IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 May 2022, accepted 6 June 2022, date of publication 17 June 2022, date of current version 30 June 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3183989

Toward Integrating Intelligence and Programmability in Open Radio Access Networks: A Comprehensive Survey

AZADEH ARNAZ[1], (Graduate Student Member, IEEE), JUSTIN LIPMAN[1], (Senior Member, IEEE), MEHRAN ABOLHASAN[1], (Senior Member, IEEE), AND MATTI HILTUNEN[2], (Senior Member, IEEE)

[1]School of Electrical and Data Engineering, University of Technology Sydney, Ultimo, NSW 2007, Australia
[2]AT&T Labs Research, Bedminster, NJ 07921, USA

Corresponding author: Azadeh Arnaz (azadeh.arnaz@student.uts.edu.au)

ABSTRACT Open RAN is an emerging vision and an advancement of the Radio Access Network (RAN). Its purpose is to implement a vendor and network-generation agnostic RAN, provide networking solutions across all service requests, and implement artificial intelligence solutions in different stages of an end-to-end communication path. The 5th Generation (5G) and beyond the 5th Generation (B5G) of networking introduce and support new use cases, such as tactile internet and autonomous driving. The complexity and innovative nature of these use cases require continuous innovation at a high pace in the RAN. The traditional approach of building end-to-end RAN solutions by only one vendor hampers the speed of innovation—furthermore, the lack of a standard approach to implementing artificial intelligence complicates the compatibility of products with the RAN ecosystem. O-RAN Alliance, a community of industry and academic experts in RAN, works on writing Open RAN specifications on top of the 3rd Generation Partnership Project (3GPP) standards. Founded on these specifications, the aim of this paper is to introduce open research topics in Open RAN that overlap the interests of both AI and telecommunication researchers. The paper provides an overview of the architecture and components of Open RAN, then explores AI use cases in Open RAN. Also, this survey includes some plausible AI deployment scenarios that the specifications have not covered. Open RAN in future cities creates opportunities for various use cases across different sectors, including engineering, operations, and research that this paper addresses.

INDEX TERMS 5G, B5G, artificial intelligence, intelligent systems, machine learning, open RAN, radio access networks.

I. INTRODUCTION

Telecommunication Radio Access Network (RAN) forms a significant part of wireless communication and has evolved considerably in the past two decades. This paper explores a new architecture of RAN named Open RAN and its role in the era of intelligent telecommunication. The RAN defines the logical group of components between the receiver and the core network in end-to-end telecommunication systems. These components communicate via interfaces. Each network generation has introduced new capabilities and possibilities to the public. As a result, the RAN evolved as a response to new requirements. Currently, the 5th Generation (5G) and Beyond 5th Generation (B5G) are in implementation and study. 5G and B5G have a much higher frequency than the previous generations. Therefore they are capable of transporting more data than the previous generations. However, high frequency limits signal attenuation through obstacles reducing non-line of sight coverage. 5G and B5G can enable more low-latency mission-critical and high data rate use cases despite challenges in coverage.

To achieve these requirements, the new RAN architecture should encourage a fast-paced and innovative approach in telecommunication. Considering the complexity and the latency of decision-making tasks in 5G and B5G,

The associate editor coordinating the review of this manuscript and approving it for publication was Jiankang Zhang.

Artificial Intelligence (AI) will play a vital role in developing a new RAN architecture called "Open RAN". Consequently, AI is increasingly becoming relevant for 5G and B5G network deployment. Open RAN defines open interfaces to encourage a multi-vendor system. This intention is to introduce flexibility and boost innovation in a competitive market. Also, Open RAN has dedicated logical components for enabling intelligent applications that control RAN communications. This approach creates a global standard and framework for all 5G and B5G vendors and operators to implement AI solutions. Nevertheless, there are challenges to implementing this architecture in practice. For example, the number of components in the Open RAN and security or compatibility risks raised by this multi-vendor architecture has created research and engineering projects across different industries such as hardware, software, machine learning, and security. As a result, Open RAN has been the subject of significant research and standardization efforts over the last few years. To the best of our knowledge, there are some topics that other published works have not covered, and this paper addresses them. These topics are a comprehensive survey of Open RAN focused on AI applications within Open RAN, the use of AI within Open RAN, deployment scenarios in artificial intelligence, and future opportunities in this area. The motivation for this work stems from the growing research interests in Open RAN and the impact it is expected to make for future wireless networks. Furthermore, the interaction of ML/AI with Open RAN introduces a tremendous opportunity to further enhance and optimise the performance of future wireless networks, in particular it opens up a pathway towards developing fully autonomous or self-driving networks. Hence, this survey paper comprehensively surveys the current research and build a pathway for future search directions over Open RAN. The specific contributions are as follows:

1) A holistic study of AI in telecommunication solutions, classified based on the Open RAN controllers classification is presented.
2) ML deployment scenarios that are not covered in the standards are discussed.
3) Future opportunities and challenges that the combination of AI and Open RAN in 5G and B5G can provide are presented.

The organization of this paper is as follows. Section II provides a brief history of the RAN, while Section III introduces Open RAN. Section IV explores the proposed Open RAN architecture that is standardized. Section V explores the use of AI to solve problems within Open RAN in telecommunications. Section VI focuses on categorizing the Open RAN architecture components for running near real-time or non-real-time intelligent applications and explores some of these AI applications. Section VII addresses an essential branch of machine learning: the deployment of machine learning pipelines in production, known as MLOps. Finally, Section VII explores opportunities and challenges for future research.

II. HISTORY OF RAN

The end goal of any network infrastructure design is to provide seamless and secure communication between devices. Radio Access Network (RAN) is one of the critical concepts in telecommunication that facilitates this goal. Conceptually, RAN resides between a device such as a mobile phone, a computer, or any remotely controlled machine and the core network (CN) [1].

A. NETWORKING GENERATIONS

In [2] Henrik A. *et al.* describe the evolution of telecommunications networking known as "Networking Generations." The public started using the first generation of mobile phones (1G) in the early 1980s. Although in the early 1990s, the 2G Global System for Mobiles (GSM) telecommunication system created a pivot point in telecommunication, its functional architecture was static. Further, network functionalities were geographically localized, and all radio-related functionalities operated within the base station.

The next generation, 3G UMTS (Universal Mobile Telecommunication System) terrestrial RAN, split radio functionalities into two parts. One part was for functionalities such as transmission and reception, which were the responsibility of NodeB (a radio base station receiver in 3G). The other part included radio resource management and higher-layer RAN user processing in 3G running on RNC (Radio Network Controller). Splitting functionalities and running them on two separate parts created latency due to control processes between NodeB and RNC. But because RNC could execute resource allocation tasks faster than the previous generation, the overall latency in 3G became less than 2G.

Upgrading 3G to an LTE-advanced (Long-Term-Evolution-advanced) version led to 4G. 4G changed expectations on data transfer rate and security management. However, new use cases demand more and faster data transportation in wireless communication that exceeds the capacity of 4G. 5G and beyond 5G (B5G) are designed to meet these capacity requirements for future capacity. The following section discusses the evolution of RAN from D-RAN to v-RAN, which led to the foundation of Open RAN.

B. D-RAN

Early generations of wireless networks had a Baseband Unit (BBU) and a Remote Radio Head (RRH) component, both physically located in Base Stations (BS). BBU and RRH were connected to the Radio Frequency (RF) antenna at the top of the tower through electrical cables. This design experienced RF signal loss. As a result, telecommunication experts designed Distributed RAN (D-RAN), in which BBU and RRH are separated. As Fig. 1 shows, in D-RAN, each Base Station(BS) includes a BBU and an RRH (Radio Remote Head), also called Remote Radio Unit (RRU). BBU connects to RRH through a Common Public Radio Interface (CPRI). BBU is responsible for baseband processing which includes processing calls and forwarding traffic. RRH is responsible

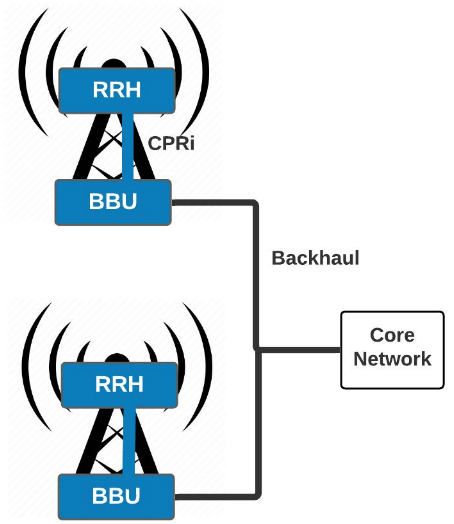

FIGURE 1. D-RAN.

for transmitting, receiving, and converting signals. Base stations are connected to the core network individually through a separated backhaul. One of the main limitations of this architecture is that the result of BBU processing can be shared only with the coupled RRH. The increase in demand for services disclosed other limitations of this architecture, such as low spectral efficiency, high cost of scaling this architecture, and inefficient use of resources.

C. C-RAN

C-RAN (Cloud-RAN or Centralized Access Network) is a response to the limitations of D-RAN. It decouples RRUs and their corresponding BBUs. RRUs stay at the cell site in this architecture and connect to a centralized but shared and virtualized group of BBUs. Every RRU connects to a BBU pool via a fronthaul link. Each BBU pool can serve more than one RRU unit and connects to the Core network through a backhaul link. As shown in Fig. 2, C-RAN comes in two types: ''Fully centralized'' and ''Partial centralized''. The main difference between these two types is the functionalities related to layer one, such as sampling, modulation, resource block mapping, antenna mapping, and quantization. These functionalities happen in BBU pools in the fully centralized type, but they run on the RRU unit in the partially centralized type.

D. vRAN

Virtualized RAN (vRAN) is a revolutionary improvement in networking. vRAN is a disaggregated RAN architecture. In vRAN, the software is separated from the underlying hardware instead of running software on expensive proprietary hardware units. This modification makes it possible to run networking functionalities dynamically and flexibly. It also reduces the cost of maintenance and operation because a modification of network functionalities does not require hardware-related amendments. Also, functions can run on

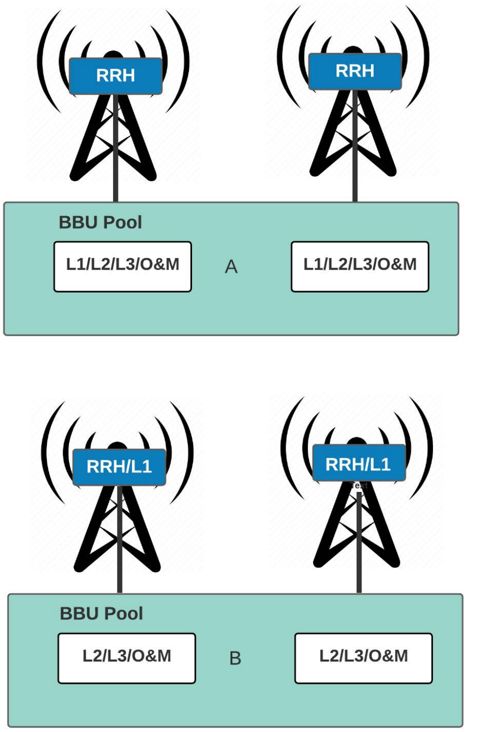

FIGURE 2. C-RAN. A: Fully centralized, B: Partial centralized.

common-off-the-shelf (COTS) hardware. Operators can use monitoring applications and manage load balancing and performance-related functionalities. Moreover, it creates an agile and scalable environment for Open RAN to achieve its goals.

III. INTRODUCTION OF OPEN-RAN

In this section, we introduce Open RAN. We have divided this section into two parts. In the first part, we talk about the research history of Open RAN and examples of some published works on this topic. In the second part, we introduce communities and organizations in the industry that implement Open RAN for 5G and B5G telecommunication systems.

A. OPEN RAN IN RESEARCH

The idea of having a RAN with open interfaces is not very recent. In 2002, the Mobile Wireless Internet Forum IP in the RAN working group developed a version of Open RAN, in collaboration with researchers from Cisco Systems and DoCoMo Communication Lab. This collaborative team published a paper describing the need and requirements of Open RAN [3]. In that paper, the authors addressed the challenges of RAN architectures in scalability and reliability. For instance, the expansion of RAN is costly. Furthermore, the centralized control system becomes the *single point of failure*. The paper proposed an Open RAN architecture that included separate mobile node control functions; and supported multiple radio technologies, including 2G, 3G, Wireless LAN (Local Area Network), and any upcoming technologies in this area. The proposed architecture also supported the operation and administration of multi-vendor solutions, hence having open interfaces. In [3], the authors described the obstacles of implementing an Open RAN architecture, such as managing and orchestrating an Open RAN architecture. Existing RANs are not globally standardized, and different operators might have different administration procedures. As a result, making components from different vendors compatible with each other and the ecosystem might become an obstacle during the installation of a system. In addition, decommissioning a network generation and replacing it with a new one is costly and requires a significant amount of time. Operators need to upgrade their networking generation alongside the running legacy network generations to avoid any negative impact on QoE for customers. A modular, multi-vendor architecture such as Open RAN should ensure that it can provide services to all current and future network generations such as 4G, 5G, and B5G.

Following on from [43], there has been a significant number of Open RAN-related publications. Virtualized Network (vRAN) created an environment to implement Open RAN architectures. In [4], the authors address the role of virtualized RAN by discussing the role of SDN/NFV (Software Defined Network/Network Function Visualization) in Open RAN. [5] explored a networking ecosystem that is multi-vendor and uses open-source software products to run an Open RAN architecture. [6] discusses the scalability and reliability of xApps in production. In [4] authors explored the role of AI in IoT (Internet of Things) and 5G within the context of Open RAN.

In [7], the authors have studied an Open RAN system and assumed that the core network and RAN in Open RAN are installed in the same data center. The authors performed a theoretical study implementing a mathematical model to solve minimum delay problems efficiently. In [8], the authors worked explicitly on integrating the blockchain technology and network sharing in Open RAN. The paper uses simulations to show that its blockchain-based Open RAN architecture addresses trust and scalability issues. In [9], authors focus on using Open RAN architecture to create high data rate cellular links for drone applications. The idea is to stream applications from drones and use Open RAN controllers to manage the end-to-end communication of their solution. In [10] authors studied the role and implementation of deep learning in Open RAN, and addressed concerns such as energy efficiency, security and scalability of Open RAN.

In [11], the authors present an overview of Open RAN and its capabilities to solve potential networking problems. In [12], the author discusses the estimated cost of $1 Trillion for implementing new generations of networking, specifically 5G. The article argues that opening RAN interfaces and deploying components from multiple vendors will create a competitive culture resulting in a lower product price. Implementing new network generations can take a few years to move from prototyping to a fully implemented solution in cities. There are standards such as 3GPP on networking for vendors and operators to facilitate this task. But these standards have many gaps. As a result, vendors have their internal specifications for implementations of networking architecture and design. For example, they use parameters from their internal specifications for connections to base stations. The goal of Open RAN should be standardizing interfaces to ensure compatibility among products made by different vendors. In addition, Open RAN implementations should be compatible with already implemented networks in any area. For instance, 5G or B5G RAN infrastructure should operate alongside the previous generations. The main reason is that replacing the entire networking infrastructure is costly and demands years of work. Also, not all user devices work with new network generations, so operators may need to provide backward compatibility with older devices (e.g., 4G) in addition to supporting newer networks (e.g., 5G or B5G). This approach has another benefit. It can let users switch between networks if any of them faces temporary problems.

As part of the Open RAN concept, operators should be able to mix components of different vendors. As a result, operators are not obliged to accept all the components from one vendor and can use and mix the best products from different vendors. Hardware disaggregation and virtualization in vRAN will help Open RAN to provide flexible environments that can welcome a multi-vendor architecture.

Open RAN's benefits of commercializing domains with multi-vendor inter-operable products will accelerate innovation. In this exciting innovation journey for telecommunication, AI applications play vital roles.

In [13], authors refer to a use case to discuss the role of AI in RAN disaggregation and virtualization. They consider components of CU (Control Unit), DU (Distributed Unit), and RU (Radio Unit) as multi actors or agents that can use intelligent programs to exchange information and collectively make a decision. The authors study challenges that implementing a multi-agent solution on Open RAN can introduce, such as information sharing and assigning the correct number of agents to a problem. [14] has an overview of RIC in Open RAN, its intelligent controller, and the overall architecture of RIC. In [15] researchers present an open source Open RAN solution which is a closed-loop control of RAN slicing for Open RAN. The solution is top-to-bottom in the POWDER (the Platform for Open Wireless Data-driven Experimental Research [16]) mobile and wireless research platform.

B. OPEN RAN IN ACTION

As mentioned previously, Open RAN has several benefits, such as unlocking operators from single-only vendor options, decreasing cost, and increasing opportunities for innovation. However, these benefits come with some problems, such as end-to-end management of RAN. In a single-only vendor solution, that single vendor is responsible for maintaining and managing the whole system, which is not the case in Open RAN. Open RAN supports the multi-vendor implementation of RAN and requires compatibility between components built by different vendors. Interoperability and reliability in Open RAN demand a global standard that all vendors and operators can use. In 2016 a foundation named XRAN was formed by AT&T, Deutsche Telekom, SK Telecom, Intel Corporation, Texas Instruments, Radisys, Altran, and Stanford University. The XRAN Foundation aimed to create a software-based, modularised architecture for RAN. XRAN was created to standardize user plane silicon and software with open interfaces and logically centralize network intelligence and state. 3GPP has also created many networking standards that vendors and operators refer. Nevertheless, there are still gaps to address, especially regarding Open RAN requirements. That is why a team of members from the six operators, AT&T, China Mobile, Deutsche Telekom, NTT Docomo, and Orange, joined together to work on global specifications and fill the gaps in 3GPP standards. The collaboration is called O-RAN Alliance and started in 2018. The O-RAN Alliance has grown and added more active members from industry and academy involved in its projects.

The O-RAN Alliance group [17] creates standards for Open RAN based on 3GPP and in cooperation with ETSI. It also has partners who help them with implementing solutions. One of the most important partners is O-RAN Software Community [18], an Open Source Software community that works on building Open RAN components and creates frameworks for different use cases.

Other teams that work on solutions for Open RAN are TIP OpenRAN which works on disaggregated and interoperable 2G/3G/4G/5G NR Radio Access Network (RAN) solutions for Open RAN, and ONF (Open Network Foundation) that has created SD-RAN to implement near-real-time controller solutions for Open RAN. SD-RAN is a project under the Open Network Foundation (ONF) consortium hosted by Deutsche Telekom (DT) in Berlin. In [19] authors present a software development kit (SDK) that enables building specialized service-oriented controllers. The SDK has a modular architecture which makes an efficient product for SD-RAN. More and more consultants and software companies are getting involved in this technology as it becomes more evident that Open RAN is the future architecture of telecommunication.

C. ARCHITECTURAL FEATURES

The key factors in the architecture of Open RAN are [20]:

- Disaggregation of RAN Hardware & Software on vendor neutral, 3GPP-based platforms.
- Open interfaces between components (e.g., RU/CU/DU/RIC) following universal specifications.
- Flexibility in separation or aggregation of components (DU, CU, RU, RIC, etc.) during implementation and installation.
- Adoption of new technologies such as Machine Learning and Artificial Intelligence (ML/AI) and Continuous Integration and Deployment (CICD).
- Multi-vendor approach for supplying components.

Open RAN has inherited many of its advantages from C-RAN and vRAN that are enablers of Open RAN.

IV. OPEN RAN ARCHITECTURE AND COMPONENTS

One of the main goals of Open RAN is enabling Operators to run multi-generation networking systems. Therefore, the Open RAN architecture should support both legacy generations and the new ones simultaneously. That is why observing LTE and 5G terminologies on the same Open RAN architectural diagrams. This section discusses terms and components specific to Open RAN in addition to LTE and 5G components and terminology. O-RAN Alliance has prepared and continues to issue specifications for different aspects of Open RAN. The architectural content of this paper is based upon O-RAN Alliance specifications defined in [21]. The overall architecture of the Open RAN and its components are shown in Fig. 3. This section describes these components in the architectural point of view.

A. NON OPEN RAN SPECIFIC COMPONENTS

We start with components that Open RAN has in common with other networking architectures. Some terms are initially from 3GGP TR 21.905 but overwritten by the O-RAN Alliance.

1) NMS (NETWORK MANAGEMENT SYSTEM)

A Network Management System for the O-RU to support legacy Open Fronthaul M-Plane deployments.

2) O-eNB

It is the hardware aspect of a 4G RAN that communicates with Near_RT RIC over E2.

3) gNB

The gNB (Next Generation NodeB) is introduced in 5G and is the succeeding Node of NodeB in 3G and eNB in 4G. The New Radio (NR) RAN has designed two diverse types of gNB, including en-gNB (the 4G Evolved Packet Core (EPC) as the core connected to the 4G LTE base station which is connected to a 5G NR base station) and ng-eNB (the 5G NG core is deployed with connection to the 4G LTE base station) to make NR and LTE compatible with each other [22]. In 5G, gNB includes CU (Central Unit) and DU (Distributed Unit).

4) O-CU, O-DU

In 5G, BBU is split into two functional units, centralized Unit (CU), and Distributed Unit (DU). DU is responsible for real-time L1 and L2 scheduling functions. CU is responsible for non-real-time higher L2 and L3 functions.

5) O-DU (O-RAN DISTRIBUTED UNIT)

A logical node hosting RLC/MAC/High-PHY (Radio Link Control/Media Access Control/High Physical layer) layers based on a lower layer functional split.

6) O-RU (O-RAN RADIO UNIT)

A logical node hosting Low-PHY layer and RF (Radio Frequency) processing based on a lower layer functional

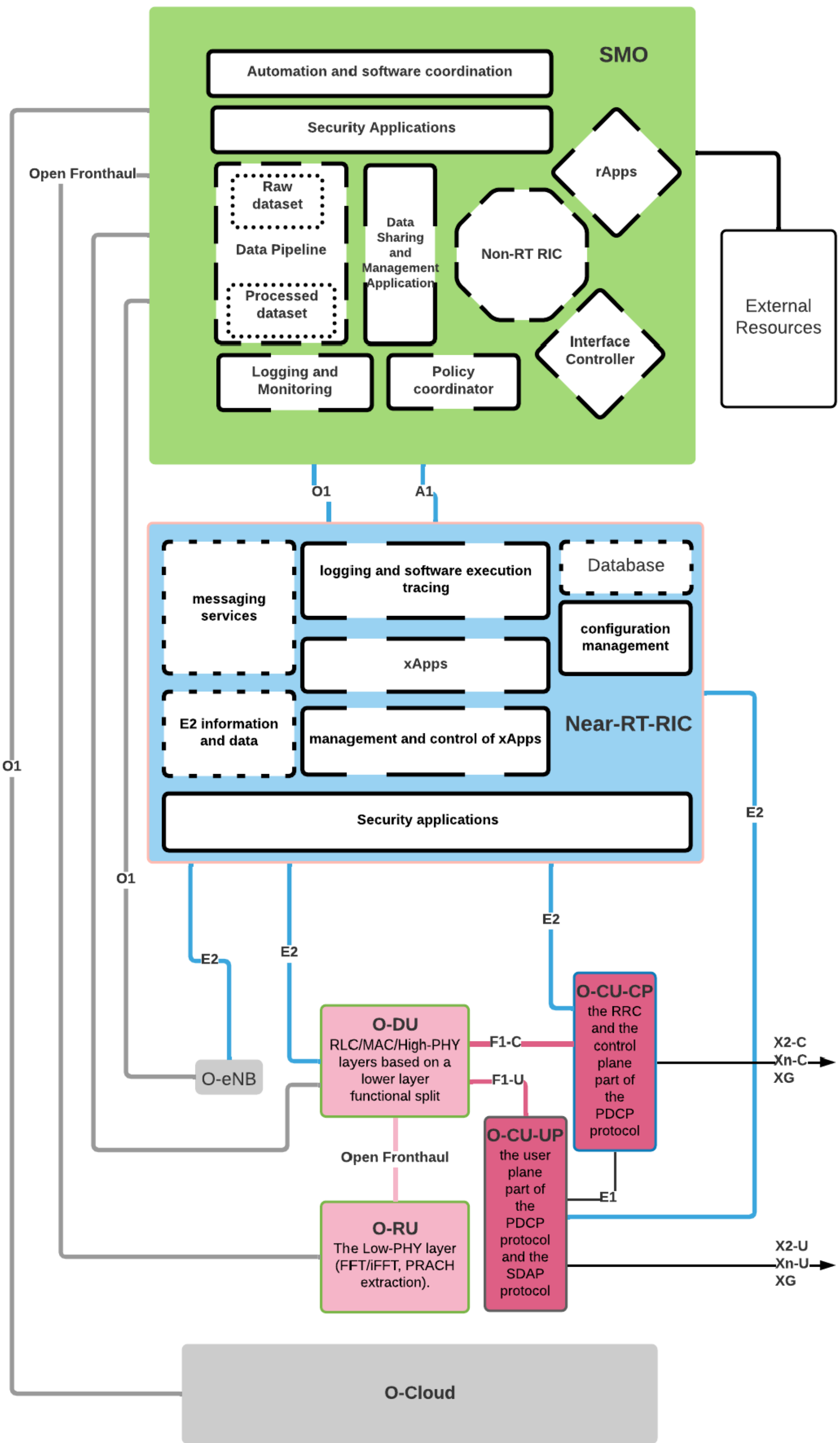

FIGURE 3. Overall architecture of Open RAN.

split. This is similar to 3GPP's "TRP" (Total Radiated Power) or "RRH" (Remote Radio Head) but more specific in including the Low-PHY layer (FFT/iFFT (Fast Fourier Transform/Inverse Fast Fourier Transform), Physical Random-Access Channel extraction).

7) O-CU (O-RAN CENTRALIZED UNIT)

Same as CU is a part of gNB in 5G, and it is responsible for functionalities such as Transferring User Data, Mobility Control, RAN sharing, Positioning, and Session management. CU communicates with DU over the F1 interface. In 3GPP, CU is split into CU-CP and CU-UP.

- **O-CU-CP (O-RAN Central Unit-Control Plane)**: a logical node hosting the Radio Resource Control (RRC) and the control plane part of the Packet Data Convergence Protocol (PDCP) protocol.
- **O-CU-UP (O-RAN Central Unit-User Plane)**: a logical node hosting the user plane part of the PDCP protocol and the SDAP (Service Data Application Protocol) protocol.

B. OPEN RAN SPECIFIC COMPONENTS

This section reviews the architecture with components and terminologies exclusive to Open RAN. The Open RAN architecture comprises two groups of components that are exclusive to Open RAN. The first group is those components that Open RAN originated. The second group is Open RAN components inherited from other network designs but modified to suit its needs.

5G and B5G networking generations promise use cases that require intelligent networking system to proactively manage networking tasks. As a result, the Open RAN includes intelligent and proactive management elements as part of its base framework. The Open RAN specification [23] categorizes applications that facilitate this purpose into three groups, based on their expected latency. The first category is those real-time controlling applications that run on the DUs. The latency for these controllers is Time To Interact (TTI) and is expected to be less than 10ms. The second controller category is near-real-time applications with a latency of less than 1s. The third category is non-real-time intelligent controllers with more than 1s latency.

Given the above expected latencies, the next question will be where is the best place within the Open RAN architecture to execute these applications.

RAN architectures currently include RU and DU to run real-time intelligent applications. The Open RAN specification has added two logical components to the RAN: Near-Real-Time Intelligent Controller (Near-RT RIC) and Non-Real-Time Intelligent Controller (Non-RT-RIC). These components are the primary hosts of near-real-time and non-real-time intelligent applications.

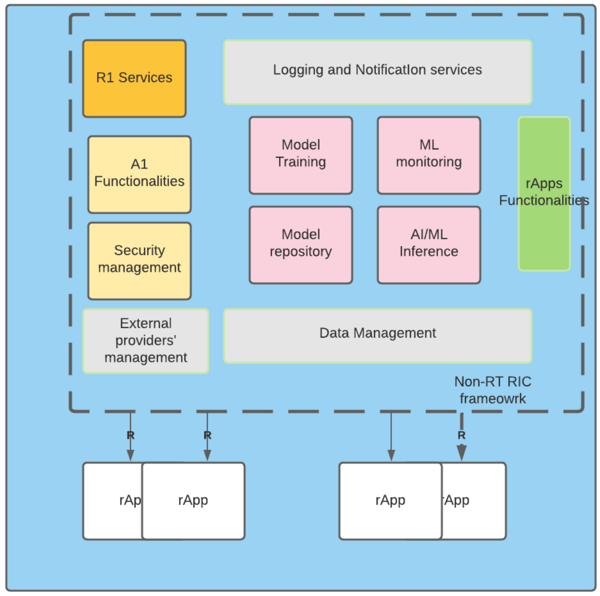

FIGURE 4. Non-RT RIC.

1) NON-RT RIC

Non-Real Time RAN Intelligent Controller, or Non-RT RIC, is a logical function made of many software products within SMO (Service Management and Orchestration). It drives the contents carried across the A1 interface and comprises Non-RT RIC frameworks and Non-RT RIC applications (rApps). Non-RT RIC is responsible for operating "content transfer" via interfaces to Near-RT RIC. Non-RT RIC influences content for the O1 interface and generates *enrichment information* for Non-RT applications. For model monitoring, the ML developer will provide metrics of the model in the form of a contract so Non-RT RIC can log and present those metrics.

2) NON-RT RIC APPS (rApps)

rApps are modular functions that leverage the functionality exposed via the R1 interface. They provide services relative to RAN, such as driving the A1 interface, recommending values and actions applied over O1/O2 interfaces. The rApps functionalities enable non-real-time control and optimization of RAN elements and resources and policy-based guidance to the applications on Near-RT RIC.

3) NON-RT RIC FRAMEWORK

This framework addresses the functionality internal to the SMO. The framework logically terminates in the A1 interface with the Near-RT RIC and is connected to rApps via its R1 interface. The Non-RT RIC Framework functionality within the Non-RT RIC provides ML/AI workflow, including model training, inference, and required updates for rApps. As shown in Fig. 4 Non-RT RIC is responsible for the security and data management of its functionalities.

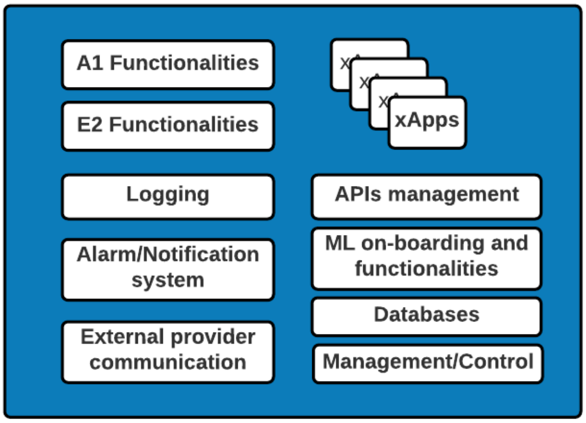

FIGURE 5. Near-RT RIC platform.

4) NEAR-RT RIC

Near Real-Time RAN Intelligent Controller or Near-RT RIC is one of the logical functionalities to control and optimize RAN elements and resources. The control happens via fine-grained data collection and actions over the E2 interface. This component may include AI and ML pipelines, including training, inference, and updates.

Near-RT RIC is primarily responsible for control activities with less than 1s latency. These use cases do not necessarily include real-time use cases. However, to run intelligent applications, Near-RT RIC facilitates many functionalities that are presented in Fig. 5, and some of them are described below:

- Database: The database collects data from UEs and other components of RAN. It will be used as an input for Near-RT applications to decide and take action. Applications also can update or put data into that database if they need to store information.
- Management components: These components are for application management tasks such as onboarding, data sharing, and response collection.
- Messaging and notification component: This component is responsible for alerting or informing operators and vendors on the status of applications and/or platform.
- Logging and monitoring facilities for applications.
- Interface management: This component is added to manage and communicate with interfaces.

5) xApp

xApp is an application designed to run on the near-RT RIC. This application consists of containers with input and outputs. xApps can be provided by an approved vendor. The E2 enables a direct association between the xApp and RAN functionalities.

C. OPEN RAN INTERFACES

Three major groups of interfaces are SMO interfaces, Near-RT RIC interfaces and Nodes interfaces. Table 1 summarises these interfaces with their functionalities and their origin.

TABLE 1. List of interfaces in Open RAN.

Interfaces			
Interface	Components	Origin	Role
A1	Non-RT, Near-RT RIC	ORAN Alliancee	Application deployment, Policy control and management
E2	Near-RT, Node	ORAN Alliance	Communication between edge and Near-RT
Fronthaul	O-RU, O-DU, SMO	ORAN Alliance	Communication between O-RU, and O-DU/SMO
O1	SMO	O-Alliance	Between SMO and other ORAN components
O2	SMO, Cloud	ORAN Alliance	Communication between Core and SMO
E1	CU-CP,CU-UP	3GPP	Communication between CU-CP and CU-UP
X2/Xn/NG	Node components	3GPP	Legacy responsibilities in previous RANs

SMO communicates with other components of the Open RAN via O1. SMO uses O2 to communicate with the cloud and support the execution of functionalities that run on the cloud. However, Non-RT RIC which is a part of SMO uses A1 to send information regarding use cases and EI (Enrichment Information) Jobs to Near-RT RIC.

Near-RT RIC in addition to A1 and O1 for interaction with SMO, uses E2 to communicate with managed elements such as O-CU, O-DU, and O-eNB. Other components that are inherited from previous generations of RAN use the same interfaces as the other RAN architectures. For example, E1 between O-CU-UP and O-CU-CP, or F1 between C-CU and O-DU. Some interfaces such as front-haul between RU and DU is also inherited from previous architecture designs but is modified by being an open interface, now being called open front-haul.

Because Open RAN is a multi generation framework, both 4G and 5G can run on this design. As a result 5G interfaces that help with this multi-generation architecture are also adopted in the Open RAN architecture. For example, X2, and Xn help with interoperability of nodes from both networking generations and NG connects 5G nodes to the core network in a standalone operation.

D. COMMUNICATION DETAILS IN A1

A1 plays a crucial role in Open RAN and communication between Near-RT RIC and Non-RT RIC. Therefore, we dedicate a section to this interface and the life cycle of data transfer via A1.

1) ENRICHMENT INFORMATION

SMO collects information from internal and external resources. This information which is called Enrichment Information (EI), is being used by both SMO functionalities including Non-RT RIC and Near-RT RIC.

2) A1 POLICY

Non-RT RIC uses a declarative policy to lead Near-RT RIC functionalities via A1. In a declarative policy, statements express the goals of the policy but not how to accomplish those goals. This policy is called the A1 policy and is non-persistent. In other words, it cannot survive if Near-RT RIC restarts. Non-RT RIC provides *Enrichment Information* (EI) to Near-RT RIC via A1, using A1 policy.

A1 Enrichment Information is a phrase to address the information which is collected or derived at SMO and Non-RT RIC. Sources of EI can be external or internal Open RAN resources. EI is either not directly available to Near-RT RIC or cannot be derived inside Near-RT RIC from network data due to processing or storage constraints. Near-RT RIC requests for delivering this information by using formal statements, called *EI jobs* created, modified, and deleted by Non-RT RIC.

3) ENRICHMENT INFORMATION FUNCTIONS

Enrichment Information (EI) functions are functions that generate and manage the delivery of the A1 EI. This function is also responsible for publishing Ids for different types of EI as EiTypeIds. Fig. 6 presents the lifecycle of EI jobs. It starts with *Registration* when Non-RT RIC generates a new EiJob from collected input information by SMO. Then in the stage of *Discovery* Near-RT RIC discovers new EI jobs and requests the detailed description of the EI. Near-RT RIC can send a request for EI to Non-RT RIC. Non-RT RIC in this stage, the stage of *Request* generates an EI Job and an Id named EiJobId that will be assigned to the corresponding the EI Job. In the last stage, the stage of *Delivery*, Non-RT RIC sets up the connection and delivers the EI job via A1 on a push delivery basis.

a: SMO

SMO stands for Service Management and Orchestration Framework. It is not specific to Open RAN, but its role is customized in Open RAN.

The 3GPP standard number TS 28.533 has covered many aspects of management and orchestration in networking. It covers components such as operation and notification, entity-specific information management, performance management, and any combination of these components. In Open RAN, SMO includes Non-RT RIC, which we discussed before. In addition, SMO is responsible for any operation and management tasks related to O1, A1, R1, O2, and any communication with external resources. SMO also offers APIs (Application Programming Interfaces) for DU, CU, RU, and RIC configurations. The interaction between SMO and O-Cloud, which hosts RAN functionalities such as RU, DU, and CU, will continuously manage and support the ongoing activities of Open RAN. In the design and implementation, the large capacity for collecting, storing, and processing a large amount of data on SMO is a primary requirement.

SMO in Open RAN can reduce security risks and improve the management of applications and their versions by being the source of truth for both rApps and xApps. It can be responsible for onboarding applications and keeping a catalog of deployed applications and their status. In this way, if any

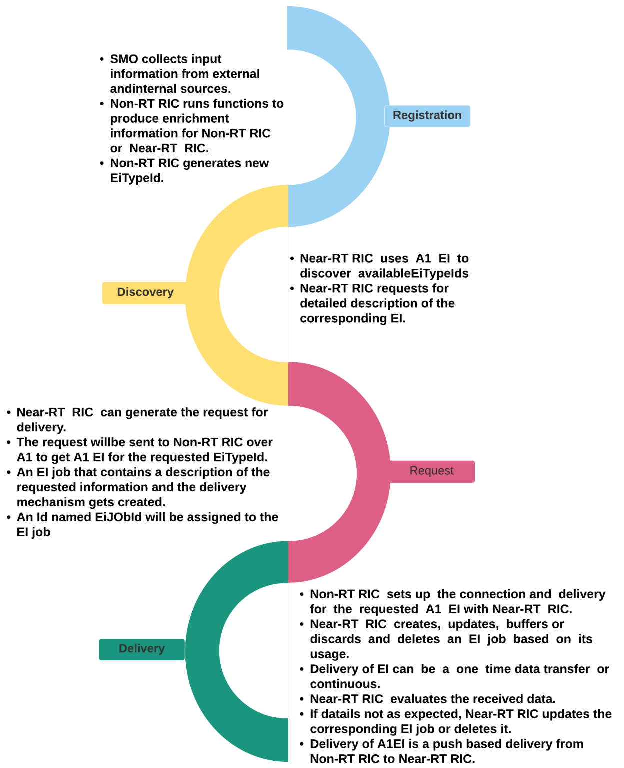

FIGURE 6. EI job lifecycle.

unknown or less known application is running on Near-RT RIC, it can send an alarm to operators. Nevertheless, implementing this logic needs enough maturity and capacity in SMO and a tendency from Operators to create the catalog of ML/AI applications in SMO. Otherwise, Near-RT RIC should maintain a trustworthy catalog of xApps. In any case, the overall system should be consistent on where to save the details of already evaluated applications that run on Near-RT RIC hosts.

b: O-CLOUD

O-Cloud is a cloud computing platform that is inherited from the concept of virtualization and cloud-native RAN.

O-Cloud is a collection of physical infrastructure nodes that host the O-RAN functionalities (such as Near-RT RIC, O-CU-CP, O-CU-UP, and O-DU), the management and orchestration related functions, and the underlying software components (such as Operating System, Virtual Machine Monitor, and Container Runtime). O-Cloud architecture and deployment architecture are affected by the use case, location, and the operators' preference. The current specification of O-Cloud approves both containerization and NFV (Network functions virtualization), but operators might decide to accept only one of the approaches from vendors in the future, to reduce the complexity of management and operational tasks.

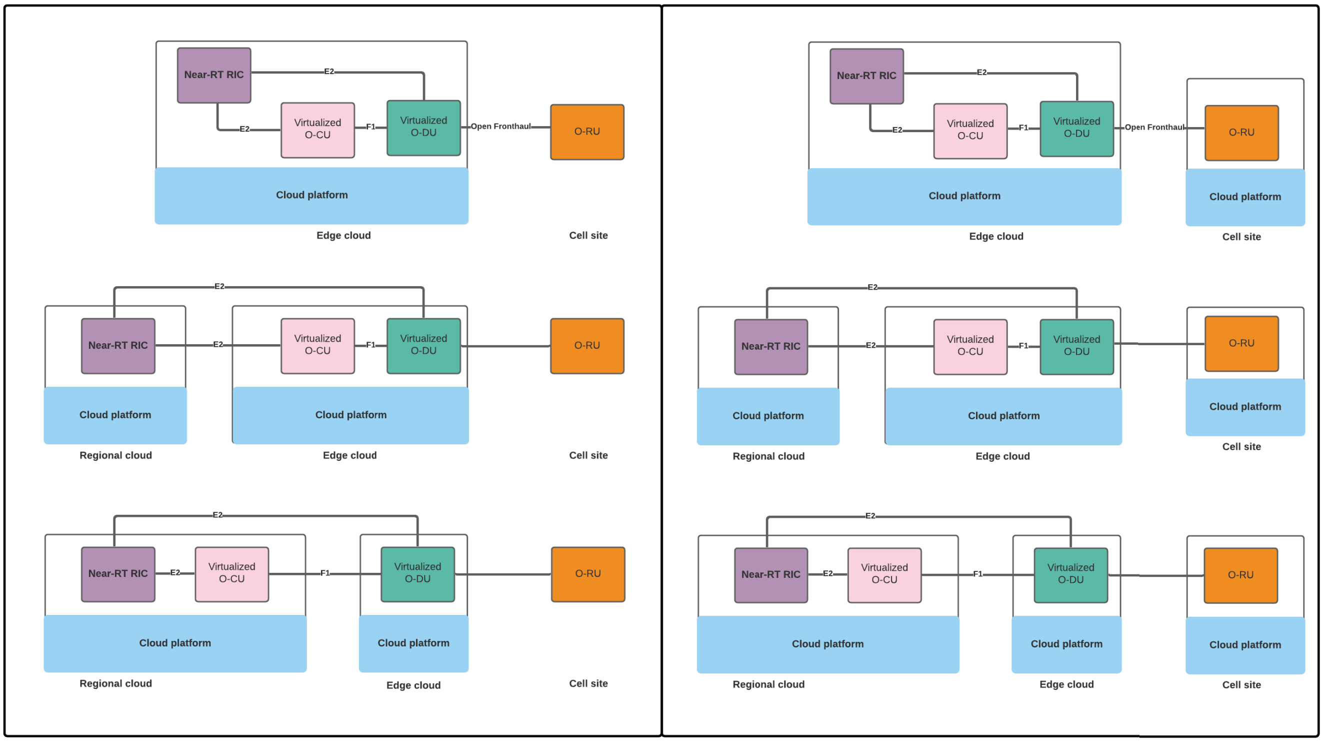

FIGURE 7. **Deployment scenarios of O-Cloud.**

Networking use cases define which combination of components should run on the same cloud platform. Fig. 7 presents different combinations of deploying Open RAN components. The figure shows all the different combinations of deploying RAN components on cloud environments.

V. ARTIFICIAL INTELLIGENCE AND OPEN RAN

In this section, we focus on how intelligent applications can improve the efficiency of RAN, specifically with the architecture of Open RAN. This section starts with a brief review of previous implementations of AI algorithms in telecommunication. The, following the O-RAN Alliance specifications [23], we categorize intelligent applications that will control RAN in Open RAN into three categories based on their expected latency. These categories are real-time controllers that run on DU, near-real-time applications, and non-real-time applications. [24] has reviewed the employment of AI in Open RAN based on different network levels. In this paper, we look at applications of AI in terms of latency requirements with more details about their background, implementation in production, and challenges.

A. AI IN TELECOMMUNICATION

ML/AI algorithms, in general, are being categorized into supervised, unsupervised, and reinforcement learning. Fig. 8 has summarised common telecommunication problems and ML/AI algorithms in each category that researchers have used to solve those problems.

1) SUPERVISED LEARNING

Supervised learning algorithms are AI algorithms that take labeled data as input for training. In this learning method, humans control what the algorithm learns. Nevertheless, this learning method needs a large amount of correctly labeled data which leads to the human capital cost and any potential human error in labeling. Nevertheless, some common supervised algorithms have been used in telecommunication extensively. In telecommunication, Support Vector Machine (SVM) is used in security and time-series forecasting, and Naive Bayes for intrusion detection, TCP (Transmission Control Protocol) enhancement, DDOS (Distributed Denial Of Service) attack, and localization. Algorithms such as Logistic regression, Random Forest, and decision trees have been used to solve security, intrusion, and DDOS attacks problems.

2) UNSUPERVISED LEARNING

Unsupervised learning is a learning method that does not require labeled input data. As a result, it reduces the cost of pre-processing steps in machine learning. However, the range of problems that this learning method can solve is minimal compared to the supervised learning method. Some of the unsupervised learning methods that have been used in telecommunication are clustering algorithms for sensor networks and locating controllers in Software Defined Network (SDN) systems or data mining algorithms to ensure the reliability and accuracy of solutions. Singular Value

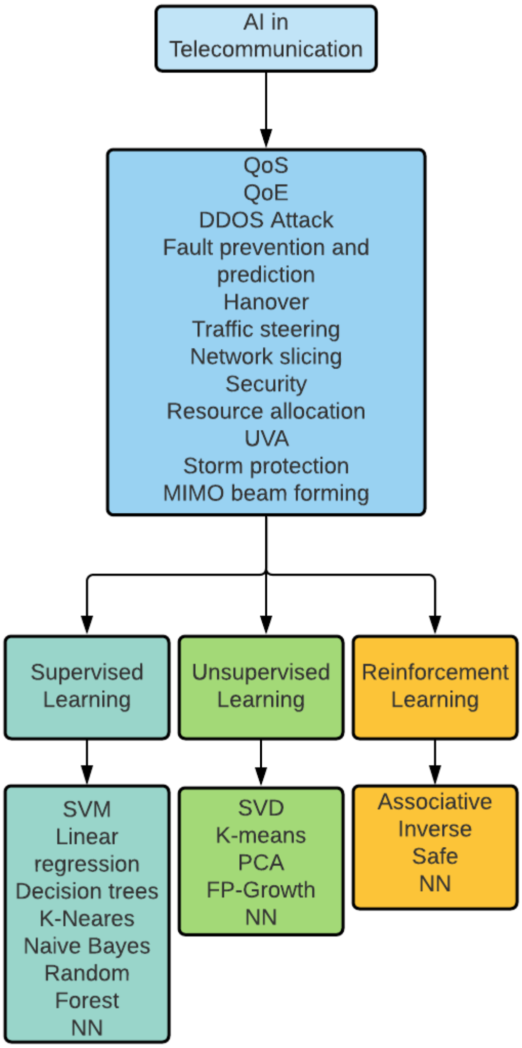

FIGURE 8. AI in telecommunication.

Decomposition (SVD) and Principal Component Analysis (PCA) that are common algorithms in data processing and run as a step in different machine learning pipelines belong to this category.

3) REINFORCEMENT LEARNING

Reinforcement learning (RL) is a method that an agent learns patterns and decision-making strategies by interacting with its environment. It does not need labeled data like supervised learning. But it requires accurate environment modeling and sometimes more iteration than supervised learning to achieve the required accuracy. RL has been used in telecommunication for many use cases such as packet routing, beamforming, Hand Over-optimization, and other use cases.

Neural Network (NN) algorithms have advanced these three learning methods. In telecommunication, any of these learning methods or a combination of them can be used to solve problems or improve the efficiency of tasks.

TABLE 2. ML/AI in telecommunication.

Learning method	Algorithm	Sample papers
Supervised Learning	SVM security and time-series	[25],[26],[27],[28]
	Logistic regression	[29],[30],[31]
	Naive Bayes	[32] ,[33],[34],[35]
	Decision trees	[36],[37],[38]
	K-Nearest neighbor	[39]
	Neural networks	[40],[41]
	Random Forest	[42],[43]
—	—	—
Unsupervised Learning	Clustering algorithms	[44],[45],[46],[47]
	Data mining algorithms	[48],[49]
—	—	—
Reinforcement Learning	Associative, Inverse, Safe	[50],[51],[52],[53],[54],[55]

Table 2 has summarised learning methods with some of the research papers on each topic.

B. RUNNING ML/AI ON OpenRAN

ML/AI algorithms can solve many problems in a telecommunication system, including the RAN. Fig. 10 shows some telecommunication problems that AI algorithms can solve. Intelligent functionalities that control RAN can be classified based on their learning approach, such as supervised or unsupervised learning. They can also be categorized as real-time applications with latency less than 10ms, Near-Real-Time (Near-RT) intelligent applications that should respond in less than 1s, and Non-Real-Time (Non-RT) intelligent applications that can have a latency of more than 1s. Real-time MAC and PHY layer intelligent applications run on DU, so we call them DU applications. Fig. 9 presents these three intelligent controllers. Following Open RAN specifications, we call Near-RT and Non-RT intelligent applications Near-RT RIC and Non-RT RIC, respectively.

One of the most important parts of any intelligent system is its data pipeline. Collecting, processing, and passing data to applications in a secure, robust, reliable, and efficient is critical. The quality of output from machine learning models and analysis programs is highly dependent on the quality of data pipelines. Therefore, we review data-related applications as part of intelligent controllers. In the following paragraphs, we start with real-time applications and move to the other two groups of RAN Intelligent Controllers (RIC). It is worth mentioning that solving many problems needs collaboration among applications from different families. Also, Non-RT RIC and SMO might include applications that respond in less than 1s, or training algorithms can happen on a Near-RT RIC host to meet some KPIs, although they usually take more than 1s to complete.

C. DU REAL-TIME INTELLIGENT CONTROLLERS

Running intelligent applications close to UEs and on DU has started before Open RAN. Light ML algorithms and small size but powerful processors have shifted the expectation for ML/AI applications. When an *application latency* is defined

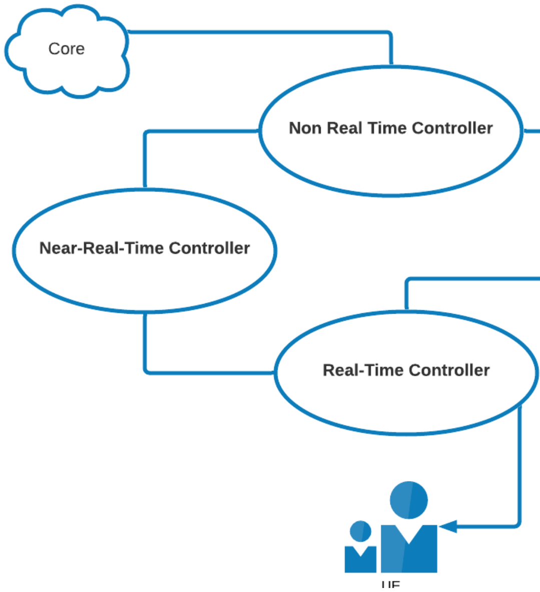

FIGURE 9. Three intelligent controllers in Open RAN design.

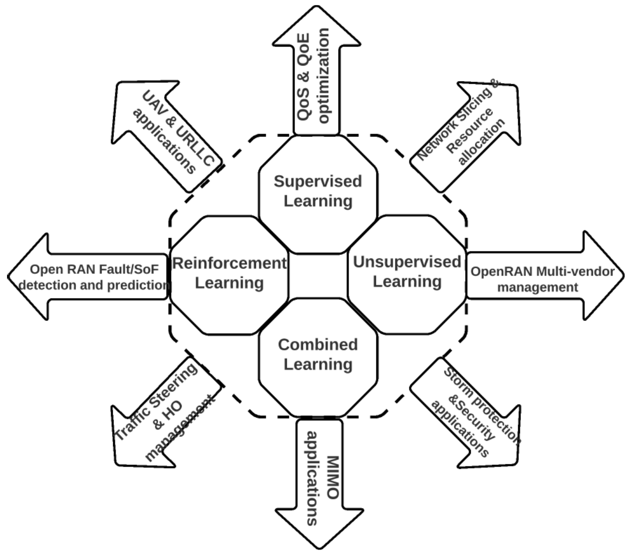

FIGURE 10. ML/AI in Open RAN.

as real-time, the program should run on components as close as possible to users. DU is one of the hosts for real-time applications in wireless networks. DU is responsible for PHY and MAC layer scheduling functions.

1) MAC LAYER FUNCTIONS

LTE introduced MAC schedulers and transferred them to the next generation, 5G. The MAC Scheduler assigns resources to user devices and ensures QoS requirements are met. ML/AI has already started to help with MAC scheduler tasks such as Link Adaptation, Massive MIMO, and multi-user MIMO. An industrial project [56] focused on building a cognitive MAC layer to predict UE's mobility patterns. After running experiments, they reported that MAC layer general tasks and model predictions compete for resources. As a consequence, meeting the 10ms time limit in response becomes challenging.

In [57] a real-time application focused on predicting encrypted packets in video streaming use cases. They combined an unsupervised clustering model with an adaptive classification approach to classify frames and used a time series forecasting algorithm to predict the streaming frames.

In general, considering the MAC scheduler's tasks and the number of ML/AI research works on those topics, DU will play an indispensable role in running intelligent applications on Open RAN.

2) URLLC APPLICATIONS

Ultra-Reliable Low Latency Communication applications (URLLC), have just started their momentum and will play a key role in our near future cities. This category includes use cases such as V2X (Vehicle to Everything), mission-critical applications, smart grid, and tactile internet. DU is essential for URLLC applications by running real-time networking functions on MAC and PHY layers. However, Near-RT RIC might also help for the management and implementation of collective wisdom solutions such as federated learning.

D. NEAR-RT RIC

This logical component is added to the RAN and is responsible for controller activities that need to respond in less than 1s. Near RT RIC is connected to Non-RT RIC to access data collected from internal and external resources. It is also connected to CU and DU that are closer to UEs.

One of the considerable benefits of Near-RT RIC is helping engineers and researchers have a universally agreed platform for running many edge applications or federated learning programs.

1) DYNAMIC HO MANAGEMENT

To optimize the handover (HO) process and perform near-real-time optimization, we need ML models which run on Near-RT RIC. In this solution, the xAPP, an AI model, will monitor the device-specific mobility, predict or detect unexpected HO events, and generate HO sequences to prevent anomalies. One of them is Vehicle to Everything, known as V2X. It is about communication between a vehicle and surrounding items that can receive or send any information from or to the vehicle. The main goal of this topic is to improve road safety and traffic efficiency and reduce energy consumption. The 3GPP standards include standards regarding V2X and V2V (Vehicle to Vehicle). Also, IEEE 802.11 has an amendment, IEEE 802.11p, for wireless access in vehicular environments. Using Open RAN on top of these standards makes some of the goals for V2X communication achievable. The IEEE standard of 802.11p supports direct communication between vehicles and their environment. On the other hand, 3GPP covers cellular V2X (C-V2X) in communication, specifically for 5G. C-V2X includes

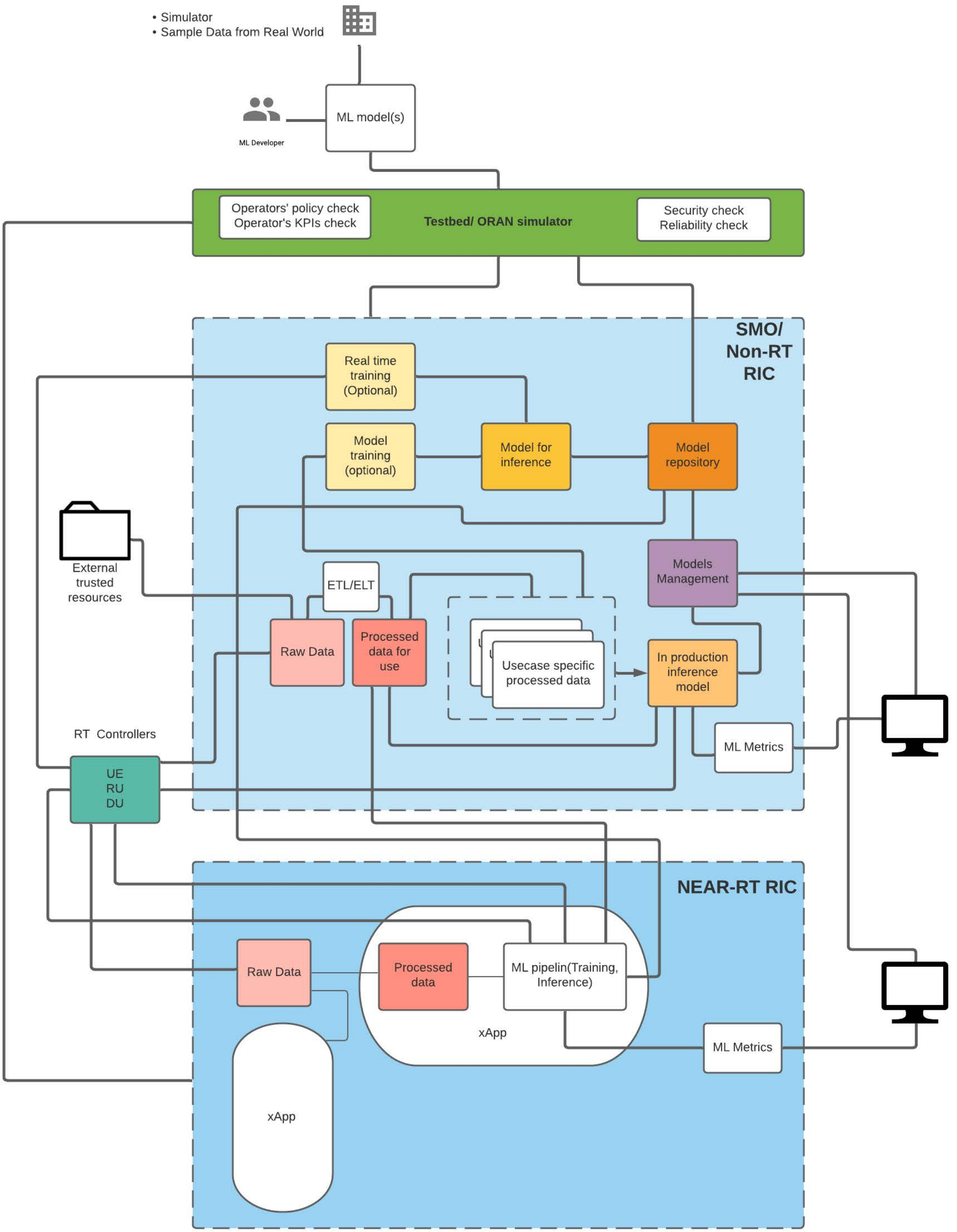

FIGURE 11. Overall ML/AI workflow.

both Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I), as well as Vehicle-to-Network (V2N) [58]. In V2X, the communication happens between the V2X application server (V2X AS) and the V2X device attached to the vehicle. The V2X AS handles services such as data delivery data V2X devices. The application also provisions the 5G core and the V2X device with parameters. At the sub-optimal HO, some anomalies might happen, such as short stay, ping-pong, and the remote cell. An xAPP should update the database maintained by V2X AS to improve the V2X user experience. This use case can run on Near-RT RIC as an xAPP. Non-RT RIC can also support this use case which we will address later in V-E4.

2) TRAFFIC STEERING

Near-RT RIC can constantly monitor users and ensure KPIs of QoE are met. Running traffic steering relies on the historical data of available cells and networks. Near-RT RIC can access this information from SMO and Non-RT RIC or if the resources are not limited, use the same host to collect and use this data. BY using this information, it can optimize UEs Received Signal Strength(RSS) and energy consumption of batteries.

In some papers such as [59]–[62], the researchers suggested a reinforcement learning algorithm that runs on UEs and, by learning from the past experiences of the device in an environment, improves its QoS and reduces its energy consummations. This approach that suggests running the model on UEs has some problems. Firstly, it threatens the network system's security or is unavailable to all users. Secondly, the underlying system and the networking system can change dynamically. Updating the model running on user devices is not always possible. A solution can be running a model as a xApp on the Near RT RIC platform, using AI algorithms to improve UEs QoS and QoE. It can also be used as part of a federated learning solution among UEs so they continuously will be updated on the current status of their environment or the latest version of the model.

E. SMO AND NON-REAL-TIME CONTROLLER

This section explores applications that do not necessarily need a near real-time or real-time response. However, in some scenarios, these applications collaborate with Near RT RIC or support it to solve some problems.

1) DATA GOVERNANCE AND ANALYSIS

Data Governance and analysis are crucial tasks for making intelligent decisions in any business. We discuss these applications in the category of Non-RT controllers because SMO is where data from all external and internal resources are collected. Therefore, SMO and Non-RT RIC platforms can implement all sorts of data-related logics and tasks, regardless of latency constraints. Nevertheless, the other two categories of ran intelligent controllers also can include such applications.

To improve QoE, operators need to understand and monitor their users in the most intelligent way. Data analysis and visualization are systematic approaches for this purpose. Engineers and scientists collect, store and process data to prepare it for analysis or ML/AI applications. Handling data pipelines securely and efficiently while complying with data governance rules and standards is crucial in Non-RT RIC and SMO.

Without a reliable data pipeline achieving a reliable, scalable, and intelligent networking system is impossible.

2) MANAGEMENT AND ORCHESTRATION

SMO is responsible for the management and orchestration of the Open RAN environment. However, managing and orchestrating an environment such as Open RAN, with different stakeholders, various components, multi-vendor solutions, a large number of parameters, and many different data flows to or from components, is complicated and challenging. Communications between elements of Open RAN are potential targets of malicious actors. Data collection and processing from internal and external sources require continuous quality and safety policies. Incident prevention and quick response to alerts without ML/AI solutions are almost impossible. End-to-end network slicing and resource allocation, especially in dense urban areas, are complex challenges for operators to overcome. These problems are not new to researchers.

[63] used SDN principles to orchestrate and coordinate resources in a 5G RAN infrastructure. Another published work [64] proposed a functional split orchestration scheme for running 5G on Cloud-RAN. The authors evaluated the results from running the proposed design on a 5G experimental prototype and claimed the suggested design reduces energy consumption and the deployment cost.

3) NETWORK STORM PROTECTION

Storm protection and control prevent LAN interfaces from interruptions caused by a broadcast storm. A broadcast storm happens when network packets flood the subnet, creating excessive traffic and degrading network performance [65]. This action degrades the network performance and interrupts crucial communications such as health-related applications. As technology advances, more mission-critical use cases use wireless networking, and any interruption can threaten someone's life. Storms can be a consequence of errors in network design or installation. However, there can be some intentional attacks causing this disturbance. In these scenarios, an attacker usually finds vulnerable devices connected to the internet and manipulates them by aggressively sending many packets to create a storm and cause an outage of an extensive network. At the moment, the most common reaction is rejecting both benign and malicious services. However, intelligent algorithms can detect the compromised device, preventing malicious actions while serving benign service requests. A good solution should include both detection and mitigation capabilities. Although detection and prevention of the applications can run on Near-RT RIC, some solutions can better run on SMO and Non-RT RIC. Information such as device type, International Mobile Equipment Identity (IMEI), and Public Land Mobile Networks (PLMN) is available on SMO. SMO also has the scheme of the attack. Non-RT RIC can feed Near-RT with this data and help it detect and stop the threat on edge. In this way, Non-RT RIC will do heavy computation tasks and send input data to near-RT xApps and help them make more intelligent and faster actions.

4) DYNAMIC HAND OVER (HO) MANAGEMENT

Efficient Handover management in networking has always been a challenging task, but new use cases such as V2X make this challenge an urgent research problem. In general, The HO sequences happen mainly based on Neighborhood Relation Tables (NRTs), which are maintained by xNB, so devices themselves do not have much information to optimize them.

ML/AI applications can optimize HO in different ways, such as navigation and radio statistics history. In this solution, data gets sent from devices through O1 to SMO, where the Non-RT RIC applications run, and databases collect data. An ML/AI application can find anomalies and suggest resolutions accordingly. NRTs are preferred to be UE-specific. In this solution, This database will be an input for the AI algorithm to learn and find anomalies. The V2X AS is responsible for data maintenance and providing information to AI algorithms. The algorithm's output should be finding the anomalies and the resolutions accordingly. [66] worked on 3G/4G cellular networks to maintain QoE for a UE that sends RSRP (Reference Signals Received Power) and RSRQ (Reference Signal Received Quality) of its serving cell. If the received power and quality are less than a defined threshold, the serving eNB selects another target cell and triggers the HO process by sending the message to MME (Mobile Management Entity), and consequently, a message will be sent from MME to NB if enough resources are available. If eNB gets the message from MME, it switches the target cell to the new one, updates the table, and releases resources from the old cell. The paper used regression and neural networks to predict handovers based on their data analytic results. Another research [67] used ML/AI in SON to identify faulty cells to improve faulty cells and reduce packet loss. They used the KNN algorithm to detect abnormal access points (AP), classify access points, and find HO delay. They compared their results with other algorithms such as SVM, Random forest, and K-mean. Another research team used RNN (Recurrent Neural Network) and LSTM (Long Short Term Memory) to learn latency and cost associated with service requests. They used a real dataset of real-world vehicle movements to create a simulation environment for training their algorithm. Their goal was optimizing vehicular fog computing by HO optimization [68]. In another research [69], authors focused on anomaly detection using a semi-supervised algorithm. They used this method to detect two types of cells that cause abnormal behavior in a system. The two types were the sleeping cell caused by very low user requests and the too busy cells caused by too many requests that cause more demand for resource allocation. In [70], researchers proposed a method that combines fuzzy logic and multiple attribute decision algorithm (MADM) to use historical data and find the optimum target. As a result, the generated model triggers the HO at the right time rather than being late or dismissed. In sum, using datasets from real-world scenarios, combined with one of the supervised or unsupervised methods, can help create an AI model that can detect abnormal cells or anomalies in the system. This detection can trigger further actions or be transferred to the Near-RT RIC platform to find anomalies in near real-time.

5) UAV APPLICATIONS

An Uncrewed Aerial Vehicle (UAV) is a component of an Uncrewed Aerial System (UAS) that can fly and move without any human pilot on board. There are many benefits to using UAS in telecommunication. RAN architectures can benefit from UAS and UAV to overcome data traffics' high load and variability. Telecommunication systems can use them as additional stations, acting either as a base (BS) or relay stations (RS), especially in unexpected, natural disasters, or special events. UAVs can operate as BS or RS to increase the coverage area, balance traffic load, and enhance network capacity. The benefits of using UAVs are not limited to their mobility and flexibility in geolocation. The flexibility in deploying various products and providing line-of-site connectivity are other benefits of this technology for telecommunication, specifically RAN. Although UAS individually has had significant progress in providing services to users, there are still many telecommunication challenges to resolve that ML/AI can come to the rescue. One scenario that needs HO is Radio Resource Allocation (RRA). Based on parameters and features such as data traffic rate, latency tolerance, and reliability that define user group, an ML/AI program can find a pattern and efficiently provide the best suggestion on resource allocation and other decision-making tasks. The selection of ML/AI algorithms depends on the nature of the business question. Predicting future behavior is better executed by a supervised learning algorithm that can learn very well from the past and make a model that can more accurately predict the subsequent requests. At the same time, an online decision in a dynamic environment might suit a Reinforcement learning algorithm trained in similar dynamic environments and can take quick actions in response to changes, incoming requests, or events. When it comes to UAVs, we cannot underestimate the impact of weather conditions on their operation. ML/AI algorithms can combine weather conditions and telecommunication parameters to generate models that can operate UAVs better than a manual controller or a hardcoded program [71], [72].

6) TRAFFIC STEERING

The increasing number of telecommunication users and imbalances in traffic caused by their UEs or various bandwidths available to users have challenged telecommunication operators. The 3GPP Self-Organizing Network (SON) function includes Mobility Load Balancing (MLB). MLB balances the load by optimizing the handover triggers and handover decisions using load information shared between neighboring cells. However, it treats different user groups equally. In addition, since 5G, the networking system can support different combinations of access technologies such as LTE (Licensed band), NR (licensed band), NR_U (unlicensed band), and WI-FI (unlicensed band). As a result, finding a solution to prioritize and provide services according to the request type becomes inevitable. Traffic Steering addresses this challenge. Based on the traffic type (e.g., HTTPS (Hypertext Transfer Protocol Secure) or HTTP), user membership profile, or the priority level of the service request, and many other factors, the request can be served by different operational activities or a combination of access technologies.

Traffic management policies with flexible configuration can help proactively manage user traffic across different technologies. SMO can collect information from UEs and, in collaboration with Non-RT RIC, can monitor user experience by measuring the UE performance and resource utilization on the cell level. Where the service requirements are not acceptable, it can locate the target cell and implement a solution such as switching cells for the affected user or offloading that cell, or increasing bandwidth if it is possible. Moreover, in multi-access systems, traffic steering can happen between different environments. Non-RT RIC can create traffic management policies specific for any UE and based on the priority of cells for each UE. Non-RT RIC sends policies to the Near-RT RIC to enforce the radio resource control. In addition, Non-RT RIC, by exploring the historical data, can extract radio fingerprint enrichment information. By monitoring the inter-frequency measurement and passing information to Near-RT RIC, the system can predict the inter-frequency measurement. This prediction can reduce the unnecessary inter-frequencies to boost traffic optimization and network performance. There are many academic works on MLB algorithms to improve QoE. Although those works consider MLB as a solution, their logic and algorithms belong to Non-RT RIC applications. The shortcomings of this solution can be used as challenges for improving QoE. For example, authors in a research [73], after studying the previously suggested algorithms for MLB and addressing the unfairness of those algorithms, provided a QoE-aware algorithm for LTE systems. They use a fuzzy logic controller to tune handover parameters to reduce QoE differences across cells and services. They argue that this method will provide a fair service across all users. They validated their algorithm in a dynamic system-level simulator of a macro-cellular LTE scenario. In [34], authors exclusively studied data analytics with Radio Access Technology (RAT) selection scheme and discussed acquiring contextual information and minimizing control signaling exchange. As we mentioned in Near-RT RIC, monitoring users and checking KPIs (Key Performance Indicator) can be executed by Near-RT RIC.

7) OPTIMIZATION OF MASSIVE MIMO BEAM FORMING

Massive multiple-input multiple-output (MIMO) wireless communications refers to the concept of implementing a large number of antennas in a cellular base station [74]. [75] describes the benefits of using MIMO to increase the capacity on the scale of ten times or more and improve the radiated energy at the same time. MIMO increases energy efficiency due to being able to focus with sharpness into small regions in space and to reduce latency in communication with user devices. As 5G and B5G are in the range of mmWave and THz and are the critical path for improving networking communications, the importance of MIMO becomes apparent. One of many papers in this area is [76] which proposes an ML/AI influenced design for Ultra Massive MIMO and intelligent surfaces. Although many research works are on MIMO and intelligent surfaces, designing an end-to-end solution from core networking to user devices is still a massive challenge for researchers and operators to overcome. In this use case, AI can help with the optimization of parameters and locations.

8) NETWORK SLICING

Network slicing has been a topic of interest for improving communication performance for many years [77], [78]. Because of technology advancements and making concepts such as NFV (Network Functions Virtualization) and SDN (Software Defined Networking) applicable, network slicing became a key role player for 5G and B5G.

"Network slicing is a paradigm where logical networks/partitions are created, with appropriate isolation, resources, and optimized topology to serve a purpose or service category (e.g., use case/traffic category, or for MNO internal reasons) or customers (a logical system created on-demand)" [79]. A network slice is a virtual network with a group of allocated services over a shared network infrastructure. 3GPP has standardized four different service types: eMBB (enhanced Mobile Broadband), URLLC, mMTC, and V2X [80].

However, network slicing is a challenging task. Services which send requests to RAN for resources are dynamic and various. They are different in terms of accepted latency, required computation, time duration, and priority in terms of emergency. A utility request related to autonomous vehicle alarms can tolerate much less latency than a device reporting the customer traffic into a shopping center. Fortunately, machine learning algorithms can be used practically in network slicing and help manage the network's life cycle.

The provisioning of network slicing includes the four phases: preparation, commissioning, operation, and decommissioning. The NSI/NSSI (Network Slice Instance/Network Slice Sub-net Instance) provisioning operations include: Create an NSI/NSSI, Activate an NSI/NSSI, De-active an NSI/NSSI, Modify an NSI/NSSI, Terminate an NSI/NSSI. Some of important use cases of network Slicing are:

1) RAN Slice SLA (Service Level Agreement) Assurance: In this use case, the flexibility of Open RAN

combined with network slicing will allow multiple vendors to request resources from RAN and serve their clients with different applications.
2) NSSI (Network Slice Subnet Instances) Resource Allocation Optimization. ML/AI algorithms and methods have been used for these scenarios and implemented in Open RAN.

[81] presented a unified model which included the network slices, data, and slice managers run by ML, communicating with SDN and the rest of the network over APIs. In Open RAN, the data and slice manager applications can be deployed as xApps on the Near RT RIC platform but can also be deployed on Non-RT RIC. In another research, [82] the authors discussed the challenges that network slicing faces, such as Heterogeneous QoS (Quality of Service) requirements of various services, the fluctuating status of networks, overhead costs caused by network slicing. The authors proposed a network slicing framework and a neural network algorithm to predict the resource allocation problem. The authors of [82] suggested a multi-layer control level for network slicing, such as gNodeB-level and package scheduling level. However, as data is scarce, they suggest using transfer learning to overcome this challenge and generate synthetic data to train and build highly accurate models.

One of the crucial tasks in network slicing is to allocate resources adequately based on the plausible usages in the region. For example, IoT devices might need resources during peak times or events, such as some sport matches that increase the number of requests considerably in a short period, demand for intelligent methods for prediction, and allocation of required resources.

Open challenges in network slicing include NSSI resource allocation, multi-vendor NS, indoor positioning, congestion prediction and management, IIoT (Industrial Internet of Things) optimization, and dynamic spectrum sharing.

Finding optimum solutions for each topic, especially for running 5G and B5G applications on Open RAN, is a complex but crucial task. Smart grids are also another new technology that can improve the RAN's network slicing and resource allocation policies. In [83] the authors describe the use of smart grid technology in 5G and B5G. The integration of AI application with IoT and smart grids in 5G and B5G, can help with the resource allocation and network slicing as well.

Use cases that we discussed are not the only potential opportunities for AI application to the Open RAN and telecommunication.

VI. MLOps IN OPEN RAN

The phrase MLOps stands for Machine Learning in Operations and refers to the efficient deployment and application of machine learning solutions in production. A ML/AI pipeline in production should be scalable, reproducible, and maintainable. Moreover, ML/AI pipeline should have a robust monitoring and versioning system for data, models, and features. In MLOps, an artificial orchestrator automates and maintains these principles. Engineers from open source communities to major cloud providers have created some frameworks and products to help data scientists run their ML/AI solutions efficiently in production based on MLOps principles. Fig. 11 shows the main concept and workflow of ML applications in production. However, these solutions are for general purpose applications where ML/AI pipelines can run on a protected data center at any location. Nevertheless, Open RAN is a new world to explore. Near-real-time and real-time ML/AI solutions are less likely to run on the currently available data centers. Low latency applications that need to be close to user devices cannot run on traditional cloud and data centers. They need to run on edge devices close to base stations and UEs (User Equipment) but have limited resources for computation and data-related activities. Also, executing all the pipeline steps on one platform might not be possible, and we might need to split pipelines to run on Non-RT RIC and Near-RT RIC platforms. Therefore, maintaining the performance of ML/AI pipelines and maintaining security in communications become inevitable challenges to manage.

We continue this section by studying commonly used ML/AI pipelines required in production which can be used in Open RAN. This section does not include different scenarios for data pipelines. Data pipelines which include data storage, data processing, and data privacy procedures, might require a large amount of memory or processing resources and can affect the location and the strategy of data pipelines. They can affect or be affected by the corresponding ML pipeline. O-RAN Alliance has published a specification on AI/ML workflow [23]. It includes multiple deployment scenarios. However, this paper covers standard deployment practices in addition to practical scenarios such as the need to run multiple versions of a model in Near-RT RIC that the specification does not cover. Fig. 12, 13, 14, and 15 present some of common workflows in MLOps that are discussed in this section.

A. SINGLE DATA SINGLE MODEL

After selecting the algorithm and defining the data source in this pipeline, the algorithm will be trained on the collected and processed dataset. The first time training of the model is usually on batch datasets, even if the inference and prediction run on a data stream. For example, a defect detection model can be trained on batches of defective items and then receive a stream of scanned images for identifying the problem on a real-time basis. There are exceptions for this practice, such as reinforcement learning algorithms that run in a simulated environment. The output of training algorithms is a generated model. This model will be deployed to production, usually as a container. The container accepts inputs based on the model requirements. Therefore the production pipeline includes a model with input data from stored data or as a stream from an external source. The outputs and metrics are stored and presented to the end-user via a monitoring system. Most of the time, new types of data over time degrades the model performance. In this case, the engineer or data scientist will add new sample data to training datasets or change training parameters to retrain the algorithm and redeploy the model.

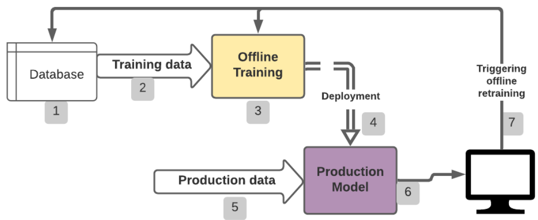

FIGURE 12. Single model ML pipeline.

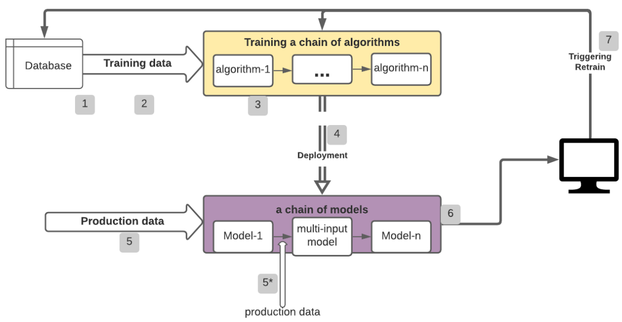

FIGURE 13. Chain of models.

Fig. 12 presents this logic. Data from datasets flows to an algorithm that is running offline. The output of this training is a model that meets the required metrics and can be used in production. The model gets deployed to production to generate inferences for the production data. This pipeline can run on any three controllers' platforms, depending on the data source and availability of computing resources for the production model. Running training activities on Near-RT RIC is possible if, firstly, the production model runs on the RAN edge and enough hardware resources are available for this purpose. Secondly, SMO and Non-RT RIC are not responsible for maintaining a model repository for the Open RAN ecosystem.

B. CHAIN OF MODELS

In this pipeline, the ML/AI solution is a combination of more than one model, that run either pure sequentially or partly asynchronous and partly sequentially. The chain of models can be treated as one model and map the previous solution to this one, with one input data source. However, there are scenarios in which, one of the models is a multi-input model. A multi-input model requires some considerations on providing the required input data. Fig. 13 shows a group of models that either generate output for another model in the same group, or use the output of one of models as input. A common example of this case is using AI models for pre-processing or post-processing of the main inference logic.

C. CHAMPION CHALLENGER AND ONLINE TRAINING

Champion challengers are trained models supposed to have a higher performance than the model currently running in the production. Online training refers to using the production

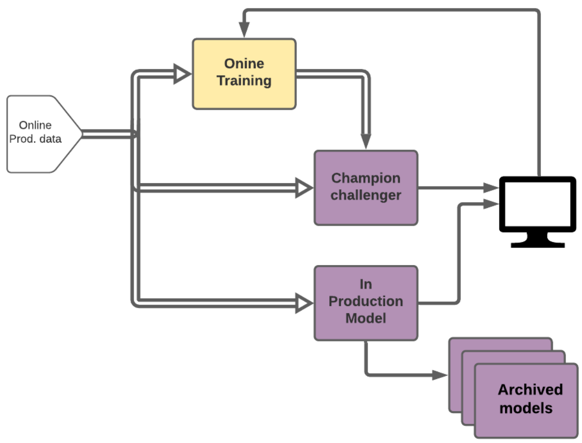

FIGURE 14. Online training with champion challenger.

stream of data to incrementally train the model and update the model to improve its performance.

Fig. 14 shows a diagram of online training combined with the concept of the champion challenger. Not all champion challengers are built from online training, but an online trained model is usually deployed as a champion challenger not to cause trouble for QoE. The traffic is split between the currently in production and the newly built models. After monitoring metrics and the new model's performance, the traffic will be routed entirely to the new model. The previously running model will be stored in an archive for future reference or any roll-back requirement.

D. A/B OR CANARY TESTING

In A/B or canary testing (Fig. 15) the goal is to compare the performance of two versions of the same model. These versions solve the same business problem with the same KPIs but are different, for instance in parameters, or the impact of a feature in the performance. In A/B testing, the ML pipeline splits the traffic between these two models (usually equally), and compares their performance in production. Then the traffic will be increased to 100% towards the superior one. The champion challenger scenario also can be called a case of A/B testing.

E. MLOPs AND ORAN

As we can see in each of the pipelines, the ML orchestration system is responsible for maintaining an ongoing performance of the ML/AI model and securing all communications between the items in the process. Splitting a pipeline to run on more than one platform will introduce security risks, but running the whole pipeline on one platform requires enough resources to store and process data besides running intelligent applications.

The architecture of Open RAN, the life cycle of data in this architecture, and the diversity of ML applications in telecommunication make MLOps solutions on Open RAN different from other industries. Automation and deployment of ML models require specific considerations, especially

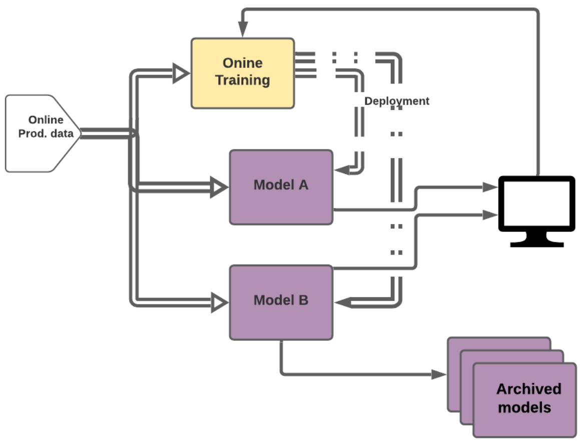

FIGURE 15. A/B and canary testing.

regarding security, availability, and reliability. This requirement emphasizes a new branch in the telecommunication industry that focuses on MLOps.

VII. CHALLENGES AND OPPORTUNITIES

In this section we explore opportunities and challenges in 5G and B5G that using AI in an Open RAN architecture can solve.

A. ARCHITECTURAL OPPORTUNITIES

We discussed the architecture of Open RAN and the variety of use cases that it supports. Open RAN introduces new interfaces such as A1 and new components such as Near-RT RIC. Moreover, combining virtualization and cloud-native architecture with Open RAN components generates many different scenarios to deploy discussed in Section IV-B. As a result, there are many different architectural options to install an end-to-end Open RAN scenario. With all the experience we have from deploying cloud solutions, including regional and edge, we cannot underestimate the challenges that O-Cloud designers and operators face. One of the challenges is the deployment of functions on the cloud platform. The current specifications and standards allow virtual functions and container deployment on cloud platforms. Although it might add flexibility to specifications, it can reduce compatibility between vendors' components and operators' available platforms. The other challenge will be connecting components. Wired or wireless, with interfaces or via API calls, components communicate with each other and their environment. Designing a robust, reliable, and scalable system requires a detailed and thorough study of requirements. Especially the requirements of different regions are not the same. For instance, the countryside with camping facilities is quite different from an urban region, considering these regions also need to be connected. In a multi-vendor system such as Open RAN, products made by different vendors have to be deployed, maintained, and secured by operators. Although standards work on Open RAN interfaces, each operator can have different requirements and configurations, such as namespaces for Kubernetes deployment. Vendors who build components and products for one operator might face difficulties in offering the same products to other operators.

B. NON-TERRESTRIAL USE CASES

Many use cases need very low latency or are in geographical locations where a solid urban network system does not exist. Utilizing mobile networking components such as aerial and vehicle-based components can be helpful in these situations.

For example, in a use case where an accident has happened, but paramedics cannot access the injured people, drones can fly over and carry resources such as robotic hands and a mobile cloud platform to run applications. Another opportunity to use a non-terrestrial implementation of Open RAN is where operators temporarily face hot spots. They can use UAS to ensure an acceptable level of QoS or QoE (Quality of Experience) with a low cost.

The main controversial challenge for non-terrestrial solutions might be providing them with enough power supply. The current enhancement in solar panels on space platforms such as satellites can enhance the efficiency of non-terrestrial networks in terms of energy [84]. Vehicles also can play the role of RAN components, and even base stations in the future of cities [85].

C. SECURITY

In Open RAN, splitting RAN components and opening interfaces in addition to accepting multi-vendor solutions will introduce vulnerabilities to the system [86]. Consequently, security experts need to focus on three major areas.

The first is the general telecommunication security problems, regardless of the RAN implementation. This one is a classic threat. In this security concern, network intrusion happens between UEs and RAN components. The DDOS attack, in which malicious actors target RAN and networking infrastructure through UEs, is a well-known example.

Secondly, security attacks are caused by Open RAN vulnerabilities. In Open RAN, many interfaces are open. Therefore, components made by different vendors can communicate. This feature makes a malicious actor capable of breaking into that communication if the interface is not managed brutally.

Thirdly, contaminated applications or components or faulty products can create intentional or unintentional threats to the system. In Open RAN, unlike previous versions of RAN, one vendor is not responsible for the end-to-end implementation and maintenance of the installed RAN. All vendors should follow the same security standards and tests before submitting any product for operation. Before deployment, operators can test products in a test environment to ensure products meet safety and quality requirements. In particular, implementing intelligent solutions that can identify anomalies and unexpected activities in real-time is vital.

Especially for cross-component communications, there is a need to create algorithms, write programs and generate models which can monitor, identify and prevent malicious activities. This opportunity is open for researchers and security engineers to find creative and efficient solutions and methods. Using certified protocols can improve security but can increase latency.

Networking operators also can cause shortcomings in the system if they do not apply strict measures to Open RAN implementation and engage vendors and third parties in deploying their products. This problem is preventable by adding processes to Open RAN installation specifications.

In sum, operators need a systematic and proactive approach for security threats in Open RAN. Also, they need to engage their vendors and monitor products actively.

D. IMPLEMENTING ORAN IN URBAN REGIONS

In combination with 5G and B5G, Open RAN will generate opportunities for serving diverse use cases, from mobile broadband to URLLC, and it will profoundly impact the future of technologies. However, more diversity in requests, more complicated to manage communications. In particular, autonomous vehicles and mission-critical applications are inseparable parts of future cities. The biggest challenge of these use cases is that they cannot tolerate interruption. Building a robust and reliable RAN creates unlimited project topics for researchers and engineers. These projects should focus on creating communications that follow the highest standard and are secure and reliable with efficient resource management.

Moreover, any of those projects should be environmentally friendly. Achieving these requirements without compromising latency and other QoS KPIs is another continuous challenge to explore.

Another crucial problem that Open RAN systems in urban areas have to overcome is attenuation in 5G and B5G. In 5G and B5G wireless generations, wireless communication will be in the range of mmWave and THz. Many wave blockers such as buildings, walls, and vehicles can create challenges for designing and implementing a reliable system in an urban environment. Weather conditions also act as blockers and interrupt mmWave and THz communication, which add to this complexity. These are topics of research that researchers have been working on since the idea of 5G started. A survey [87] discussed how some weather conditions such as heavy rain, temperature, and fog could affect the quality of mmWave communications. [88] also has studied the impact of rain on 5G communication. Considering that rain is a common phenomenon in many places and some cities even experience heavy rainfall for months, its severe impact on millimeter-wave propagation and signal loss cannot be overlooked.

Another atmospheric condition that can severely impact a telecommunication system is a hurricane. In [89] hurricanes and rain are mentioned as severe threats for 5G communication.

Future cities and people living in them require a wide range of telecommunication use cases, and any interruption can cause severe damages and huge costs, not only a financial cost but also loss of lives. Therefore, building and optimizing Open RAN in an urban area requires ongoing innovation and optimization research and engineering works. The end goal is to run a reliable, scalable, and secure telecommunication system with robust contingency plans that can serve all requests with a high level of QoS and QoE in any weather condition.

E. ZERO-TOUCH NETWORKS

New networking generations give rise to new use cases and new expectations. 5G and B5G promise to have faster speed and larger capacity while requiring very low latency and a high level of security. These requirements make automation and AI applications crucial parts of new telecommunication systems. A zero-touch networking system tries to delegate repetitive tasks to machines and let AI help with functionalities that humans cannot operate as fast or as accurately.

1) AUTONOMOUS NETWORKING

Over the past decades, networking controllers have evolved from hardware controllers to software-defined networks controllers and will soon enter the era of AI-controlled networks. The role of AI is expanding the flexibility that SDN provides and adds elasticity, self-maintenance, and higher performance. *Autonomous networking* is a phrase that started with the introduction of SDN as a means to use software products to manage and orchestrate networking in a region or for a use case. The other term common in mobile networking to address a network system without predefined network infrastructure and parameters is Self-Organizing Network (SON), widely used in mobile networking. 3GPP has explained SON in detail, in TS 32.500 [90]. Autonomous Network and SON are not new concepts. More than 20 years ago, we can find research discussing these concepts in terms of design and implementation. As a result of advancements in hardware and ML/AI algorithms, both SON and Autonomous networking are taking advantage of AI algorithms [91], [92].

The ETSI organization has published a white paper [93], dedicated explicitly to end-to-end autonomous networks and more interested in the operation and maintenance of telecommunication. It defines four levels for autonomous networking. From level one to four, it gradually decreases the role of humans in tasks and increases the role of artificial intelligence. Artificial intelligence's responsibilities in the system grow at each level from being assistive tools and failure recovery systems based on logs and alerts to failure diagnosis applications. Eventually, in the fourth level, being intelligent decision-makers and predictors. In the fourth level, the system autonomously can take precautionary actions and be in charge of failure and fault prediction. Using Open RAN in operation, levels three and four will become more crucial. In Open RAN, there will not be any specific vendor responsible for the end-to-end implementation of RAN. Therefore, it is vital to have a system that accurately diagnoses the problem and maps it to the source. Also, incorporating an application that can predict faults and failures can give potentially affected operators and vendors time to take action in a reasonable time. Once more, ML/AI applications become inevitable tools of future networking systems.

F. MANAGEMENT

Each of the individual topics mentioned before is a management challenge and opportunity. Operators and vendors can solve many of these problems using products and research results that researchers and engineers have created for other RAN implementations. We also have talked about managing configurations and the importance of preparing solid contingency plans. Nevertheless, some problems are unique to Open RAN. One of the unique problems is managing different vendors and their products. The multi-vendor concept helps operators take advantage of the competitive market and negotiate price and quality with their providers. However, it introduces many challenges.

Firstly, ensuring that products meet the required quality, pass security measures, and are compatible with other operators running products in their system. As we discussed, testbeds before deployment, universal specifications and standards, and engaging stakeholders and vendors in discussions can resolve or minimize this problem.

Secondly, open interfaces generate opportunities for malicious actors, and operators maintain communications' security intact. In addition to in-depth investigation and test on purchased components, they need to implement ML/AI applications and other intelligent systems to predict and detect flaws or malicious activities in real-time and react in real-time.

Thirdly, a RAN system's end-to-end maintenance and responsibility is not the vendors' responsibility anymore. Operators should establish procedures covering all parameters, communications, and activities for a test and monitoring. In addition, they need an intelligent application that can quickly map any flaw or problem to its possible sources. In this way, operators can quickly bring the related vendor on board and resolve the issue.

Also, in Open RAN, three intelligent controllers generate massive opportunities to optimize and innovate for telecommunication applications and use cases. Its orchestration has many use cases for software engineers to implement the best practices on deploying zero-touch data or ML pipelines and set up the best practices for deployment, continuous delivery, reliability, and scalability of software products in Open RAN.

Moreover, conflict resolution scenarios will become challenging as more applications for different telecommunication use cases run on the RIC platforms. Applications can have compatibility issues with the running infrastructure, or have conflicts on optimization and adjustment of parameters in their target network. If two AI applications directly or indirectly try to change the same parameter or change QoE of the same UE, there needs to be an intelligent system to prevent or resolve this issue.

As we can see, managing an Open RAN system more than other versions of RAN requires a strong collaboration between many sectors in engineering and business. This efficient collaboration becomes more crucial if we remind that an operator might not be able to implement what they build in one area to another because of many factors such as populations, diversity of use cases, type of user requests, different regulations, and different weather conditions.

G. DIGITAL TWIN FOR TEST AND IMPROVEMENT

We discussed the need for creating a test environment before deploying components and applications into production. Operators need to create a test environment to monitor the quality and reliability of products before integrating them into a running system. The current test environments are based on historical data. The ideal test could be monitoring the applications in the currently running environment while not deploying them into the actual production environment. A digital twin environment can assist with this problem. Based on [94] "A digital twin is a virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning, and reasoning to help decision-making.". As a result, applications can run in a digital twin environment and show how they behave in different circumstances.

The other benefit of creating a digital twin will be troubleshooting and fault finding in applications or components. If the outcome of a running application is not as expected, the digital twin environment can repeat the scenario, and operators can find the source of the issue or optimize parameters to prevent the same poor outcomes.

Creating digital twins increases the required resources, but by advancing hardware design and implementing data lifecycle policies, the benefits of utilizing digital twins, such as enhancing security and troubleshooting activities, can outweigh its cost for operators.

H. ENERGY CONSERVATION

In end-to-end telecommunication, the RAN consumes more than 80% of the wireless network power [95]. As a result of the growth in networking systems and their traffic which generate a large amount of data, energy consumption increases. Therefore, reducing the energy consumption in the RAN is a challenge in 5G and B5G that AI can address. For instance, intelligently allocating tasks to servers can economize energy consumption in the RAN. In [96] authors proposed an algorithm to optimize the offloading selection, radio resource allocation, and computational resource allocation in Mobile Edge Computing (MEC). In [97] researchers designed an autonomous control method to reduce the energy consumption of networking systems. [98] proposes a green cellular network by using Markov chain and modeling possible load variation. This information is used to select the most efficient base station. In sum, AI applications can help operators use the available resources efficiently with total capacity. Also, they can reduce redundancy in the system, which consequently can contribute to creating a green wireless networking system. Radio Frequency devices, the RAN components, and AI algorithms can be improved to conserve energy.

VIII. CONCLUSION

This paper started with an overall view of Open RAN and its architecture. We discussed the benefits of Open RAN in telecommunication. We argued that its multi-vendor and open interfaces would boost innovation. In addition, Open RAN helps operators to run multi-generation networking systems. In this paper, we focused on the role of AI in 5G, and B5G and how AI applications can run on an Open RAN architecture. We explained three families of applications responsible for the intelligent control of Open RAN. Telecommunication use cases in 5G and B5G are complex in terms of latency requirements, QoS KPIs, and the large number of parameters that they use and manage. AI can make faster and more accurate decisions than human-beings or static programs in such complex and dynamic environments with strict KPIs.

The last section of the paper is dedicated to the challenges and opportunities for implementing intelligent solutions in 5G and B5G telecommunication. This survey has discussed a new area of research area on AI and wireless network. The topics in this area can be topics of future research and implementation for researchers and engineers in the telecommunication and computer science industries. The future of telecommunication is painted by revolutionary communication methods that make implausible critical scenarios, possible. In addition to the advancements in AI algorithms and hardware, the Open RAN architecture can play a vital role in this significant technical evolution. The Open RAN architecture reduces cost and encourages multi vendor solutions that boosts the innovation rate, and the standard adaptability of new systems.

REFERENCES

[1] Wikipedia. *Radio Access Network*. Accessed: Jan. 2022. [Online]. Available: https://en.wikipedia.org/wiki/Radio_Access_Network

[2] H. Asplund, D. Astely, P. Butovitsch, T. Chapman, M. Frenne, F. Ghasemzadeh, M. Hagström, B. Hogan, G. Jöngren, J. Karlsson, F. Kronestedt, and E. Larsson, "Architecture and implementation aspects," in *Advanced Antenna Systems for 5G Network Deployments*, H. Asplund, D. Astely, P. Butovitsch, T. Chapman, M. Frenne, F. Ghasemzadeh, M. Hagström, B. Hogan, G. Jöngren, J. Karlsson, F. Kronestedt, and E. Larsson, Eds. New York, NY, USA: Academic, 2020, ch. 12, pp. 527–559. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128200469000125

[3] J. Kempf and P. Yegani, "OpenRAN: A new architecture for mobile wireless internet radio access networks," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 118–123, May 2002.

[4] B.-S.-P. LinI, "Toward an AI-enabled O-RAN-based and SDN/NFV-driven 5G & IoT network era," *Netw. Commun. Technol.*, vol. 6, no. 1, p. 6, Jun. 2021.

[5] L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, "Open, programmable, and virtualized 5G networks: State-of-the-art and the road ahead," *Comput. Netw.*, vol. 182, Dec. 2020, Art. no. 107516. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128620311786

[6] A. Huff, M. Hiltunen, and E. P. Duarte, "RFT: Scalable and fault-tolerant microservices for the O-RAN control plane," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage. (IM)*, Jun. 2021, pp. 402–409.

[7] N. Kazemifard and V. Shah-Mansouri, "Minimum delay function placement and resource allocation for open RAN (O-RAN) 5G networks," *Comput. Netw.*, vol. 188, Apr. 2021, Art. no. 107809. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128621000037

[8] L. Giupponi and F. Wilhelmi, "Blockchain-enabled network sharing for O-RAN," Jul. 2021, *arXiv:2107.02005*.

[9] L. Bertizzolot, T. X. Tran, J. Buczek, B. Balasubramanian, Y. Zhou, R. Jana, and T. Melodia, "Streaming from the air: Enabling high data-rate 5G cellular links for drone streaming applications," Apr. 2021, *arXiv:2101.08681*.

[10] B. Brik, K. Boutiba, and A. Ksentini, "Deep learning for B5G open radio access network: Evolution, survey, case studies, and challenges," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 228–250, 2022.

[11] S. K. Singh, R. Singh, and B. Kumbhani, "The evolution of radio access network towards open-RAN: Challenges and opportunities," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2020, pp. 1–6.

[12] M. Koziol, "The clash over 5G's first mile: The wireless industry is divided on open RAN's goal to make network components interoperable," *IEEE Spectr.*, vol. 58, no. 5, pp. 40–46, May 2021.

[13] P. E. I. Rivera, S. Mollahasani, and M. Erol-Kantarci, "Multi-agent team learning in virtualized open radio access networks (O-RAN)," Feb. 2021, *arXiv:2012.04861*. [Online]. Available: https://doi.org/10.48550/arXiv.2012.04861

[14] B. Balasubramanian, E. S. Daniels, M. Hiltunen, R. Jana, K. Joshi, R. Sivaraj, T. X. Tran, and C. Wang, "RIC: A RAN intelligent controller platform for AI-enabled cellular networks," *IEEE Internet Comput.*, vol. 25, no. 2, pp. 7–17, Mar. 2021.

[15] D. Johnson, D. Maas, and J. K. Van der Merwe, "Nexran: Closed-loop ran slicing in powder—A top-to-bottom open-source open-RAN use case," in *Proc. 15th ACM Workshop Wireless Netw. Testbeds, Exp. Eval. CHaracterization (WiNTECH)*, 2022, pp. 17–23.

[16] J. Breen, A. Buffmire, J. Duerig, K. Dutt, E. Eide, A. Ghosh, M. Hibler, D. Johnson, S. K. Kasera, E. Lewis, and D. Maas, "Powder: Platform for open wireless data-driven experimental research," *Comput. Netw.*, vol. 197, Oct. 2021, Art. no. 108281. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128621003017

[17] *O-RAN Alliance*. [Online]. Available: https://www.o-ran.org/

[18] *O-RAN Software Community*. Accessed: Jan. 2022. [Online]. Available: https://wiki.o-ran-sc.org/

[19] R. Schmidt, M. Irazabal, and N. Nikaein, *FlexRIC: An SDK for next-Generation SD-RANs*. New York, NY, USA: Association for Computing Machinery, 2021, pp. 411–425, doi: 10.1145/3485983.3494870.

[20] *Tip Project for Open RAN*. Accessed: Jan. 2022. [Online]. Available: https://telecominfraproject.com/openran/

[21] *O-Ran Architecture Description V03.00*, Telecom Infra Project OpenRAN, 2020.

[22] *O-RAN Use Cases and Deployment Scenarios*, O-RAN Alliance, White Paper, Feb. 2020.

[23] *AI/ML Workflow Description and Requirements, O-RAN.WG2.AIML-V01.03*, 2021.

[24] R. Ferrus, O. Sallent, J. Perez-Romero, and R. Agusti, "Applicability domains of machine learning in next generation radio access networks," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2019, pp. 1066–1073.

[25] P. Somwang and W. Lilakiatsakun, "Computer network security based on support vector machine approach," in *Proc. 11th Int. Conf. Control, Autom. Syst.*, 2011, pp. 155–160.

[26] T. Zhi, Y. Liu, J. Wang, and H. Zhang, "Resist interest flooding attacks via entropy–SVM and Jensen–Shannon divergence in information-centric networking," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1776–1787, Jun. 2020.

[27] S. Y. Mehr and B. Ramamurthy, "An SVM based DDoS attack detection method for Ryu SDN controller," in *Proc. 15th Int. Conf. Emerg. Netw. EXp. Technol.*, 2019, pp. 72–73.

[28] R. Guerra-Gomez, S. R. Boque, M. Garcia-Lozano, and J. O. Bonafe, "Machine-learning based traffic forecasting for resource management in C-RAN," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2020, pp. 200–204.

[29] P. Ghosh and R. Mitra, "Proposed GA-BFSS and logistic regression based intrusion detection system," in *Proc. 3rd Int. Conf. Comput., Commun., Control Inf. Technol. (C3IT)*, Feb. 2015, pp. 1–6.

[30] S. Yadav and S. Selvakumar, "Detection of application layer DDoS attack by modeling user behavior using logistic regression," in *Proc. 4th Int. Conf. Rel., Infocom. Technol. Optim. (ICRITO) (Trends Future Directions)*, Sep. 2015, pp. 1–6.

[31] R. Shah, Y. Qian, D. Kumar, M. Ali, and M. Alvi, "Network intrusion detection through discriminative feature selection by using sparse logistic regression," *Future Internet*, vol. 9, no. 4, p. 81, Nov. 2017. [Online]. Available: https://www.mdpi.com/1999-5903/9/4/81

[32] N. Ben-Amor, S. Benferhat, and Z. Elouedi, "Naive Bayes vs decision trees in intrusion detection systems," in *Proc. ACM Symp. Appl. Comput.*, 2004, pp. 420–424.

[33] L. Koc, T. A. Mazzuchi, and S. Sarkani, "A network intrusion detection system based on a hidden naïve Bayes multiclass classifier," *Exp. Syst. Appl.*, vol. 39, no. 18, pp. 13492–13500, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417412008640

[34] G. K. Reddy and S. P. Thilagam, "Naïve Bayes classifier to mitigate the DDoS attacks severity in ad-hoc networks," *Int. J. Commun. Netw. Inf. Secur.*, vol. 12, no. 2, pp. 221–226, Apr. 2020.

[35] D. Liu, Z. Liu, and Z. Song, "LDA-based CSI amplitude fingerprinting for device-free localization," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Aug. 2020, pp. 2020–2023.

[36] J. Jiang, X. Zhu, G. Han, M. Guizani, and L. Shu, "A dynamic trust evaluation and update mechanism based on C4.5 decision tree in underwater wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9031–9040, Aug. 2020.

[37] J. Wang, Q. Yang, and D. Ren, "An intrusion detection algorithm based on decision tree technology," in *Proc. Asia–Pacific Conf. Inf. Process.*, vol. 2, Jul. 2009, pp. 333–335.

[38] M. H. Mazhar and Z. Shafiq, "Real-time video quality of experience monitoring for HTTPS and QUIC," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Apr. 2018, pp. 1331–1339.

[39] Y. Xie, Y. Wang, A. Nallanathan, and L. Wang, "An improved K-nearest-neighbor indoor localization method based on Spearman distance," *IEEE Signal Process. Lett.*, vol. 23, no. 3, pp. 351–355, Mar. 2016.

[40] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.

[41] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2595–2621, 4th Quart., 2018.

[42] R. Primartha and B. A. Tama, "Anomaly detection using random forest: A performance revisited," in *Proc. Int. Conf. Data Softw. Eng. (ICoDSE)*, Nov. 2017, pp. 1–6.

[43] S.-H. Choi, D.-H. Hwang, and Y.-H. Choi, "Wireless intrusion prevention system using dynamic random forest against wireless MAC spoofing attack," in *Proc. IEEE Conf. Dependable Secure Comput.*, Aug. 2017, pp. 131–137.

[44] *K-Means*. Accessed: Feb. 2022. [Online]. Available: https://en.wikipedia.org/wiki/K-means_clustering

[45] E. Balevi and R. D. Gitlin, "A clustering algorithm that maximizes throughput in 5G heterogeneous F-RAN networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[46] H. Harb, A. Makhoul, D. Laiymani, A. Jaber, and R. Tawil, "K-means based clustering approach for data aggregation in periodic sensor networks," in *Proc. IEEE 10th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2014, pp. 434–441.

[47] G. Wang, Y. Zhao, J. Huang, Q. Duan, and J. Li, "A K-means-based network partition algorithm for controller placement in software defined network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[48] L. Tong-yan and L. Xing-ming, "The study of alarm association rules mining in telecommunication networks," in *Proc. Int. Conf. Commun., Circuits Syst.*, May 2008, pp. 1030–1034.

[49] C. Li and X. Huang, "Research on FP-growth algorithm for massive telecommunication network alarm data based on spark," in *Proc. 7th IEEE Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Aug. 2016, pp. 875–879.

[50] Q. Wang, K. Feng, X. Li, and S. Jin, "PrecoderNet: Hybrid beamforming for millimeter wave systems with deep reinforcement learning," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1677–1681, Oct. 2020.

[51] H. Vaezy, M. Salehi Heydar Abad, O. Ercetin, H. Yanikomeroglu, M. J. Omidi, and M. M. Naghsh, "Beamforming for maximal coverage in mmWave drones: A reinforcement learning approach," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1033–1037, May 2020.

[52] C.-H. Zhong, K. Guo, and M. Zhao, "Online sparse beamforming in C-RAN: A deep reinforcement learning approach," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.

[53] D. Guo, L. Tang, X. Zhang, and Y.-C. Liang, "Joint optimization of handover control and power allocation based on multi-agent deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13124–13138, Nov. 2020.

[54] T. M. Ho and K.-K. Nguyen, "Joint server selection, cooperative offloading and handover in multi-access edge computing wireless network: A deep reinforcement learning approach," *IEEE Trans. Mobile Comput.*, vol. 21, no. 7, pp. 1–15, Jul. 2020.

[55] M. Sana, A. De Domenico, E. C. Strinati, and A. Clemente, "Multi-agent deep reinforcement learning for distributed handover management in dense mmWave networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 8976–8980.

[56] *Intelligent 5G L2 MAC Scheduler, Powered by Capgemini Netanticipate 5G on Intel Architecture*. Accessed: Feb. 2022. [Online]. Available: https://builders.intel.com/docs/networkbuilders/intelligent-5g-l2-mac-scheduler-powered-by-capgemini-netanticipate-5g-on-intel-architecture-v13.pdf

[57] H. Saki, N. Khan, M. G. Martini, and M. M. Nasralla, "Machine learning based frame classification for videos transmitted over mobile networks," in *Proc. IEEE 24th Int. Workshop Comput. Aided Model. Des. Commun. Links Netw. (CAMAD)*, Sep. 2019, pp. 1–6.

[58] *Vehicle to Evrything*. Accessed: Feb. 2022. [Online]. Available: https://en.wikipedia.org/wiki/Vehicle-to-everything#802.11p_(DSRC)

[59] M. Jiang, "Device-controlled traffic steering in mobile networks," in *Proc. 9th Int. Conf. Next Gener. Mobile Appl., Services Technol.*, Sep. 2015, pp. 7–12.

[60] K. Adachi, M. Li, P. H. Tan, Y. Zhou, and S. Sun, "Q-learning based intelligent traffic steering in heterogeneous network," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1–5.

[61] S. Barmpounakis, A. Kaloxylos, P. Spapis, C. Zhou, P. Magdalinos, and N. Alonistioti, "Data analytics for 5G networks: A complete framework for network access selection and traffic steering," *Int. J. Adv. Telecommun.*, vol. 11, nos. 3–4, p. 101, 2018.

[62] F. D. Priscoli, A. Giuseppi, F. Liberati, and A. Pietrabissa, "Traffic steering and network selection in 5G networks based on reinforcement learning," in *Proc. Eur. Control Conf. (ECC)*, May 2020, pp. 595–601.

[63] A. Rostami, P. Ohlen, K. Wang, Z. Ghebretensae, B. Skubic, M. Santos, and A. Vidal, "Orchestration of RAN and transport networks for 5G: An SDN approach," *IEEE Commun. Mag.*, vol. 55, no. 4, pp. 64–70, Apr. 2017.

[64] S. Matoussi, I. Fajjari, S. Costanzo, N. Aitsaadi, and R. Langar, "5G RAN: Functional split orchestration optimization," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 7, pp. 1448–1463, Jul. 2020.

[65] "Configuring storm control," in *Catalyst 4500 Series Switch Software Configuration Guide, 12.2(53)SG*. San Jose, CA, USA: Cisco, 2018. [Online]. Available: https://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst4500/12-2/53SG/configuration/config/bcastsup.html

[66] L. V. Le, B.-S. Lin, L.-P. Tung, and S. Do, "Enhanced handover clustering and forecasting models based on machine learning and big data," *Trans. Mach. Learn. Artif. Intell.*, vol. 6, no. 5, p. 43, Oct. 2018.

[67] L. Hao and B. Ng, "Self-healing solutions for Wi-Fi networks to provide seamless handover," in *Proc. IFIP/IEEE Symp. Integr. Netw. Service Manage. (IM)*, Apr. 2019, pp. 639–642.

[68] S. Memon and M. Maheswaran, "Using machine learning for handover optimization in vehicular fog computing," in *Proc. 34th ACM/SIGAPP Symp. Appl. Comput.*, 2019, pp. 182–190.

[69] B. Hussain, Q. Du, and P. Ren, "Semi-supervised learning based big data-driven anomaly detection in mobile wireless networks," *China Commun.*, vol. 15, no. 4, pp. 41–57, Apr. 2018.

[70] Q. Liu, C. F. Kwong, S. Zhang, L. Li, and J. Wang, "A fuzzy-clustering based approach for MADM handover in 5G ultra-dense networks," *Wireless Netw.*, vol. 28, no. 2, pp. 965–978, Feb. 2019.

[71] Z. Kaleem, M. Khaliq, A. Khan, I. Ahmad, and T. Duong, "PS-CARA: Context-aware resource allocation scheme for mobile public safety networks," *Sensors*, vol. 18, no. 5, p. 1473, May 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/5/1473

[72] H. Nishiyama, Y. Kawamoto, and D. Takaishi, "On OFDM-based resource allocation in LTE radio management system for unmanned aerial vehicles (UAVs)," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2017, pp. 1–5.

[73] M. L. Marí-Altozano, S. Luna-Ramírez, M. Toril, and C. Gijón, "A QoE-driven traffic steering algorithm for LTE networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 11271–11282, 2019.

[74] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Jun. 2014.

[75] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[76] S. Nie, J. M. Jornet, and I. F. Akyildiz, ''Intelligent environments based on ultra-massive MIMO platforms for wireless communication in millimeter wave and terahertz bands,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 7849–7853.
[77] D. Bega, M. Gramaglia, C. J. B. Cano, A. Banchs, and X. Costa-Perez, ''Toward the network of the future: From enabling technologies to 5G concepts,'' *Trans. Emerg. TeleCommun. Technol.*, vol. 28, no. 8, p. e3205, Jun. 2017.
[78] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, ''From network sharing to multi-tenancy: The 5G network slice broker,'' *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 32–39, Jul. 2016.
[79] *Management and Orchestration; Concepts, Use Cases and Requirements*, document Rec. TS 28.530, 3GPP, Dec. 2017. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3273
[80] *Technical Specification Group Services and System Aspects; System Architecture for the 5G 30SystemStage 2*, document Rec. TS-28.501, 3GPP, 2019.
[81] W. Jiang, S. D. Anton, and H. Dieter Schotten, ''Intelligence slicing: A unified framework to integrate artificial intelligence into 5G networks,'' in *Proc. 12th IFIP Wireless Mobile Netw. Conf. (WMNC)*, Sep. 2019, pp. 227–232.
[82] J. Mei, X. Wang, and K. Zheng, ''An intelligent self-sustained RAN slicing framework for diverse service provisioning in 5G-beyond and 6G networks,'' *Intell. Converged Netw.*, vol. 1, no. 3, pp. 281–294, Dec. 2020.
[83] E. Esenogho, K. Djouani, and A. M. Kurien, ''Integrating artificial intelligence Internet of Things and 5G for next-generation smartgrid: A survey of trends challenges and prospect,'' *IEEE Access*, vol. 10, pp. 4794–4831, 2022.
[84] M. Giordani and M. Zorzi, ''Non-terrestrial networks in the 6G era: Challenges and opportunities,'' *IEEE Netw.*, vol. 35, no. 2, pp. 244–251, Mar./Apr. 2021.
[85] C. Li, Q. Luo, G. Mao, M. Sheng, and J. Li, ''Vehicle-mounted base station for connected and autonomous vehicles: Opportunities and challenges,'' *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 30–36, Aug. 2019.
[86] (Jun. 2021). *5G Open Ran Ecosystem Whitepaper*. [Online]. Available: https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/whitepaper_5g_open_ran/OREC_WP.pdf
[87] N. P. Narekar and D. M. Bhalerao, ''A survey on obstacles for 5G communication,'' in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2015, pp. 0831–0835.
[88] M. Ghanim, M. Alhilali, J. Din, and H. Y. Lam, ''Rain attenuation statistics over 5G millimetre wave links in Malaysia,'' in *Proc. 5th Int. Conf. Electr. Eng., Comput. Sci. Informat. (EECSI)*, Oct. 2018, pp. 266–269.
[89] K. Yasin. *5G Mobile Communication Systems: Fundamentals, Challenges, and Key Technologies*. Singapore: Springer, 2019.
[90] *3GPP, Self-Organizing Network*. Accessed: Feb. 2022. [Online]. Available: https://www.3gpp.org/technologies/keywordsacronyms/105-son
[91] S. Fuhrmann, Y. Kogan, and R. A. Milito, ''An adaptive autonomous network congestion controller,'' in *Proc. 35th IEEE Conf. Decis. Control*, vol. 1, Dec. 1996, pp. 301–306.
[92] Y. Chun, L. Qin, L. Dongsheng, and S. MeiLin, ''QoS routing in self-organized network,'' in *Proc. Int. Conf. Commun. Technol. (WCC ICCT)*, 2000, pp. 1304–1309.
[93] *Autonomous Networks, Supporting Tomorrow's ICT Business*, 1st ed., ETSI, Oct. 2020.
[94] *Cheat Sheet: What is Digital Twin?* Accessed: Feb. 2022. [Online]. Available: https://www.ibm.com/blogs/internet-of-things/iot-cheat-sheet-digital-twin/
[95] A. E. Azzaoui, S. K. Singh, Y. Pan, and J. H. Park, ''Block5Gintell: Blockchain for AI-enabled 5G networks,'' *IEEE Access*, vol. 8, pp. 145918–145935, 2020.
[96] P. Zhao, H. Tian, C. Qin, and G. Nie, ''Energy-saving offloading by jointly allocating radio and computational resources for mobile edge computing,'' *IEEE Access*, vol. 5, pp. 11255–11268, 2017.
[97] K. Awahara, S. Izumi, T. Abe, and T. Suganuma, ''Autonomous control method using AI planning for energy-efficient network systems,'' in *Proc. 8th Int. Conf. Broadband Wireless Comput., Commun. Appl.*, Oct. 2013, pp. 628–633.
[98] R. Li, Z. Zhao, X. Zhou, G. Ding, Y. Chen, Z. Wang, and H. Zhang, ''Intelligent 5G: When cellular networks meet artificial intelligence,'' *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 175–183, Oct. 2017.

AZADEH ARNAZ (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with the School of Electrical and Data Engineering, University of Technology Sydney (UTS). She has more than six years of industry experience in software engineering, big data engineering, and machine learning. She has worked on machine learning projects in various industries, such as health and entertainment. Her research interests include reinforcement learning and machine learning.

JUSTIN LIPMAN (Senior Member, IEEE) received the Ph.D. degree in telecommunications engineering from the University of Wollongong, Australia, in 2004. He is currently an Industry Associate Professor at the University of Technology Sydney (UTS) and a Visiting Associate Professor at the Graduate School of Engineering, Hokkaido University. He is also the Director of research translation at the Faculty of Engineering and IT and the RF Communications Technologies (RFCT) Laboratory, where he leads industry engagement in RF technologies, cybersecurity, privacy preserving technologies, the Internet of Things, and tactile internet. He was previously the Deputy Chief Scientist of the Food Agility Cooperative Research Center. Prior to joining UTS, over a 12 year period, he held a number of senior management and technical leadership roles at Intel and Alcatel driving research and innovation, product development, architecture, and IP generation. His research interests include ''things'' adaptive, connected, distributed, ubiquitous, and secure. He serves as a Committee Member for Standards Australia contributing to International IoT standards and Digital Twins.

MEHRAN ABOLHASAN (Senior Member, IEEE) is currently an Associate Professor with the University of Technology Sydney. He has served in various leadership roles over the past ten years, such as the Director of research programs with the Faculty of Engineering and IT, representing Australia as the Head of Delegates for ISO/IEC IoT Standardization meetings. He is also serving as the Deputy Head of School for Research with the School of Electrical and Data Engineering. He has expertise in a number of different areas, such as the IoT and sensor network technologies, wireless networking and analytics, software-defined networking, algorithm development, simulation and modeling, and development of test-beds. His experience and contribution to the above areas has resulted in significant research outcomes, which have been published in highly prestigious IEEE journals and achieving high citations. He has authored over 160 international publications and won over U.S. $5 million in research funding. He has led and worked as the Chief Investigator on a diverse range of projects, such as ARC Discovery Project, ARC Linkage, ARC LIEF, CRC, and various government and Industry funded projects. His current research interests include software-defined networking, the IoT, WSNs, intelligent transportation systems, wireless multi-hop networks, 5G networks, and beyond. He regularly serves as a reviewer for various IEEE, ACM, and Elsevier journals. He also serves many conferences as a reviewer, a TPC, and an Organizer.

MATTI HILTUNEN (Senior Member, IEEE) is currently a Lead Inventive Scientist at AT&T Labs Research, Florham Park, NJ, USA. He acted as the PTL for the RICAPP Project at O-RAN Software Community, from 2019 to 2021. His research interests include Open RAN, dependable distributed systems and networks, and adaptive systems. He is a member of the IEEE Computer Society, and IFIP Working Group 10.4 on Dependable Computing and Fault-Tolerance.