includes quality assurance and user's satisfaction. Software testing was then incorporated as a component in the ***Software Lifetime Development (SLD)*** and encompassed major changes in the last decade due to transformations in the software nature and role as follows:

- **Higher complexity:** Application software grows both in size and function making the testing more challenging and time consuming;
- **Integration with third parties:** Software makes use quite often of functions and capabilities of multiple third parties through APIs. Proper testing requires adequate testing environment;
- **Failures lead to significant economic, public health and safety:** Application software plays a crucial role both in public and private sectors. A failure can have serious consequences and must absolutely be avoided. Testing is then put under high pressure;
- **Higher market and customer expectations:** As application software gets more sophisticated the market and customer expectations grow higher and demands for more suitable software are urging;
- **Faster, better and cheaper deliveries:** In a fierce competition, software vendors are forced to reduce constantly the delivery time while keeping the price low;
- **Frequent updates:** In a dynamic market in which trends and demands are changing quickly, application software needs to be updated frequently. Continuous development and incremental deliveries are the solution and require continuous monitoring and testing.

One important contribution to Software Lifetime Management (SLM) is done by ***ALM (Application Lifetime Management)*** with models and tools that support both waterfall and agile development model. ALM covers the entire lifecycle from the idea conception to the development, testing, deployment, to support and ultimately retirement of systems. Indeed, testing is entirely integrated in the process. Testing was dramatically changed by the ***Agile Software Development*** [2, 3] which is an iterative, incremental and evolutionary approach that breaks down the development process into several iterations including all functions: planning, analysis, design, coding, unit testing, and acceptance testing.

1.1.2 DevOps

By 2016, **DevOps** [4] hit the mainstream and as stated in its name it results from the combination of development and operation in software engineering. DevOps proposes to automate the development, testing and deployment process through ***Continuous Integration*** and ***Continuous Delivery (CI/CD)***. As shown in Figure 1, testing is now turned to ***Continuous Testing*** and efforts have been put to achieve ***Automate testing***. Further, due to constant changes, it is necessary to introduce ***Regression Testing***, which verifies recent changes, either to the program or code, in order to not negatively impact the existing features of the software.

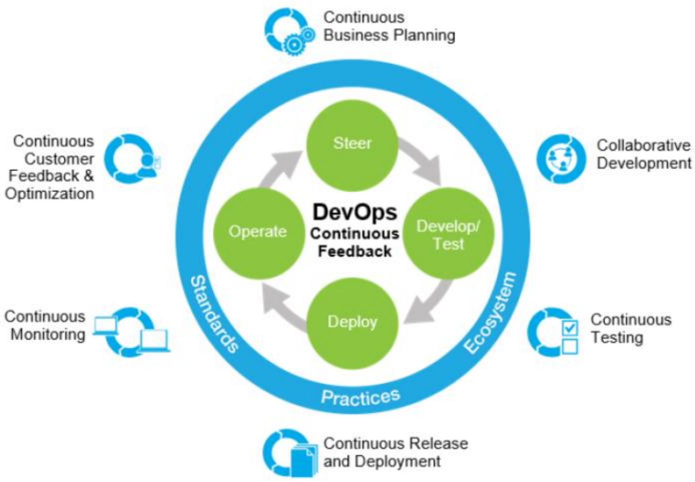

Figure 1: The DevOps Reference Architecture [4]

1.1.3 DevSecOps/DevQualOps

Recently, the COVID-19 pandemic has intensified demands on higher software quality, stronger security and improved privacy since most of people's daily tasks are carried out online on the Internet. Requirements on the integration of quality assurance (QA) and security (Sec) in early phases and throughout the development and operation process had resulted in ***DevQAOps*** or ***DevSecOps*** [6]. The security testing of applications, networks, and systems not only ensures secure transactions but also protects the privacy of end-users. Testing will have to evolve to ensure both higher quality and stronger security. Consequently, testing has become more demanding and complex.

1.1.4 Test automation

One major trend in testing is to replace manual test with automated testing because testing has become complex, time consuming and error prone with the introduction of continuous testing. The need for ***Test automation*** at least for regression test is getting urgent. Automation tools such as Selenium, Katalon and TestComplete continue to improve to provide more features and functions.

The demands for the automation of API and Service Test have increased considerably following the decoupling of client and server in the design of both Web and mobile applications. Since APIs and services are used in more than one application and any change will have consequences on all these applications, the necessity of testing becomes crucial. However, it is more effective to test these APIs and services than to test all the clients. Here AI/ML will be used to ensure the quality of test automation.

When API and services are used across client applications and components, testing them is more effective and efficient than testing the client. The trend is that the need for API and services test automation continues to increase, possibly outpacing that of the functionality used by the end-users on user interfaces.

1.1.5 Low code/No code testing

Codeless automated testing aims to reduce or completely eliminate the need of coding when creating tests. Codeless automated testing allows test teams to automate the generation of test scripts regardless of their programming skills.

Codeless testing makes use of AI/ML and visual modelling to let tester create test cases that are optimised for automation testing. Test case simulations can be generated without coding expertise and the development time for repeating test cases can be considerably reduced.

Codeless testing is effective, easy to assess and can be done by non-technical testers. There is currently a plethora of codeless testing tools in the market such as Testsigma, TOSCA, tests.ai, Ranorex, Ghost Inspector, TestComplete, Perfecto Scriptless, etc.

1.1.6 Cloud testing

With the advent of cloud computing, it is not surprising that testing is also cloudified. Indeed, ***Cloud testing*** or more precisely ***Cloud-based testing*** enables the assessment of performance, security, scalability, and reliability of applications on a quality assurance testing platform hosted on a cloud computing environment.

Cloud testing allows to carry out tests on a shared environment at any time by a geographically distributed team that may be remotely located or even mobile. Further, tester can scale their applications to support workload from a large number of users as in a real situation. Cloud testing brings reduced delivery time, flexibility and pay-per-use costs at the same time as the technologies and processes used in functional testing for cloud testing are not very different from traditional on-premises testing. The biggest challenge is, however, to ensure adequate protection for data integrity and security processes.

Cloud testing paves the way for Testbed as a Service [7] such as OneLab/FIT1, GENI 2 and Fed4Fire+3. These tesetbeds offer their testing platform to application developers and manufacturers. Companies like Cap Gemini, IBM, KPMG, etc. also offer Testing Platform-as-a-Service as a commercial service for their customers.

1.1.7 Mobile device testing

As the popularity and the penetration of mobile phones continues to grow mobile testing is getting more and more important. ***Mobile device testing*** allows the assessment of mobile applications aka. mobile apps on functionality, usability and consistency. The testing can be done manually or in an automated way, on-site or on cloud. While the mobile apps grow both in number and complexity their development time and time-to-market must be shorten. Further, mobile apps usually have a larger and more diversified user base than other applications and need to run on a broader range of devices using different operating systems such as IoS, Android, Linux, etc. This increases testing complexity and calls for mobile test automation and also the use of AI/ML.

1.1.8 IoT device testing

Although IoT or Internet of Things is not a new concept its breakthrough materialized by the boom of IoT devices has only occurred recently due to the deployment of 5G combined with the emergence of COVID 19 pandemic which accelerates the need for remote digital assistance and management. According to TechJury [5], there are around 46 billion connected devices in 2021 compared to 14,2 billion in 2019. Testing has become more complex due to several reasons. These devices belong to a variety of vertical markets such as transport, health, energy, finance, security, etc. and diverge in terms of functionality, capability and usability. Further, these devices make use of a variety of software applications, are connected to the Internet and generate large amount of data. They are also vulnerable to both physical and cyber-attacks and need adequate protection. Consequently, it is necessary to adopt an effective IoT testing strategy that takes into account the large number of testing combinations, test devices, communication protocols, operating systems and platforms. Again, AI/ML will be both useful and efficient in reducing the complexity.

1.1.9 Use of AI/ML in testing

As mentioned earlier, software testing meets considerable challenges due to the increase in number, type and size of application software and devices, higher requirements from customers in terms of quality, security and privacy, shorter time-to-market, continuous and automated testing and codeless testing. Fortunately, cloud computing and the availability of big data pave the way for applying Artificial Intelligence and Machine Learning (AI/ML) methods in the solutions addressing the mentioned challenges. Although the use of AI/ML in software testing is still in its infancy, it is already used in the following use cases:

- Generating test cases based on user behavior, which helps businesses create software tailored according to customer needs;
- Helping to detect faults, to understand test coverage, areas of high risk, etc.;
- Executing regression tests that collect data and apply performance metrics to provide accurate assessment of the software;
- Ensuring proper authoring, implementation, and continuity of automated tests;
- Running tests in simulated environments, such as mobile browsers and some specific devices.

1.2 AI used in mobile wireless networks

1.2.1 The needs for network automation

The deployment, operation and management of mobile wireless networks, in particular 5G, is becoming more challenging due to the introduction of new radio technologies such as 5G New Radio, millimetre waves, massive MIMO (Multiple Input Multiple Output), beamforming and new network technologies like MEC (Multiple Access Edge Computing), NFV (Network Function Virtualization), SDN (Software Define Network), Network Slicing combined with the dynamic network quality and performance offering tailored according to the demands of a variety of vertical markets. One of the solutions to overcome these challenges is mobile operators network automation.

Network automation, as defined by Cisco, is the process of automating the configuring, managing, testing, deploying, and operating of physical and virtual devices within a network. The network service is continuously improving thanks to daily network tasks, automated functions and repetitive processes that are automatically controlled and managed.

Advanced network automation solutions make use of AI/ML to learn about network behaviours, to deliver predictive analysis and to provide recommendations to network operations teams. They can be configured to take mitigation and recovery actions or even prevention actions. Consequently, network automation improves operational efficiency, reduces the potential for human error, increases network service availability, and delivers a better customer experience.

1.2.2 Network autonomy

A more advanced alternative to Network Automation is **Network Autonomy**. Although closely related Network Autonomy is not just Network Automation.

With Network Automation the network is operating within well-defined parameters or with predefined constraints while with Network Autonomy the network is operating independently, reflecting and adapting behaviour according to the environment beyond well-defined parameters or predefined constraints.

An **autonomous network** runs with minimal to no human intervention—able to configure, monitor, and maintain itself independently. Automation itself, and the idea that technologies could be self-provisioning, self-diagnosing, and self-healing, has been around for some time [8].

AI/ML together with virtualization and cloud computing play a key part in the realisation of first automation and then autonomy. ITU endorsed by GSMA, ETSI and TM Forum proposes 6 levels of intelligence as shown in Figure 2.

Since AI/ML methods are used for both network automation and network autonomy they must also be integrated in the testing platforms at the same time as their functions and performance should be tested and validated by sound methods.

	Level 0 Human-operated network	Level 1 Assisted-automation network	Level 2 Advanced-automation network	Level 3 Partially autonomous network	Level 4 Advanced autonomous network	Level 5 Fully autonomous network
Dimensions						
Context awareness and analysis	Human	AI/ML	AI/ML	AI/ML	AI/ML	AI/ML
Non-real-time prediction and inference	Human	Human	AI/ML	AI/ML	AI/ML	AI/ML
Decision making and execution	Human	Human	Human	AI/ML	AI/ML	AI/ML
Real-time prediction and interference	None	None	None	None	AI/ML	AI/ML
Exception handling	Human	Human	Human	Human	AI/ML	AI/ML
Features						
Human-network interface	Rule based	Rule based	Open loop	Open and closed loop	Mostly closed loop, partly intent based	Closed loop, intent based
System	Element	Element	Element, subsystem	Element, subsystem	Mostly subsystem	End-to-end

Figure 2: Six-level model for the assessment of network intelligence - https://senzafili.com/autonomous_networks/

2. Key Research topics related to the use of AI/ML for testing mobile wireless networks

2.1 Introduction of autonomic functions in mobile wireless networks

The exploitation of AI/ML in the testing of mobile wireless networks requires access to large data sets that are obtained by collecting, aggregating and storing data at multiple locations in different forms, i.e., structured and/or unstructured, in different formats and protocols for certain periods of time. Fortunately, we can certainly benefit from previous work on network autonomy focusing on architecture and functions. In the following sections, we explore the different working activities aiming at integrating AI/ML in mobile networks.

2.1.1 Architecture enhancements for 5G System (5GS) to support network data analytics services

3GPP has defined in 5G Release 16 and Release 17 *an Architecture enhancements for 5G System (5GS) to support network data analytics services* [13] offering the collection, analysis and storage of data on 5G networks.

The ***Network Data Analytics Function (NWDAF)*** has been specified further in TS 23.501 [11] to include more of the following functionalities:

- Support data collection from NFs and AFs;
- Support data collection from OAM;

* NWDAF service registration and metadata exposure to NFs and AFs;
* Support analytics information provisioning to NFs and AFs;
* Support Machine Learning (ML) model training and provisioning to NWDAFs (containing Analytics logical function).

The NWDAF fetches data from other NFs such as AMF, SMF, UDM, PCF, NRF and AF using a Service-Based-Interface **Nnf** while it exposes an **Nnwsdaf** interface that allows 5GC NF to request network analytics information from NWDAF with ***Analytics logical function (AnLF)*** [13].

This data collection service is used directly in order to retrieve behaviour data for individual UEs or groups of UEs e.g., UE reachability, and also to retrieve global UE information, e.g., number of UEs present in a geographical area.

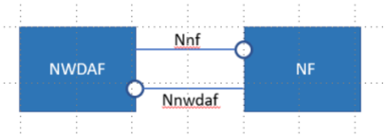

Figure 3: Data collection and Network Analytics Exposure

Single or multiple instances of NWDAF may be deployed in a 5G network. In the case of multiple NWDAF instances, all the NWDAFs may play equal role and collaborate with each other; alternatively, a central NWDAF may act as an Aggregator NWDAF that collects analytics information from other NWDAFs, which may have different Serving Areas, to produce the aggregated analytics (per Analytics ID), possibly with Analytics generated by itself.

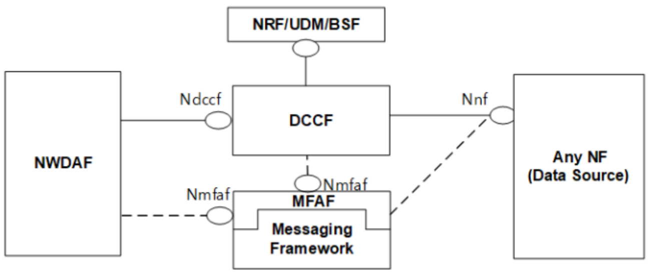

Figure 4: Data Collection architecture using Data Collection Coordination

As shown in Figure 4 the data collection can be done by a ***Data Collection Coordination Function (DCCF***) via Ndccf interface. If the data is not already being collected, the DCCF requests the data from the Data Source using Nnf services. The DCCF may collect the data and deliver it to the NWDAF or the DCCF may rely on a messaging framework to collect data from the NF and deliver it to the NWDAF.

To provide analytics to 5GC NFs and to operation and management, the NWDAF is equipped with two components [12]:

* ***Analytics logical function (AnLF):*** can perform inference, derive analytics information such as derive statistics and/or predictions based on Analytics Consumer request and expose analytics service such as Nnwdaf_AnalyticsSubscription or Nnwdaf_AnalyticsInfo;
* ***Model Training logical function (MTLF):*** trains Machine Learning (ML) models and exposes new training services e.g., providing trained model;

The storage of data is assumed by the ***Analytics Data Repository Function (ADRF)*** that enables a consumer to store and retrieve data and analytics. There are also defined procedures for analytics exposure.

Evaluation: 3GPP has specified an enhanced 5G architecture that includes the necessary collection, analysis and storage of data but has not specified how these functions should be used in achieving automation and autonomy in 5G network. Further works are required to ease the implementation and deployment of autonomic functions. Nevertheless, nothing is specified regarding neither the use of AI/ML in testing of mobile wireless networks not the testing of AI/ML-based autonomic functions.

2.1.2 NGMM 5G E2E Architecture Framework

The Next Generation Mobile Networks (NGMN) Alliance [10] [1628] proposes a ***Generic autonomic networking architectural model*** that enables a variety of flexible implementation and deployment strategies for autonomic functions as depicted in Figure 5. The Generic autonomic networking framework provides autonomic functions to all layers of every end-to-end network slices.

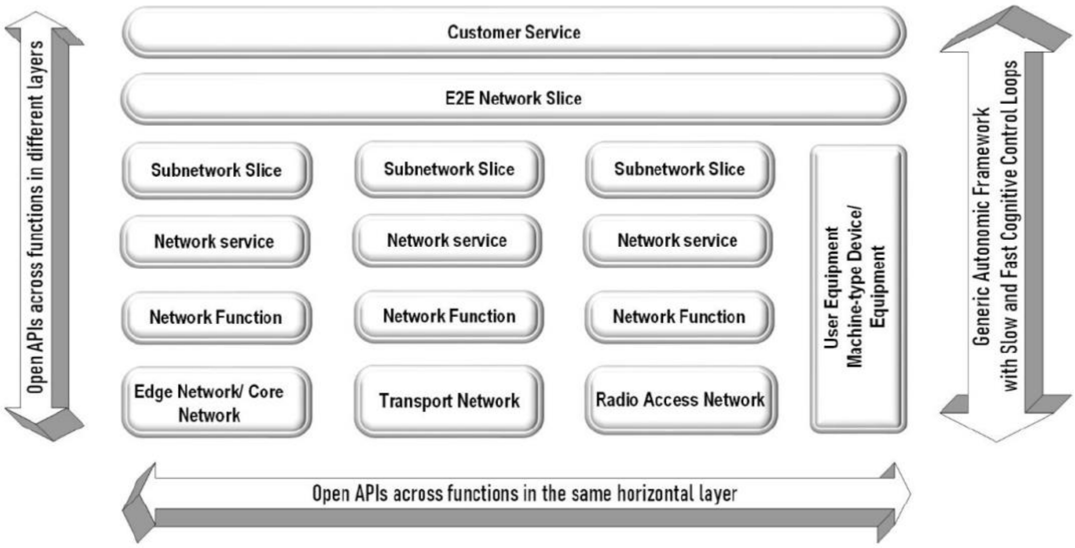

Figure 5: Generic autonomic networking architectural model

Autonomic networking referred to as ***Knowledge Plane (KP)*** includes autonomic management and control, and enables for cognitive capabilities that may be arranged in a centralized or in a distributed manner in terms of ***Decision Elements (DEs)***, which are part of a ***Network Element (NE)***, or a ***Network Function (NF)***. The ***Cognitive Modules (CMs)*** that are based on AI and ML algorithms and enable autonomic behaviours by leveraging data analytics and feedback control loops for the DEs. The CM may either be shared by one or more DEs or may be embedded within the DE. The centralized DEs, which constitutes the Knowledge Plane (KP) interact and interwork with distributed DEs for the policy control of the distributed DEs, contained in NEs and NFs.

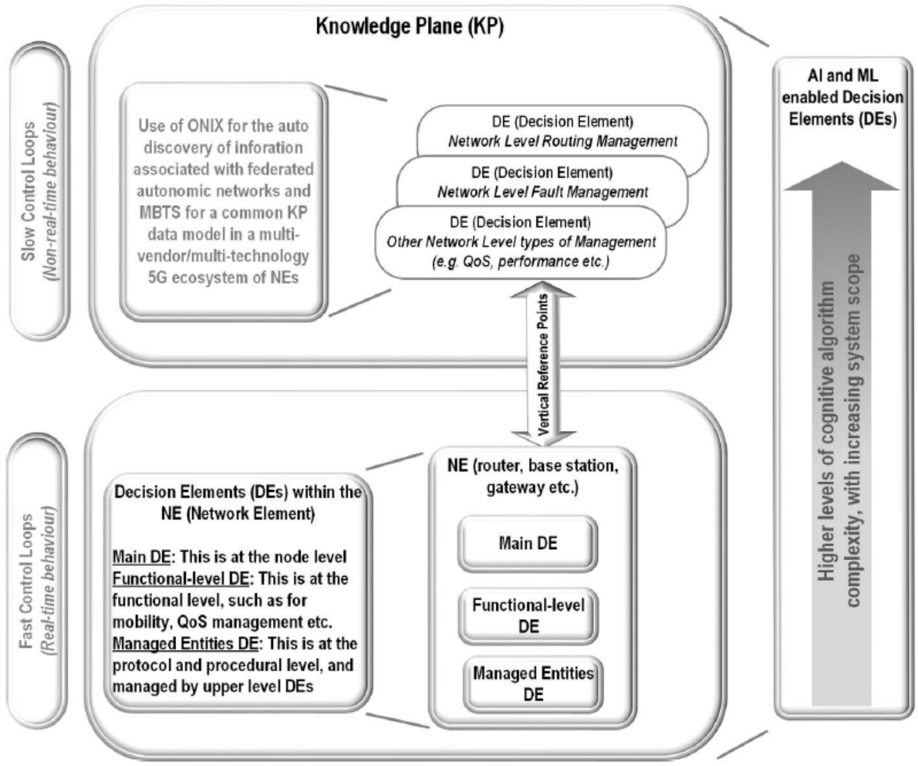

Figure 6: High-level architectural model of AI and ML enabled KP

The nature and scope of the Knowledge Plane (KP) as shown in Figure 6, requires different levels of inference and insight, which could effectively harness the different modalities of AI and ML for system-wide cognition that in turn facilitates self-organization and optimization. The KPs may be arranged to support different levels of network segmentation and distribution to suit specific deployment objectives. A federation of KPs that harness Federated Learning models may be leveraged for a cooperative operation of KPs, across collaborating domains e.g., mobile, fixed-wireless, non- terrestrial etc. types of connectivity [28].

Evaluation: The NGMM model provides a reference for the implementation of autonomic functions for mobile wireless networks. It also shows that the proposed Knowledge Plane (KP) can support all layers for all network slices. However, it is not specified how such KP can be integrated in a 5G networks since data collection and storage are not specified.

2.1.3 ITU Architectural framework for machine learning in future networks including IMT-2020

ITU defines a set of standards to enable AI/ML integration in 5G and future networks. The ITU Y.3172 Architectural framework for machine learning in future networks including IMT-2020 [28] introduces the basic toolsets in relation to the underlying network such as ML Pipeline for model optimization and serving, ML Sandbox to trial models before deployment and ML Function Orchestrator (MLFO) to control AI/ML integration as shown in Figure *7*. The standard also describes an example of realisation in IMT 2000 or 5G network as shown in Figure 8.

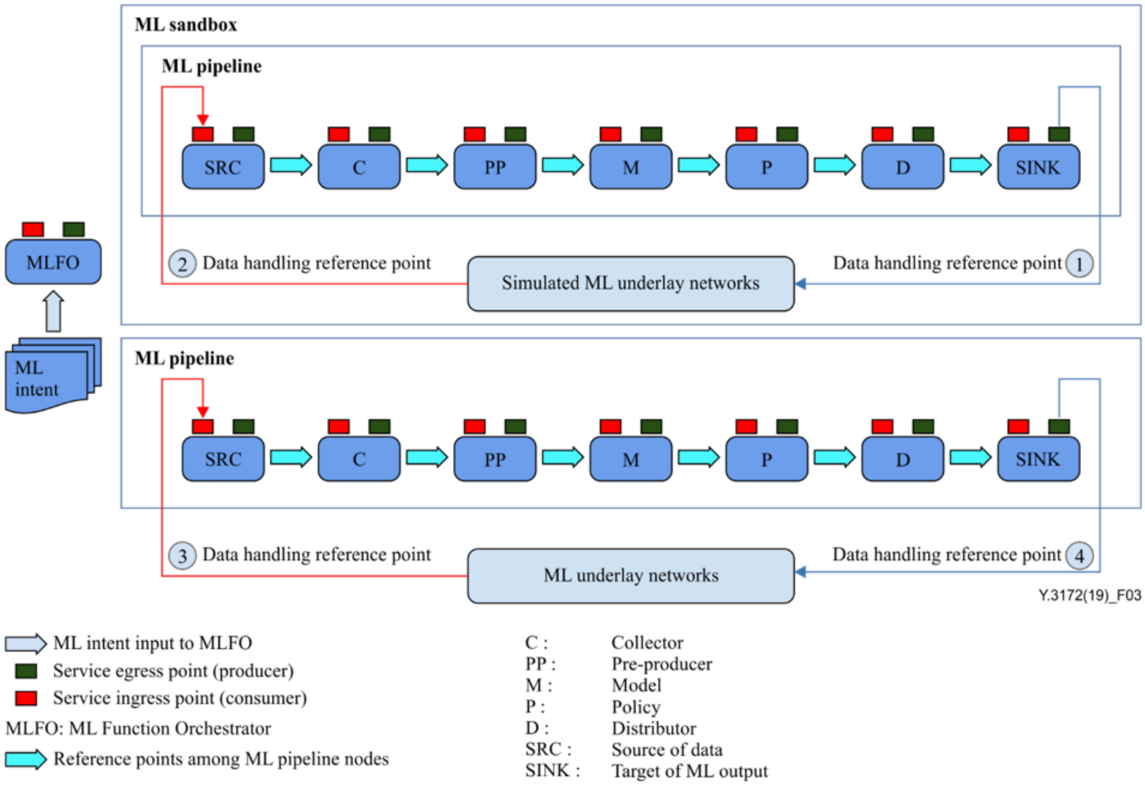

Figure 7: High-level architectural components

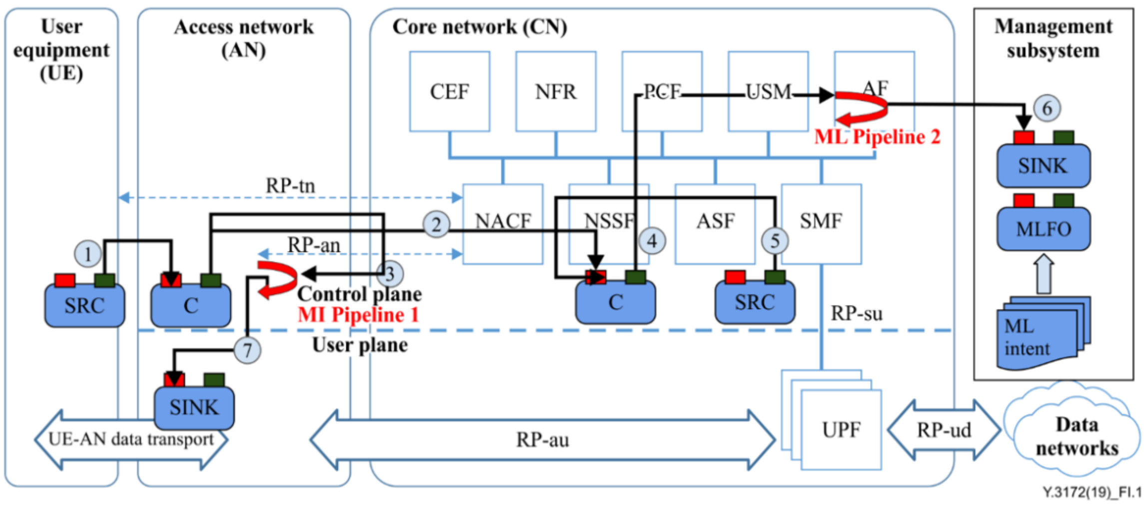

Figure 8: Example of realization of the high-level architecture in an IMT-2020 network

Evaluation: The proposed ITU framework is very generic and although one example of integration to the 5G networks is provided it is not sufficient for the implementation and integration in 5G networks due to the lack of details especially regarding the collection of data.

2.1.4 ETSI Experiential Networked Intelligence (ENI)

The Experiential Networked Intelligence Industry Specification Group (ENI ISG) defines a Cognitive Network Management architecture, called ***ENI System Architecture*** [18] using Artificial Intelligence (AI) techniques and context-aware policies to adjust offered services based on changes in user needs, environmental conditions and business goals. It therefore fully benefits the 5G networks with automated service provision, operation, and assurance, as well as optimized slice management and resource orchestration. Figure 9 shows a high-level functional architecture of ENI.

Evaluation: Although the ENI System Architecture is useful for the development of AI/ML for mobile wireless networks the practical details regarding the implementation and deployment in 5G networks are lacking.

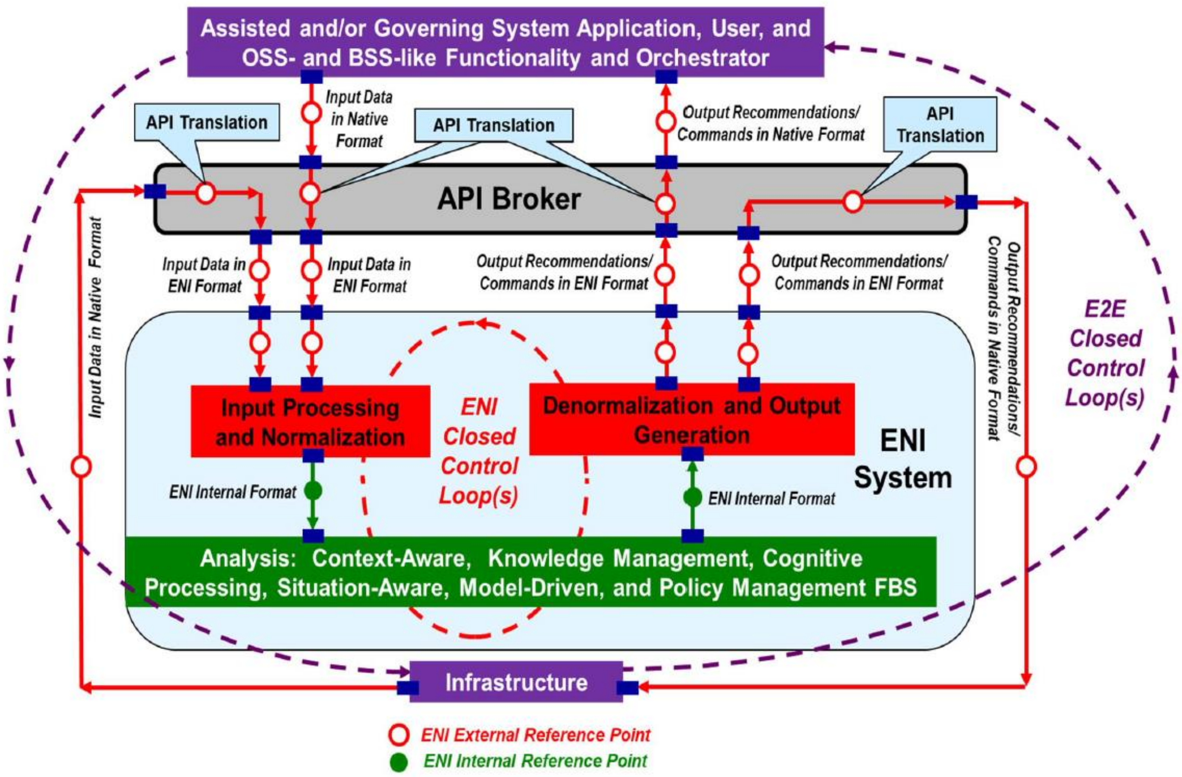

Figure 9: High-Level Functional Architecture of ENI When an API Broker Is Used

2.1.5 ETSI Generic Autonomic Network Architecture (GANA)

ETSI specifies the ***Generic Autonomic Network Architecture (GANA)*** [20][19] as an architectural reference model for autonomic networking, cognitive networking, and self-management. GANA defines so-called ***Autonomic Functions (AFs)*** as **autonomous and autonomic decision-making elements (DEs)** for network management and control that can be instrumented at four basic complementary abstraction levels for self- management within network nodes or elements/functions and in the outer management and control realm as depicted in Figure 10. The main goal of the GANA reference model is to propose the design and operational principles for Decision Elements (DEs) as the drivers for cognitive, self-managing and self-adaptive network behaviours. This should enable to achieve OPEX reduction and other benefits "Artificial Intelligence/Cognition in AMC (autonomics)" bring to Network Operators and End User Customers, and to Enterprise Networks as well, such as:

- Dynamic and Analytics-Driven Service Fulfilment and Closed-Loop (Autonomic) Service Assurance.
- Predictive, Proactive and Advanced Customer Experience.

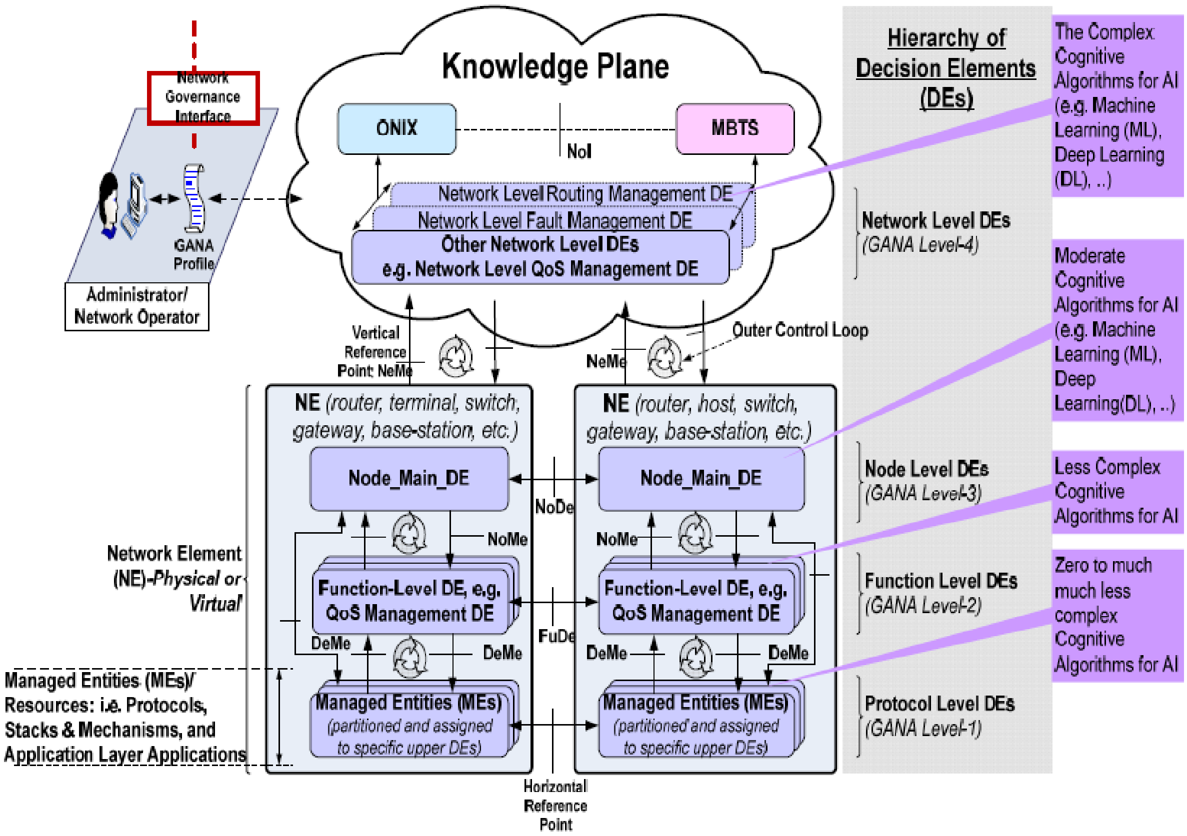

Figure 10: Snapshot of the GANA Reference Model

Decision Elements: are responsible for autonomic management and adaptive control of systems and network resources, parameters, and services. Autonomic behaviours of a DE (also called autonomic manager) include secure auto-discovery of the following items: network objectives and policies specified by the human operator, other DEs it requires to collaborate with, and capabilities of the DE's assigned Managed Entities (MEs), i.e., the information that gets available at run-time. Then after auto-discovery, a DE performs the self-* operations on its assigned MEs (by-design) by orchestrating (launching and/or configuring) the MEs when required, and adaptively (re-)programming the MEs via the effectors of their management interfaces.

Managed Entities (ME): are managed resources but not managed elements that denote a physical network element. ME can be physical or virtual. MEs have several variants. They can be fundamental MEs at the bottom of the management hierarchy (at the fundamental resources layer) such as individual protocols or stacks, OSI layer 7 or TCP/IP application layer and other types of resources or managed mechanisms hosted in a network node (NE) or in the network in general. MEs can also be composite MEs such as whole NEs themselves (i.e., MEs that embed sub-MEs).

The GANA Knowledge Plane (GANA KP): enables advanced management and control intelligence at the Element Management (EM), Network Management (NM) and Operation and Support System (OSS) levels by interworking with them or enhancing and evolving the intelligence of the systems at these levels by way of replaceable and (re)-loadable autonomics modules (DEs) that can be loaded at specific abstraction levels of management and control operations.

Evaluation: The ETSI GANA model is rather close to the NGMM Generic autonomic networking architectural model but provides higher level of details with more use cases and examples. While the standard specifies the architecture and the DEs in details, it does not recommend how many and where DEs should be deployed in 5G networks. Although this is probably done in purpose to promote creativity and innovation it poses difficulties for both mobile manufacturers and operators. Further, DEs relies on MEs at lower layer to monitor and collect data and their implementation and integration with existing 5G NFs are not sufficiently specified.

2.1.6 TM Forum Autonomous Networks framework

TM Forum proposes an ***Autonomous Networks framework*** that aims to provide fully automated ***zero wait***, ***zero touch***, ***zero trouble***, innovative network and ICT services for vertical industry users and consumers [21]. In addition, they need to support ***self-serving***, ***self-fulfilling*** and ***self-assuring*** telecom network infrastructures for internal users across various departments including planning, marketing, operations and management.

As shown in Figure 11Figure 11, the TM Forum Autonomous Networks is composed of a simplified network architecture, characterized by autonomous domains and automated intelligent business as well as network operations for the closed-loop control of digital business. This offers the best-possible user experience, full lifecycle operations automation/autonomy and maximum resource utilization.

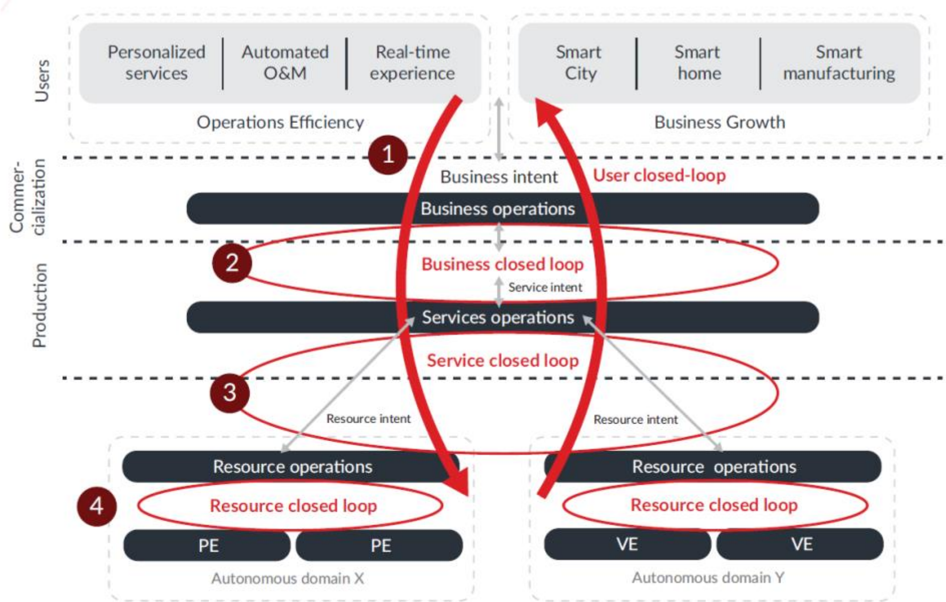

Figure 11: TM Forum Autonomous Networks framework

Evaluation: The TM Forum autonomous framework gives a clear overview architecture of autonomous network that describes the necessary functions and domains to realise automation and autonomy. Unfortunately, the framework does not provide sufficient details for implementation and deployment in mobile wireless networks such as 5G networks.

2.1.7 GSMA Layered Framework

To provide efficient operation automation, GSMA promotes a layer framework as shown in [22] that includes closed-loop automation on each layer namely network element layer, network domain layer, and cross-domain network layer. To improve performance and reduce complexity, the network domain layer must provide an open simplified interface to upper domain layers such as cross-domain layer. The data exchanged between interfaces will progressively be narrowed down. The simplification of the open interfaces relies on autonomous network capability in each domain and layers. GSMA also describes a series of use cases in the report *AI in Network - Use Cases in China* [23], which illustrates the use of the layer framework.

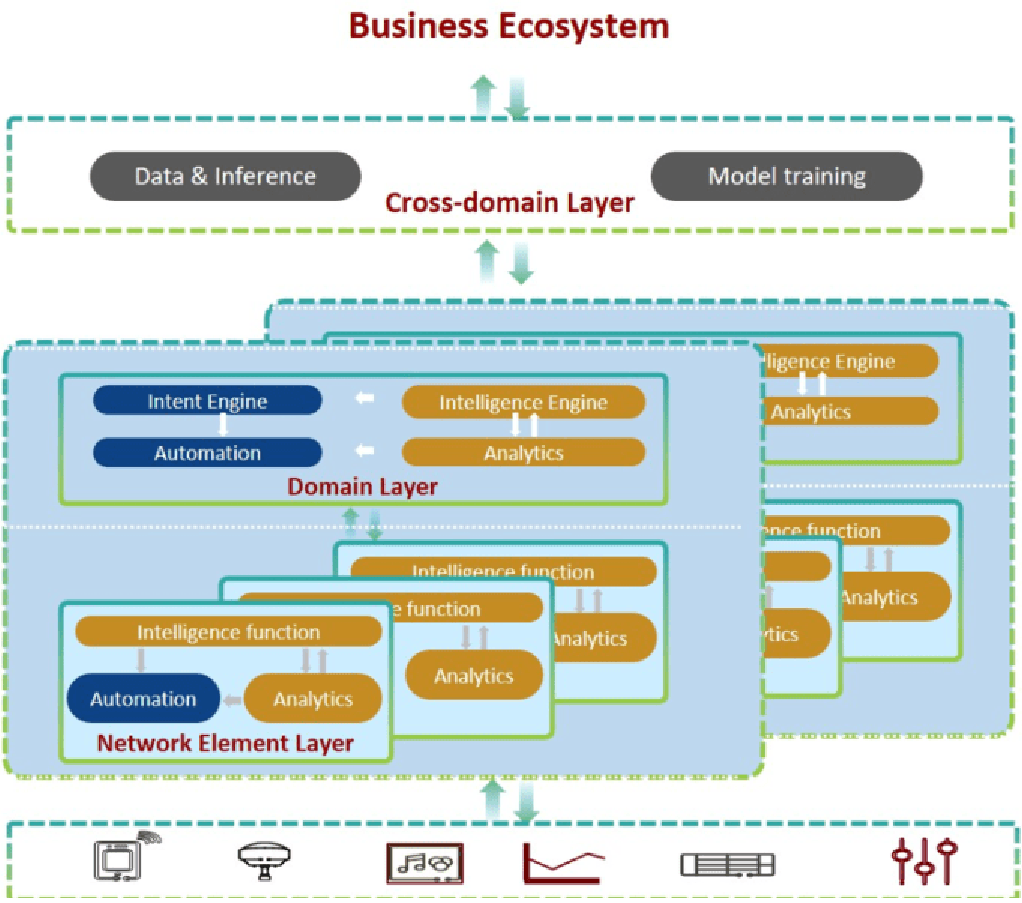

Figure 12: Layered Framework to support AI in Network

Evaluation: The GSMA Layered Framework as the TM Forum Framework is a conceptual framework and even with the description of the use cases in China does not provide sufficient details for the implementation and deployment of autonomic functions in mobile wireless networks such as 5G. The most important contribution is probably the open simplified interface between the network domain layer and the upper domain layers that will have to standardised or at least adopted as de-facto industry standard.

2.2 Elaboration of testing frameworks for mobile wireless networks

The testing of mobile wireless networks is complex, resource and time-consuming task that calls for the use of AI/ML in testing. Fortunately, there are currently a few initiatives that aims at proposing frameworks for testing future networks.

2.2.1 5GENESIS

As part of the 5G-PPP initiative, the EU-funded 5GENESIS project [25] realised a 5G facility composed of five different testbeds in Europe, accessible for both per-component and E2E experimentation purposes. A reference architecture for common and lightweight access to the 5GENESIS testbeds has been already defined along with and E2E methodology for testing and validation of 5G technologies and KPIs.

The 5GENESIS testing methodology follows a modular approach and it includes three logical components, referring to three configuration/input information classes required for running an experiment, namely, the test cases, the scenarios, and the slices. Altogether, they identify an experiment, the definition of which is also formalized in a global template, referred to as Experiment Descriptor (ED). From the architectural point of view, the information enclosed in the ED feeds a core functional block of the 5GENESIS reference architecture, the Experiment Life-cycle Manager (ELCM). ELCM manages the testing procedure and allows for automatic execution of experiments.

Subsequently, a further architectural block, i.e., the Monitoring and Analytics (M&A) framework oversees collecting KPI samples and complementary measurements during the experiment (Monitoring) and process them for a statistical validation of the KPI under test (Analytics). The M&A components are shown in Figure 13Figure 13.

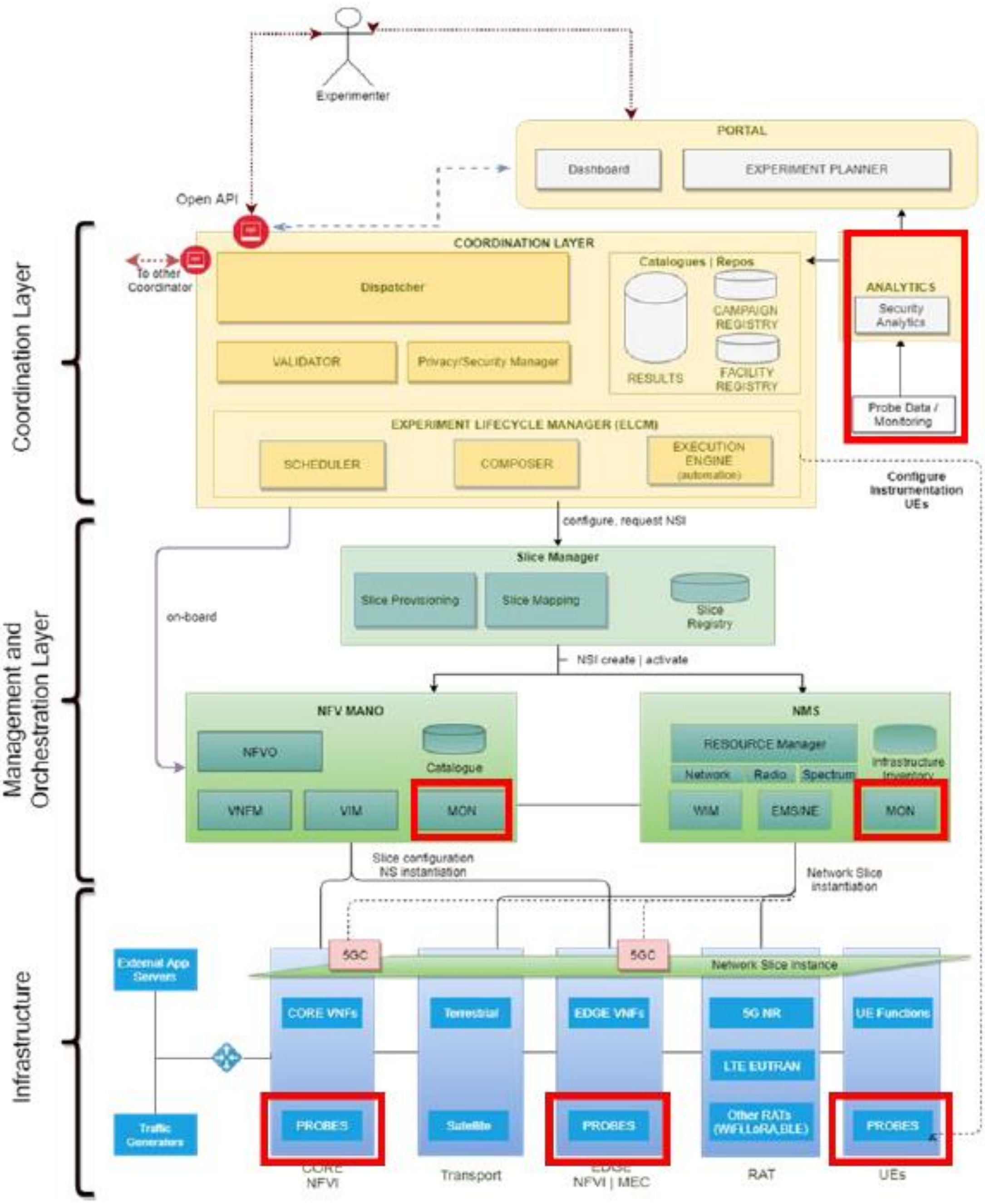

Figure 13: 5GENESIS reference architecture with M&A components highlighted in red

The 5GENESIS M&A framework is designed and implemented as 3 main interoperable functional blocks [26]:

- **Infrastructure Monitoring (IM**): focuses on the collection of data that synthesize the status of architectural components, e.g., end-user devices, radio access and networking systems, computing and storage distributed units;
- **Performance Monitoring (PM)**: is devoted to the active measure of E2E QoS/QoE indicators.
- **Storage and ML Analytics**: enables efficient management of large sets of heterogeneous data and drives the discovery of hidden values and correlation among them;

Evaluation: The 5GENESIS M&A framework is a considerable contribution for the testing of 5G networks. The collection of data is done by the infrastructure (IM). However, it is worth noting that the collected data are metrics that record a state or change in state. They are neither user traffic data nor control traffic data. The 5GENESIS M&A framework focuses more on traditional testing than on AI/ML based testing.

2.3 Study on the Benefits that AI/ML brings to Test Systems

The benefit provided by test systems embedding AI/ML compared to the traditional ones, such as reduction of test suite execution time in functional and performance testing of complex systems, should emphasize the metrics of such benefits. This topic is researched by ETSI TC INT (Technical Committee Core Network and Interoperability Testing) and ETSI TC MTS (Technical Committee Methods for Testing and Specification).

2.3.1 ETSI work item in AI in Test Systems and Testing AI Models

ETSI TC INT and ETSI TC MTS have defined as part of the deliverables of the Newly Launched Work Item in AI in Test Systems and Testing AI Models a ***General Guide on the Benefits of Artificial Intelligence (AI) in Test Systems***, with illustrations of Artificial Intelligence (AI) in Test Systems and the Benefits [24].

This is an ongoing work but the metrics on what AI brings to Performance Testing Systems for Complex Systems, i.e., Targets for AI in Performance Test Tools can be summarized as follows:

1. Speeding up and improving quality of evaluations of performance measurements;
2. Supporting or producing correct configuration of different performance measurements based on the purpose and conditions of a measurement, i.e., the user specifies what are the targets and conditions of an intended performance measurement and the test tool with AI does the rest;
3. AI could also find out all additional performance characteristics and KPIs that can be produced from the captured data of a performance measurement;
4. AI will also enable completely autonomous performance measurements.

Evaluation: The ETSI work item is very relevant and valuable but the identified metrics will have to be standardised to enable the test and certification of AI/ML models. This calls for more studies and collaboration in the scientific community.

2.4 Testing and certification of AI/ML models

The AI/ML models play a decisive role in the performance, precision, accuracy (i.e., small number of false negative), precision (i.e., small number of false positive), sensitivity and efficiency of the autonomic functions. Therefore, they must be tested properly and certified. Testing AI Models involves the validation of the input data. The effectiveness of the decision-making AI in Test Systems depends on the quality of the training data. Further, the various AI/ML algorithms constituting the core AI engine must be subject to validation after rigorous tests using various input data. Performance tests must also be performed to determine the efficiency of the AI/ML models. Integration tests must be carried out to verify the ability of the AI/ML to work with the other parts of the system.

2.4.1 ETSI work item in AI in Test Systems and Testing AI Models

ETSI TC INT and ETSI TC MTS has initiated a work item to define a *"General Approach to Designing the Test Systems/Components for Testing Autonomics AI Models, and Challenges in AI Models ability to cope with 5G Network Dynamics that need to be taken into consideration"*.

They identify some general aspects that need to be taken into consideration by both, the developer of an AI Model and the Tester of an AI Model as follows:

- Time it may take for an AI Model for autonomic management and control to meaningfully be applicable and be able to keep pace with dynamics of the network;
- Time it may take for an AI Model embedded in a Test Component/System to meaningfully be applicable and be able to keep pace with dynamics of the network;
- Verdicts Passing in Testing AI Models, and How Suppliers of AI Models (e.g., Cognitive GANA DEs) to be Tested and Certified can produce "Assertions/Claims Specifications of Measurable Metrics/KPIs and certain observable and verifiable outputs" on what the AI Model can achieve under certain conditions during its operation as "assertions/claims" by the supplier of the AI Model;
- Idea of using the concept of a "Qualified Automated Test Component(s) or System" that exhibits best quality AI capabilities, in testing comparable capabilities of AI Component(s)/System Under Test.

There are various metrics that can be used for assessment and differentiation of AI Models such as Machine Learning and should be standardised as follows:

- Stability of the AI Model;
- Speed of Learning of the AI Model;
- Speed of Decision-making cycle of the AI Model after receiving triggering inputs;
- Speed of Convergence of multiple interacting AI Models/components in a larger AI System;
- Quality of Decision-Making of the AI Model.

Evaluation: This work item is necessary for the adoption of AI/ML models and a lot more work is needed to standardise the metrics for assessment.

2.5 Testing and certification of AI/ML autonomic functions

2.5.1 ETSI Generic Test Framework for Testing ETSI GANA Model's Multi- Layer Autonomics & AI Algorithms for Closed-Loop Network Automation

ETSI has specify a Generic Autonomic Networking Architecture (GANA) Standard in ETSI TS 103 195 and starts now working on the specification for a Generic Test Framework for Testing ETSI GANA. The Generic Test Framework's objective is to provide guidance on various aspects such as the following ones:

- Conformance Testing and Interoperability Testing for GANA Functional Blocks (DEs, ONIX, MBTS) based on their Reference Points and Characteristic Information exchange expected on the Reference Points; i.e., Conformance Testing and Interoperability Testing is required on the Reference Points for Autonomics instantiated in a target architecture and environment;
- Criteria for use in Verdicts passing when Testing Autonomic Functions (AFs), i.e., GANA Des;
- The role of a GANA Meta-Model in generation of Data Types for use in Test Suites Development.

Evaluation: This work item is quite important not only for the ETSI GANA but also for every autonomic system including future mobile wireless networks. It is however ongoing and not yet completed.

2.5.2 The NGMM Generic framework for testing and certifying autonomic functions

As described earlier, NGMM proposes a *Generic autonomic networking architectural model* that enables a variety of flexible implementation and deployment strategies for autonomic functions. To develop the autonomic capabilities of self-configuration and self-adaptation, embodied within the self-CHOP (Configure, Heal, Optimize, Protect) characteristics in a 5G and beyond system, the standardization and testing methods for AI and ML models are required for consistent cognitive capabilities in an autonomic network. NGMM proposes a test framework for testing multi-layer autonomics, and associated AI and ML algorithms for closed-loop network automation. The objective is to ensure that the decision-making functions are compliant with the KPI targets objectives for cross-domain self-optimization behaviours for resource utilization in the relevant network segments of an end-to-end system.

Evaluation: As for the ETSI Test framework the NGMM Generic framework for testing and certifying autonomic functions is ongoing and will deserve to have more efforts to complete.

2.6 Framework for evaluating network intelligence capabilities

With the adoption of autonomy functions, mobile wireless networks become more and more autonomous and intelligent. Consequently, it is important to know and be certain about how intelligent a network is to use and rely on it. Well-aware on the importance of this demand, ITU has initiated an activity aiming at standardising a framework for the evaluation.

2.6.1 Framework for evaluating intelligence levels of future networks including IMT-2020

As networks become more and more intelligent [27], it is important to adopt a standard method for evaluating network intelligence levels.

A standard method for evaluating network intelligence levels has the following implications:

- It provides an evaluation basis for measuring the intelligence levels of a network and of its components;
- It helps the industry to reach a consensual and unified understanding of network intelligence concepts;
- It provides a reference for industry supervisors to formulate relevant strategies and development planning of future networks including IMT-2020 in various countries;
- It provides a decision mechanism to operators, equipment vendors and network industry participants for planning of network technology features and products' roadmaps.

Network intelligence level		**Dimensions**				
		Action implementation	**Data collection**	**Analysis**	**Decision**	**Demand mapping**
L0	Manual network operation	Human	Human	Human	Human	Human
L1	Assisted network operation	Human and System	Human and System	Human	Human	Human
L2	Preliminary intelligence	System	Human and System	Human and System	Human	Human
L3	Intermediate intelligence	System	System	Human and System	Human and System	Human
L4	Advanced intelligence	System	System	System	System	Human and System
L5	Full intelligence	System	System	System	System	System
NOTE 1 – For each network intelligence level, the decision process has to support intervention by human being, i.e., decisions and execution instructions provided by a human being have the highest authority. NOTE 2 – It is to be noted that this table may be used to only determine the network intelligence level for each dimension (and not the overall network intelligence level).						

Figure 14: Network intelligence levels [27]

Evaluation: ITU has defined 6 levels of intelligence which provides a common understanding of the intelligence of a network. However, more details will have to be specified in order to determine which intelligence level a network actually has. More work is needed in this area.

2.7 Intent-driven management

With the emergence of autonomic networks, a new concept called Intent-driven management has been introduced. Intent-based networks enable service providers to define the behaviours expected from their network such as “improving network quality” without having to specify the QoS parameters. The intent-driven network understands the service provider’s intention and will act accordingly in real time to ensure that the intent is fulfilled. Intent-driven management is currently focused work item at 3GPP and TM Forum.

2.7.1 3GPP Study on scenarios for Intent driven management services for mobile networks

According to [28] an Intent driven management service (MnS) allows its consumer to express desired intent for managing the network and services. The Intent driven MnS producer transfers the intent to executable actions for service assurance and deployment.

The executable actions can be one or more of the following:

- Perform network management tasks;
- Identifying, formulating and activating network management policies.

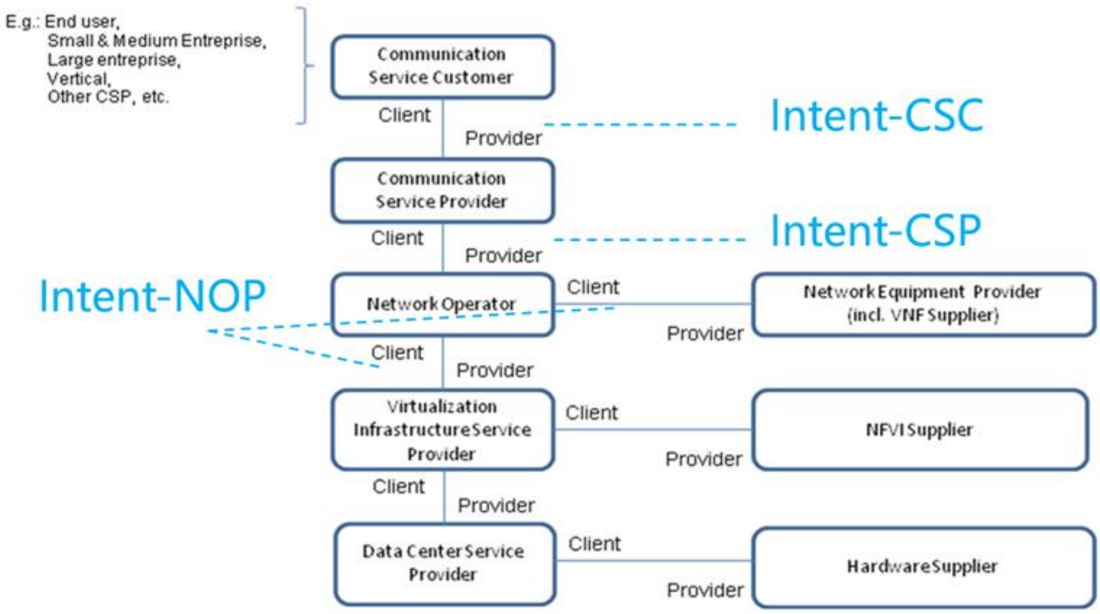

Figure 15: Concept for utilization of intent

As shown in Figure 15, 3GPP defines different kinds of intents applicable to different kinds of standardized reference interfaces as follows:

- Intent from Communication Service Customer (Intent-CSC);
- Intent from Communication Service Provider (Intent-CSP);
- Intent from Network Operator (Intent-NOP).

Evaluation: Although the 3GPP study on Intent-driven management service is very important, it is only paving the way for further work on Intent-driven networks. Indeed, standardisation is required for a few items as follows:

- Intent driven management related definition and concepts;
- Roles that are related to the intent driven management;
- Typical scenarios and management requirements for intent driven management, which can reduce the management integration complexity;
- The mechanisms for reuse the generic management service CRUD operation to realize Intent driven MnS;
- Specification of the modelling solution including Intent Driven Action and Intent Driven Object for the identified scenarios.

2.7.2 TM Forum Intent-driven interaction

Autonomous Networks need to be able to adapt their operation to the business objectives of the operator as well as to the expectations of customers and users. The role of intent is to communicate all these expectations to the Autonomous Networks. Intent defines goals, targets, requirements, and constraints that constitute machine-processable knowledge.

Intent defines what Autonomous Networks are expected to achieve, but it leaves the details of how a network is designed and operated to the internal operations of the network platform. This means that the smart software in the platform can constantly optimize how the service is delivered and we can incrementally add new technologies like analytics and machine learning to constantly improve the implementation.

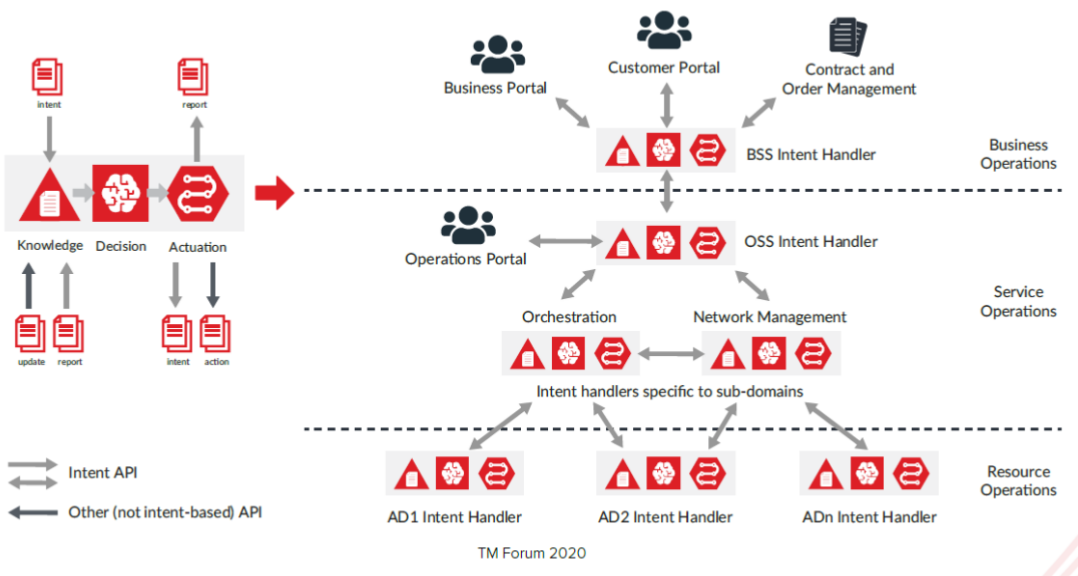

Figure 16: TM Forum Autonomous Networks operation based on intent handling [21]

TM Forum has defined an intent-interaction model as shown in Figure 16 [21]. For Autonomous Networks, the intent-interaction model is a necessary requirement. The role of intent goes beyond relieving the burden of the user knowing implementation details. More importantly, it sets the autonomous system's internal goal such that the system can take proactive actions to achieve the stated goal based on its observation of the environment.

Intents represent the concerns and objectives of the users of an Autonomous Networks. It therefore varies with the diversity of user types and roles:

- **Business intent** represents the objectives of a business user. This includes for example the delivery of a custom application defined by SLAs. Operators expect their Autonomous Networks to operate service contracts while meeting revenue targets. Their customers expect a good user experience.

- **Service intent** represents the objectives of a service user. A service is expected to deliver functional as well as non-functional attributes. This includes targets on areas including connectivity, bandwidth, latency, or availability for example.
- **Resource intent** represents the objectives of resource users. Resources are expected to be allocated so that performance and quality of service targets are met.

Evaluation: Intent-driven management is definitely a major step towards business and customer orientation of future mobile wireless network. While it is compelling to make use only of the business intents or service intents, these intents will need to be translated to resource intents to be properly enforced. There is obviously a need for testing and validating of the translation between the three types of intent in both ways.

2.8 Network simulators

As mentioned earlier the use of AI/ML in mobile wireless networks will enable the introduction of autonomic functions that make the network capable of self-configuring, healing, optimising and protecting (CHOP). This is very important for operators because of the complexity required by advanced mobile wireless networks such as 5G. However, the usability and efficiency of AI/ML depends on their training that, again, rely on the quality of the training data. Unfortunately, for newly deployed networks and networks that have yet to be deployed, the training data are scarce or even not available. It is also challenging to train and transfer ML models between networks due to configuration and operation differences.

In this situation network simulators can be presented as a possible solution. By simulating the planned future network, a network simulator can generate data necessary for the training and elaboration of ML models. ITU has addressed the ML training issue in its architectural framework for machine learning.

2.8.1 ITU Architectural framework for machine learning in future networks including IMT-2020

The ITU–T Y.3172 recommendation [29] proposes a ML sandbox as a solution for enhancing confidence in machine learning application. A machine learning sandbox is defined as an environment in which machine learning models can be trained, tested and their effects on the network evaluated.

An ML sandbox is an isolated domain that allows the hosting of separate ML pipelines to train, test and evaluate them before being deployed in a live network. For training or testing, the ML sandbox can use data generated from simulated ML underlay networks and/or live networks.

As shown in Figure 17, the framework consists of subsystems that are connected together via **reference points** as:

1, 2: These are the data handling reference points between simulated ML underlay networks and an ML pipeline in an ML sandbox subsystem;

3: This is the reference point between an ML sandbox subsystem and an ML pipeline subsystem;

4: This is the reference point between an ML pipeline subsystem and ML underlay networks; This reference point represents the data handling reference points shown as arrows 3 and 4 in Figure 17.

5, 6: These are the reference points between the management subsystem and, the ML pipeline subsystem and ML sandbox subsystem, respectively.

7: This is the reference point between the MLFO and other management and orchestration functions of the management subsystem.

8, 9: These are the reference points between ML pipeline nodes located in different levels.

Reference point 3 is quite important since it reflects the connection between the sandbox and the real system.

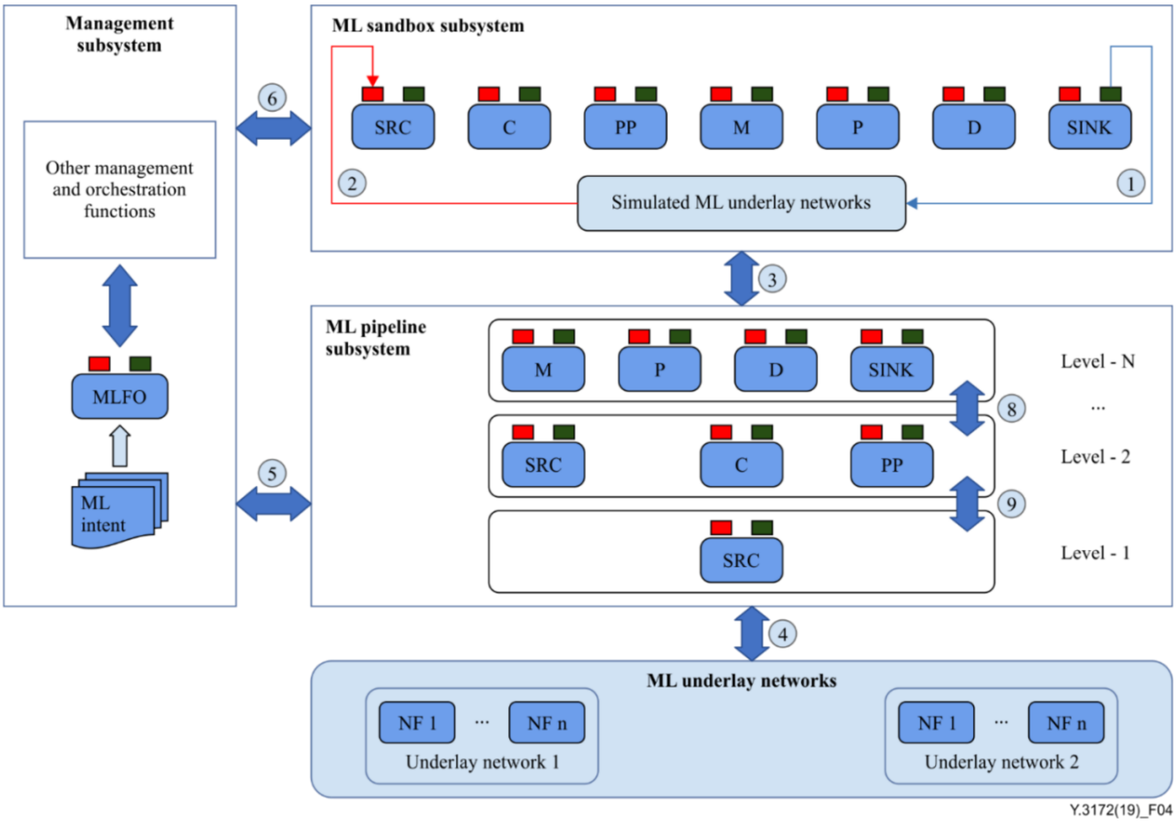

Figure 17: High level architecture of ITU Architectural framework for machine learning [29]

Evaluation: Network simulators will play an essential role in building reliability and trust in ML and most importantly in the realisation of future autonomic mobile wireless networks. Although ITU has specified a ML sandbox in their high-level architecture, more details are required for the deployment of such a sandbox. There are currently very few studies about network simulators for AI/ML in mobile wireless networks (see [30]).

3. Conclusion and Recommendations

This paper investigates the motivations for introducing AI/ML in the testing of mobile wireless networks. The rationale for the exploitation of AI/ML in testing is twofold. First, in order to match the demand for an increased flexibility and customisation, the software development process -including testing- must be agile. This can be achieved by adopting AI/ML as a solution. Second, AI/ML can equip future wireless networks with some degree of autonomy. Work is in progress but deserve some more studies especially in the following areas some topics that need to be researched further as follows:

3.1.1 Data collection

Although data processing and use are specified in several proposed architectures and frameworks, some issues are still pending:

- How are data actually monitored and captured? In the 3GPP's architecture enhancements, a NWDAF is introduced to collect data from every NF. The NF is then responsible for the monitoring and capture of the data but it is not clear whether this function should be included inside the NF or whether it should co-exist side by side with the NF. The data capture may be different depending on the type of data, i.e., protocol, format, length, etc. but its interfaces should be defined and standardised;

- What kind of data needs to be captured? In the 5Genesis architecture, a Monitoring and Analytics (M&A) framework has been specified that carries out the collection of data. However, only the collection of metrics that provide status information is collected, not containing sufficient details as required by AI/ML. Indeed, for the realisation of autonomic functions it may be necessary to monitor and capture both control traffic and user traffic data. Currently, there is a lack of SoTa on this topic;
- How much and for how long? The data needed for different purposes such as performance & load optimisation, cyber security, privacy protection, are different. Collecting a larger amount of data than needed will create storage problem. In addition, it is not clear how long the data should be stored or achieved;
- How the data aggregation is done? Data is captured at different locations in the network and must be aggregated properly. Quality check and proper time stamping are required to ensure consistency.

3.1.2 Guideline for AI/ML testing and validation

The EC's white paper "WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust" [31] states that "due to the non-deterministic and context-specific nature of AI, traditional testing is not enough". As ETSI also specified in its white paper [32] it is necessary to study and possibly generalize the available best practices and propose specification methods for AI capabilities in new standards, that are "testable" by design, or to promote the use of AI to improve the traditional testing models as well management and selection of test suites.

3.1.3 Transfer of knowledge and AI/ML models

It may be desirable to transfer and reuse knowledge gained in one mobile wireless network to another one. Indeed, it is rather straightforward to transfer AI/ML models between networks of same type, e.g., 5G. Unfortunately, networks differ in number of users, applications and usages, amount and frequency of traffic, etc. The relevant attributes in one network may not be valid for another one. Consequently, a dataset from one network may not be usable in other ones. An AI/ML platform for a newly deployed mobile wireless network will need to be trained before being fully operational. The training time could be long and there is a need for research aiming at reducing it.

3.1.4 Network simulators

Network simulators may be an efficient tool to reduce the training time of new mobile wireless networks, since they may emulate not yet deployed networks and enable the collection of data necessary for the training of AI/ML functions in the autonomous network. There exists an important need for research and development of simulators for future networks, capable of simulating normal operations and diverse anomalies.

References

1. Ovum: Ovum Decision Matrix: Selecting an Application Lifecycle Management and DevOps Solution, 2019–20
2. J. Chard & B.P Douglass: Agile Product Development For Dummies, IBM Limited Edition, ISBN: 978-1-119-17737-1 (ebk), John Wiley & Sons, Inc.
3. Kent Beck; James Grenning; Robert C. Martin; Mike Beedle; Jim Highsmith; Steve Mellor; Arie van Bennekum; Andrew Hunt; Ken Schwaber; Alistair Cockburn; Ron Jeffries; Jeff Sutherland; Ward Cunningham; Jon Kern; Dave Thomas; Martin Fowler; Brian Marick (2001). "Manifesto for Agile Software Development". Agile Alliance. Retrieved 14 June 2010.
4. S. Sharma & B. Coyne: DevOps for Dummies, IBM Limited Edition, ISBN: 978-1-119-41588-6 (ebk), John Wiley & Sons, Inc.
5. TechJury: How many IoT devices are there in 2021; 04.01.2022; accessed: 12.01.2022- https://techjury.net/blog/how-many-iot-devices-are-there/#gref
6. Perfecto: The buyer's guide to Web & Mobile Test Automation Tools; accessed 12.01.2022– http://www.perfecto.io
7. Empower: EMPOWER White Paper - Empowering Transatlantic Platforms for Advance Wireless Research, October 30, 2019
8. J. O. Kephart and D. M. Chess, "The vision of autonomic computing," in Computer, vol. 36, no. 1, pp. 41-50, Jan. 2003, doi: 10.1109/MC.2003.1160055
9. ITU: ITU News Magazine No.5 2020 - AI and Machine Learning in 5G Lessons from the ITU Challenge
10. NGMN Alliance: 5G End-to-End Architecture Framework (Phase-3) V4.31, Final Deliverable (approved), 12-Nov-2020
11. 3GPP: 3GPP TS 23.501 V17.1.1 (2021-06) - System architecture for the 5G System (5GS); Stage 2 (Release 17)
12. 5GPPP: AI and ML – Enablers for Beyond 5G Networks ver 1.0 11.05.2021 - DOI 10.5281/zenodo.4299895
13. 3GPP: 3GPP TS 23.288 V17.1.0 (2021-06) - Architecture enhancements for 5G System (5GS) to support network data analytics services (Release 17)
14. 3GPP: 3GPP TR 23.791 V16.2.0 (2019-06) - Study of Enablers for Network Automation for 5G (Release 16)
15. 3GPP: 3GPP TS 23.502 V17.1.0 (2021-06) - Procedures for the 5G System (5GS); Stage 2 (Release 17)
16. NGMM: https://www.ngmn.org/
17. ITU: Y.3172 (06/2019) Architectural framework for machine learning in future networks including IMT-2020
18. ETSI: ETSI GS ENI 005 V1.1.1 (2019-09) - Experiential Networked Intelligence (ENI) System Architecture
19. ETSI: ETSI GS AFI 002 V1.1.1 (2013-04) - Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture (An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management)
20. ETSI: ETSI TS 103 195-2 V1.1.1 (2018-05) - Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture; Part 2: An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management.
21. TM Forum: White Paper Autonomous Networks: Empowering digital transformation for smart societies and industries, Release 2, October 2020
22. GSMA: AI & Automation: An Overview - https://www.gsma.com/futurenetworks/wiki/ai-automation-an-overview/
23. GSMA: AI in Network Use Cases in China, Oct 2019
24. ETSI: White paper No. 5 Artificial Intelligence (AI) in Test Systems, Testing AI Models and ETSI GANA Model's Cognitive Decision Elements (DEs) via a Generic Test Framework for Testing GANA Multi-Layer Autonomics & their AI Algorithms for Closed-Loop Network Automation – 29.03.2020
25. https://5genesis.eu
26. 5GENESIS: Deliverable D3.5 Monitoring and Analytics (Release A)
27. ITU: ITU–T Y.3173 Framework for evaluating intelligence levels of future networks including IMT-2020, 02/2020
28. 3GPP: 3GPP TR 28.812 V17.1.0 (2020-12) Technical Report - Study on scenarios for Intent driven management services for mobile networks (Release 17)

29. ITU: ITU-T Y.3172 Architectural framework for machine learning in future networks including IMT-2020 06/2019
30. Francesc Wilhelmi, Marc Carrascosa, Cristina Cano, Anders Jonsson, Vishnu Ram, and Boris Bellalta: Usage of Network Simulators in Machine-Learning-Assisted 5G/6G Networks, innoarXiv: 2005.08281v2 [cs.NI] 2 Mar 2021
31. "WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust". COM(2020) 65 final. Published 19th February 2020. https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en
32. ETSI White Paper No. #34 Artificial Intelligence and future directions for ETSI 1st edition – June 2020 - ISBN No. 979-10-92620-30-1