Telco Data opportunities and current limits in 3GPP networks: A brief review

Luís Carlos Gonçalves*
*Work done at Institut Polytechnique de Paris, Télécom SudParis, Palaiseau, France
May 2025
Emails: luisgo@luisgo.pro
https://www.luisgo.pro

***Abstract*—A Telecommunication network generates a large amount of different types of data, such as user data or network control data (terabytes to petabytes of data on a daily basis depending on the telecommunications company) that can be used for instance, to provide user/client assistance and prevent or recover the network from outages, help in the network resource configuration, and also in some high-level user applications such as user mobility and generated traffic behavior. With this huge volume of data, the recent advances in Artificial Intelligence associated with Massive Parallel Computing provided by Graphics Processing Units, Field-Programmable Gate Arrays and Neural Processing Units have opened new avenues in telecommunication processes to rely on Artificial Intelligence based automation to reduce network configuration complexity and introduce new innovative Artificial Intelligence based approaches in networks and services. This study addresses several topics related to Data Analytics in Telecommunications. It attempts to provide a short, essential, not exhaustive and updated coverage of the subject. It is a good first read, jointly with the suggested bibliography, for newcomers on the subject, as it is an update to experienced ones.**

***Index Terms*—Data Analytics in Telecommunications, Telco Data, Telco Datasets, Edge Computing, Cloud Computing, Dataset Repositories, Artificial Intelligence, High Performance Computing.**

I. Introduction

The large amount of data generated by telecommunications networks can be used to make demographic studies[1], to be input to software agents that manage of the network in order to provide user/client assistance, heal the network etc. The newcomer technology for 5G and 6G Telecommunications Networks that enforces the paradigm for telco Data Analytics is Artificial Intelligence (AI) mature solution. In addition, breakthroughs in Artificial Intelligence are only possible with Massive Parallel Computing provided by Graphics Processing Units (GPUs), Field-Programmable Gate Arrays (FPGAs) and/or Neural Processing Units (NPUs). GPUs and FPGAs, beyond their use in Artificial Intelligence, permit the implementation of any algorithms requiring high computing complexity and/or the need to process large amounts of data.

Another new concept is the movement of processing from standalone computers on premises to virtualized computing in the cloud, thus opening the avenue for the new Telco Cloud approach, where classical telco architecture is evolving with cloud concepts and architectures. This permits scalability of the processing capability and recovery of software and hardware problems (healing). However it is still necessary to perform processing on premises for critical latency and/or secure applications.

Although the new telco cloud architecture opens large possibilities via open interfaces to hook innovative AI-based applications in the telco architecture, one problem in the research of Telecommunications Big Data Analytics is the scarce availability of public domain telecom datasets due to privacy and competition concerns.

These topics and other related are covered in the next sections.

Some considerations regarding 5G networks and beyond are presented in Section II. The major players in software and/or cloud providers for Data Analytics are presented in Section III. The use of Artificial Intelligence in telecom data analytics is described in Section IV. The usage of datasets and related repositories is presented in Section V. Finally, the conclusions are presented in Section VI.

II. The 5G Network

The advantages and drawbacks of the network functional splits at the physical layer or Data Link Layer of 4G and 5G Networks in O-RAN networks, in terms of data rate in the split, energy savings in each split, etc are analyzed in [1]. The split divides the processing

[1] https://cros.ec.europa.eu/MNOdata4OS

between the Distributed Units, near the Radio Interface, and the Central Unit, near or in the cloud. The architecture, interfaces and algorithms of O-RAN networks are presented in [2]. The near real-time intelligent controller, the non real-time intelligent controller named RICs, part of the O-RAN architecture are also presented. Non real-time RICs run algorithms in named *rApps* applications and in near real-time RICs named *xApps*. These software agents can control and manage 3GPP-defined RANs in closed loop and with AI algorithms to find and apply control policies and actions on the RAN, such as RAN slicing, load balancing, handovers and scheduling policies. 3GPP has recently defined the Network Data Analytics Function (NWDAF) which is part of the 5G core network and is designed to streamline core network data. NWDAF uses AI algorithms and automation techniques to analyze data and provide insights. The topic of NWDAF for power decisioning in the 5G core and RIC to drive intelligence in the RAN is both timely and extremely important. However, advances in AI for Mobile Networks are slow because of the lack of large-scale datasets, experimental testbeds, or real networks needed to perform training. Instead Deep Reinforcement Learning (DRL) agents, which are based on trial and error training, are successfully used. The non real-time RIC is integrated with the network orchestrator and operates with timescales of more than 1 second. The near real-time RIC has a time scale from 10 ms to 1 s and controls the RAN nodes. These software agents interact with the Central Unit, Distribution Unit and RAN nodes through REST APIs interfaces. The evolution of O-RAN considering 6G requirements is presented in [2].

Slice Management, Resource Allocation and Orchestration in 5G and in the future 6G could involve many simultaneous or disjoint degrees of freedom such as revenue, throughput, latency, reliability etc., whose equilibrium or join minimization/maximization solving can be achieved with optimization-based techniques, Machine Learning (ML) or Game Theory [3] on a single or multiple slice basis. The variables to be optimized simultaneously can spread from several domains, namely, the Core Network, Transport Network and Radio Access Network, as an end-to-end optimization. However, as the number of optimization variables increases, the problem becomes intractable. Optimization-based techniques only cover known network changes, requiring a new complete model to deal with new events [3] that can be impractical. In [3], a survey of articles on Resource Management over single and multiple domains using the above techniques (Optimization, ML and Game Theory) was presented.

III. From Big Data to Cloud and Edge Cloud

Nowadays, data can be taken from mobile communications users or the Internet of Things (IoT) provided by Mobile Network Operators (MNOs), email users provided by email providers (*e.g.* Google), and social networks providers to complement demographic statistics, such as population or other census, or for new ones. In addition, data can be obtained from WiFi providers at an University Campus or provided by city coverage. The target clients are from National Statistics Institutes to private sector companies that (the latter one) want to make market studies. In addition, it can be used in under developed or developing countries to provide statistics considering that they can not afford the normal census or statistics of unemployment that are usual in developed countries [4].

Related to Big Data Analytics (BDA) there is a trend of migration of software based on premises to the cloud in named software microservices.

The cloud providers provide three types of services:

- Service level multitenancy: In this case the hardware infrastructure and a single software host are shared between several clients.
- Dedicated bare-metal machines: One machine exclusive to one client. It is not a flexible, elastic or scalable solution.
- A container platform to schedule stateless applications (*e.g.* microservices) in a virtualized cloud. This solution exhibits software stability problems as claimed in [5].

A centralized deployment of software in the cloud benefits of commercial off the shelf (COTS) hardware savings, auto-scaling, load-balancing, self-healing and a high scale availability of hardware. Some drawbacks are lost of performance due to the virtualization/abstraction layer(s) and added complexity to troubleshooting and debuging. Critical features such as modularity, scalability, reliability and operability are now addressed by cloud-native-based software instead of hardware. Articles on the migration of standalone applications to microservices to the cloud document are [6], [7]. The migration to microservices best practices and features to take into account from a standalone platform software to a cloud platform software is shown in Figure 2, reproduced from [6].

The players and providers of Cloud, are Amazon, Google and Microsoft in USA, OVHCloud, Scaleway in Europe, ALIBABA Cloud in China. They also provide software solutions for Data Analytics which also include RapidSpace[2] and Apache Software Foundation[3].

[2] https://www.rapid.space/

[3] https://projects.apache.org/projects.html?category

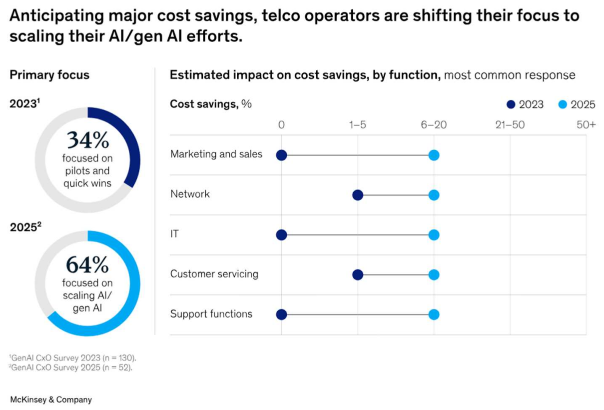

Figure 1. AI adoption cost benefit for telco operators activities.

The major impediment to BDA in the context of telecommunications is the regulation and price that access Telco Data might have, while the cost savings for telco operators adopting an AI-based network and application functions is established as plotted in the Figure 1[4].

Big Data Analytics is an expensive, resource-intensive and complex process that requires highly qualified and diversified human resources and is subject to a high failure rate. This scenario can be changed with new applications that are faceted to be used in a BDA pipeline: MongoDB, Redis, Hbase, Apache Spark, Kafka, Apache Flink, Storm, Apache Hadoop, Nexedi Wendelin[5] and SimpleRAN[6]. Some of the big data tools as follows:

- *Apache Hbase*: Apache HBase provides non-relational distributed bigtable-like database capabilities on top of Hadoop and Hadoop Distributed File System (HDFS)[7]. The Apache HBase is based on the model described in [8].
- *Apache Hive*: Apache Hive is a distributed, fault-tolerant data warehouse system that enables inquiries of large databases over Hadoop. Hive allows inquiries (reading, writing and managing) of petabyte databases using SQL.
- *Apache Pig*: Pig's language currently consists of a textual language called Pig Latin with an attached compiler that produces sequences of Map-Reduce programs[8], for which large-scale parallel implementations already exist in the Hadoop subproject.
- *Apache Spark*: Spark is a multi-language engine for executing data engineering, data science, and machine learning on single-node machines or clusters. It supports In-memory computing[9] and acyclic data flow (DAG). A DAG engine organizes the tasks to run them jointly faster. Spark is ideal for applications requiring a high throughput [5].
- *Apache Kafka*: Kafka is an open-source engine for real-time low-latency stream processing written in Scala and Java. It runs on top of Apache Spark, Apache Flink or Apache Storm.
- *Apache Flink*: Flink is a distributed realtime framework used to processing streams. Flink has low latency, permitting its use in real-time monitoring, reports, stream data analysis and data warehousing [5].
- *Nexedi Wendelin*: Wendelin is a convergent platform for Big Data and Machine Learning, and a variant of ERP5[10] with extensions for ndarrays, a core module managing RAM beyond physical limits and interfaces with libraries such as scikit-learn, jupyter, pandas, fluentD or embulk. Wendelin is hosted on SlapOS and uses NEO for data storage, allowing the management of the data lifecycle from ingestion to commercialization. It was developed and maintained by Nexedi.
- *SimpleRAN SimpleSpace* deals with AI based heterogeneous data interoperability and access for third parties.

Apache Hadoop, Apache Spark and Apache Flink can run over Kubernetes, that is in containers in the cloud. However, a monolithic architecture is not efficient when handling large amounts of data and lacks agility in scaling analytics processes [9].

Because of its disk-based data processing with no in-memory processing, no low-level processing API, and Java only support, Hadoop is unsuitable for massive-scale and complex big data processing [5].

Another Big Data Analytics Bundle (as Apache Hadoop, Apache Spark or Apache Flink) is Wendelin from Nexedi.

Amazon also includes in their services Big Data Analytics, such AWS DataSync, AWS Glue, AWS Lambda, Amazon SageMaker, Amazon Redshift and Amazon QuickSight. These services includes[11]:

[4]https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/scaling-the-ai-native-telco

[5]https://wendelin.nexedi.com/

[6]https://www.nexedi.com/RS-Presentation.SimpleRAN.Taipei

[7]HDFS is the primary storage system used by Hadoop applications.

[8]Map-Reduce is a Java-based, distributed execution framework within the Apache Hadoop bundle. It have two processing steps that developers implement: Map and Reduce.

[9]In-memory processing is a computer architecture in which data operations are locally done, with a local processor, on the data memory, rather than having to be transferred to CPU registers first.

[10]ERP5 is an open source Enterprise resource planning (ERP) software based on Python and Zope.

[11]https://aws.amazon.com/fr/blogs/industries/enable-analytics-and-insights-for-telecom-networks/

- Data migration from local site data format to Amazon Cloud data format.
- Data discovery, preparation and integration.
- Business Intelligence reports generation, display and notification.
- Build complex AI models to predict anomalies and root cause analytics.

Amazon also has tools to automate the management of networks in the cloud, such as AWS Telco Network Builder (which automates the deployment and management of telecom networks).

For latency-sensitive applications, Amazon provides hardware to be deployed on client sites (*i.e.*, MNO Network Edge and Basestations) named AWS Wavelength, AWS Local Zones, or AWS Outposts. Can be cited[12]: "AWS Outposts are fully managed and configurable compute and storage racks built with AWS-designed hardware, which allows customers to run compute and storage on-premises while seamlessly connecting to AWS's broad array of cloud services."

Microsoft provides similar services as Amazon, such as automate management of networks in the cloud, deployment of the services in client sites, near edge and far edge and AI Data Analytics processing through the Azure Operator Nexus Platform[13].

In Table I, a comparison between Wendelin and Apache Hadoop performed by Nexedi is presented[14].

Table II presents a comparison of services lifecycle automation (SLM) platforms given by Nexedi SimpleRAN. Wendelin is the only one that uses Open-Source/Free SLM SlapOS, which is a full-fledged platform that can be deployed in most cloud services and even in an owned and standalone hardware platform. However, contrary to the software, Nexedi's support services are paid. Amazon Big Data Analytics can only be deployed in the Amazon cloud.

The most well-known AI frameworks include Google TensorFlow (with or without Keras), Facebook PyTorch, and Apache MXNet. The Wendelin bundle uses Scikit-learn[15] for the AI. These frameworks rely on GPUs processing.

There are several examples of using Big Data not only for statistics but also to influence people in polls/elections as happened with the Trump election at 2016 with Facebook user data explored by the contracted company Cambridge Analytica.

Citations of articles and documentation until 2019, whose subject is the application of Telco Data to compute statistics of population roaming, population in events, etc., can be found in [10].

[12] https://aws.amazon.com/wavelength/faqs/

[13] https://azure.microsoft.com/en-us/products/operator-nexus

[14] Some care must be taken about an eventual existence of bias in the filling of this table.

[15] https://scikit-learn.org/stable/

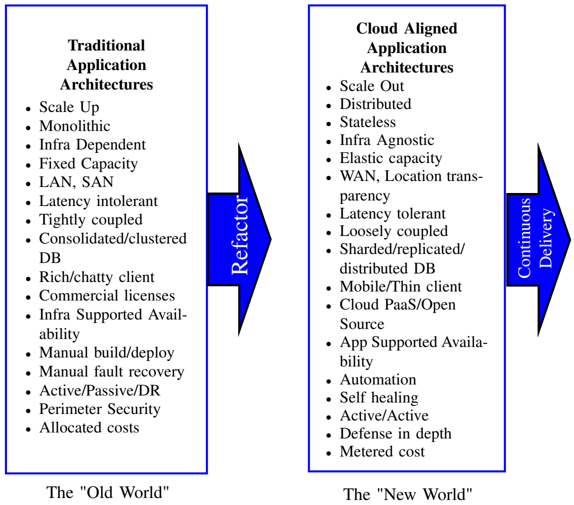

Figure 2. Migration from a standalone platform to a Cloud platform suggested requirements and features to modify.

In [11] a Literature Review of BDA from data of Learning Management Systems (LMS) is made.

In [12], a Literature Review of BDA was conducted in telecommunications until 2019. A survey of the latest BDA concepts and technologies was presented in [5].

In the context of telco data-related innovation, telco data API access is key, as it might offer multi-access to multiple players and allow different innovative applications and services. Securing access to the data and identifying access players is technically challenging as it requires security and interoperability in data access. One can cite the DIN 27070 International Data Spaces Association (IDSA) framework [13] to offer standard secure and interoperable access to multi party players, and it can also cite in the context of Industry 4.0 the OPC-UA framework [14] that also offers standard data access.

Another important work is the technical specification of ETSI related to the 3GPP API [15], which proposes ways to open some APIs in 3GPP 5G networks to access data for innovative services.

In the context of the SIMPLERAN Project [16], a framework for online and offline data access by multiple parties ensuring security and interoperability is under development in the context of 5G the innovative aspect of telco-data secure and interoperable data access for new innovative and automatic network function automation.

IV. Artificial Intelligence and Telecommunications

The deployment of an Artificial Intelligence Network has two stages: training stage and inference. The training stage is more demanding in terms of resources, in memory and processing complexity, and attending also that

	Wendelin	Hadoop
High-level programming language	Python	Java
Low-level language	C/C++/FORTRAN	N/A
Standard data structure	Numpy	N/A
Native x86 compiler	Numba	N/A
GPU compiler	Parakeet	N/A
Machine learning	Scikit-learn	Weka
Distributed storage	NEO	Spark
Distributed processing	ERP5 Activity	Job Tracker
Management portal	ERP5 Data	Cloudera Manager
Natural language processing	NLTK	Lucene
Video processing	OpenCV-python	N/A
Financial statistics	Pandas	N/A
Distributed index	MariaDB TokuDB Spider Sphinx	Solr
Cloud deployment and orchestration	SlapOS	Zookeeper

Table I
COMPARISON, GIVEN BY NEXEDI, OF WENDELIN VERSUS HADOOP.

	SlapOS	OpenStack	Kubernetes	Jupyter	NixOS	AWS
IaaS	✔	✔				✔
PaaS	✔		✔	✔	✔	✔
Service App Store	✔				✔	✔
Orchestration	✔		✔		✔	✔
Virtualisation	✔	✔				✔
Network management	✔	✔				✔
Resilient networking	✔					✔
Bare metal encapsulation	✔		✔		✔	✔
Portability	✔		limited			✔
Multi-tenant services (eg. CDN)	✔					✔
Edge computing	✔					✔
Continuous integration	✔					✔
Self-monitoring	✔					✔
Autonomous convergence	✔					?
Automated DR	✔					✔
Accounting & Billing	✔					✔

Table II
COMPARISON, GIVEN BY NEXEDI, OF SERVICES LIFECYCLE AUTOMATION (SLM) PLATFORMS.

in the transition of the training stage to inference stage many optimisations can be made as: weight pruning and model quantization. These and other optimizations can also be done during training phase as: gradient sparsification, gradient quantization and lightweight model design [17].

A weight ternary quantization for the inference phase was proposed in [18] and another at [19], for huge Large Language Models (LLM) (neural) networks, with a reduction in memory, latency and energy spent and with little impact or in some cases good impact on the accuracy.

Due to the increasing size and processing complexity of AI Models and the bulk amount of data to train them, the data and/or model must/may be split among several processing units in almost all cases being GPUs [17]. The data parallelism split consists of replicating the model, one in each processing unit, and training each one with exclusive subsets of the data [17]. The model parallelism split breaks apart the model through several processing units and provides all the data to each processing unit [17] . The split of the model can be vertical or horizontal division. In the vertical division of the AI network model, one or more layers of the network model are included in each processing unit. In the horizontal division all layers are included in a

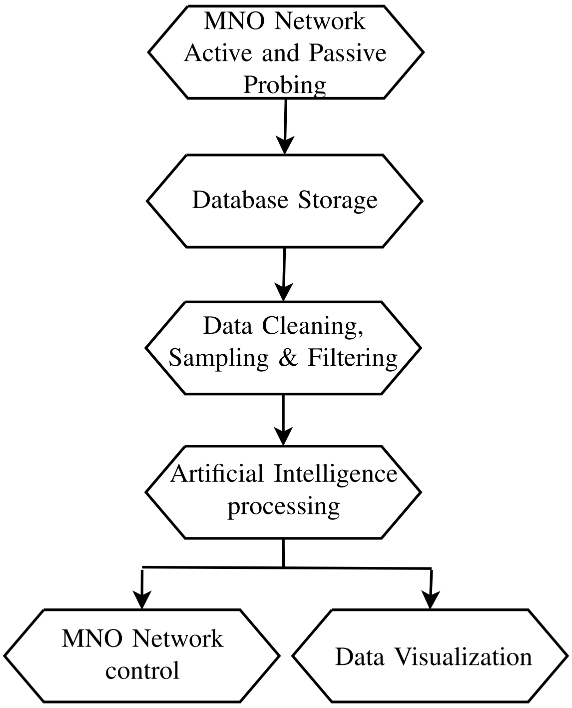

Figure 3. Operations of a typical modern OSS service.

processing unit but each layer is split over the processing units [17] . In addiction there is a hybrid parallelism split with both data and model parallelism [17] .

With the advent of end-edge-cloud computing there is a need to divide the processing and storage of the training or inference of the AI network through the end device (mobile device), edge server (*e.g.* router, basestation) and cloud server. Including training or inference in the Mobile Device can guarantee the privacy of the data.

It was claimed in [17], [20] that FPGA-based hardware is better suited for Deep Neural Network (DNN) acceleration than GPUs because of its specific customizability and non-uniform parallelism. In [20], it was claimed that FPGA outperforms GPUs in throughput parallelism and energy efficiency in edge computing since it enables Convolutional Neural Network (CNN) acceleration with arbitrary convolution window sizes.

In the context of 5G networks, one challenge of the use of Deep Reinforcement Learning (DLR) in network Slice Management is presented in [21] which can be cited: "most existing work is restricted to simulations; extensive evaluations are needed using a real-world testbed to judge the effectiveness of these DLR approaches in practice" and "DRL typically include Deep Neural Networks with fixed input and output dimensions based on the states and actions. ... when the conditions change (e.g., network topology change or change in length of Virtualized Network Functions chain), the input and output layers need to be changed to reflect the updated state and action space, and the new model needs to be re-trained from scratch.". The solution to deal with this variable input and output size in a Deep Reinforcement Learning Network is presented in [22] which uses neuron hot plugging, dynamically increases the size of the layers and performs incremental training.

Reinforcement Learning is motivated comes by the lack of availability of a large amount of *a priori* labeled training and testing data, which are difficult to obtain for real life scenarios, used in supervised Machine Learning (ML) methods. Deep Reinforcement Learning (DRL) has been used to solve sequential decision-making problems in wireless communications [23] through trial and error. Figure 4 shows a schematic of the concept of a Deep Reinforcement Network [24]. Basically, there is a DNN model (agent) in which, with the system state given by sensors and a reward, it is generated an action. This model mimics the optimization of an objective function to compute a minimal/maximal and it provides a value/action that applies to the environment to generate another reward and state with which another iteration begins (Figure 4). The agent chooses its next action with a higher probability or expected reward. After receiving the reward from the environment, the DNN parameters are updated accordingly. The generalization feature of the DNN permits the solution of high-dimensional discrete or continuous state space problems. Reinforcement Learning has two phases: *training* and *inference*. In the training phase, the agent interacts with the environment to collect experience. The environment is often a simulator that mimics a real-world system because it is expensive to directly interact with the real system [23]. DRL has advantages over traditional optimization methods due to its real-time inference. When training is complete the agent can make decisions in real-time.

Model-based Reinforcement Learning (MBRL) involves modeling the environment dynamics and utilizing the learned model to compute an optimal policy (agent). MBLR is more robust to changes in environmental dynamics or rewards and exhibits better exploratory behavior [23]. MBRL has been reported to have good sample efficiency[16]. However, in most applications, it is difficult to learn an accurate model of the world. When the optimal policy is found in a real environment, it is called Model Free Reinforcement Learning (MFRL). The Environment can be fully or partially observable giving rise to full or partial states, respectively.

Another approach is to run a Digital Twin agent in a Edge Node (Mobile Edge Computing (MEC)) that mimics the real Mobile System RAN to compute

[16]Measured in terms of the minimum samples needed to achieve a near-optimal policy.

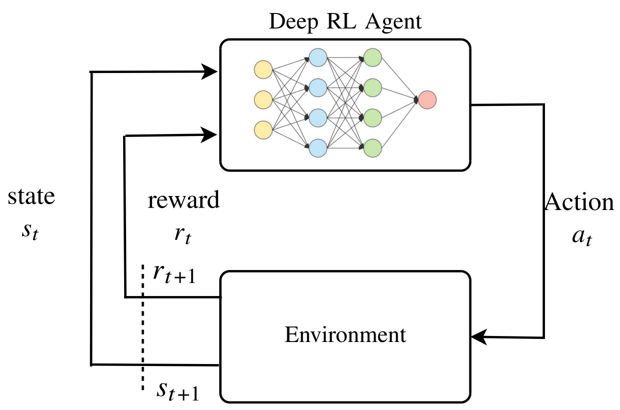

Figure 4. Sketch of the concept of a Deep Reinforcement Network

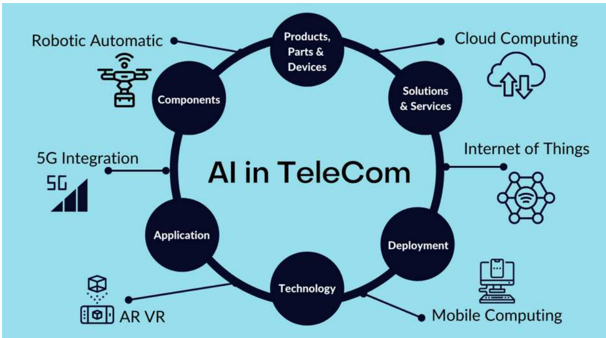

Figure 5. Options of AI integration in telecommunications.

improvements to deployment in the RAN [25] in order to increase profits.

The Extended Berkeley Packet Filter (eBDF) empowers programs to execute at the kernel execution level in Linux or Windows operating systems enabling direct kernel instrumentation in real time, plugin dynamically without needing to reboot and avoiding crashes [26]. It permits effective networking, tracing, and data profiling as observability and security, which can be potentiated by AI.

AI is used in a Mobile Virtual Network Operator (MVNO) to perform data usage prediction, data reselling optimization, and costumer churn profiling and mitigation [27]. Solving, by statistical analysis, of inaccurate billing and delayed billing state updates, was also reported.

Some outputs from the 5G-Xcast project, in which Ericsson is one of the partners, are stated in [28]. The use of AI for slice allocation and configuration with low latency was reported. In addiction, the use of AI with a DNN was reported for capacity forecast to better match better the capacity needed for each slice [29]. There is a good match between the forecast and the need of capacity with almost no capacity violations (underprovisioning).

Different possibilities of AI integration in the context of networks and telecommunications are shown in Figure 5 [30], [31].

The most recent advances in AI in telecommunications have introduced the ideas of LLM and generative AI. The AI paradigm in future telecommunication networks such as 6G will be reasoning driven, in which proactive decisions can be performed instead of data driven, which is training dependent and has limited generalization capacity [32]. The representation and transmission of model and no-model data relies on semantic content instead of the data content itself.

A good example is the ambitious objective in future 6G to integrate the known "AI-native" concept to integrate distributed and efficient AI in 6G components and architecture and offer the full power of AI to the extremely complex 6G architecture [33], [34]. The concept of AI-native has the objective of synthesizing the entire protocol stack, air interface, network core, etc. of the wireless system using AI techniques [32]. In other words, the AI-native concept consists of the transition from AI as a side feature in 5G to the core of both hardware and software in 6G [34].

V. The Actual Limits of Datasets Access in Telco Data

In the last sections, several articles have been cited using datasets. Below are only a few examples of them.

In IDSA report[17] some examples of data connectors and data market avenues are cited but lack of real telco data availability which limits the use and exploitation by future telco AI algorithms.

Due to the unavailability of datasets from Mobile Network Operators, privacy and competition concerns have resulted in the appearance of datasets available from simulators of wireless networks [35]–[37].

However, these simulations are limited to a few basestations and mobile phones, limiting the studies to local optimizations on the control plane through AI. These datasets, even in the future with greater land coverage, are not suitable for demographic studies. Although the simulator [38] uses physical Radio Frequency transceivers the propagation scenario is generated by Ray Tracing and the Doppler Frequency Offset is not emulated.

In [35], simulations on Colosseum, claimed to be the world's largest wireless network emulator, demonstrated closed-loop integration of real-time analytics and control through deep reinforcement learning agents. A dataset [39] representing network data for four Software Defined Radio (SDR) basestations and 40 SDR user equipment was generated. From the channel conditions, with

[17] https://internationaldataspaces.org/category/data-connector-report/

a pre-trained AI network, the best packet scheduler was chosen in order to achieve the maximum transmission rate.

A Deep Neural Network (DNN) for Internet Data Traffic forecasting with real geographic data, real-time data and real correspondent traffic logs was trained in [40]. It uses the dataset [41] of data traffic over time in a geographic location and also uses information on the location of the local base stations in [42].

In the sections above we cited the utilization of some datasets in telecommunications data analytics but also can be found in the repositories from Harvard University[18] and IEEE[19]. Datasets of many other subjects outside telecommunications can be found in the Harvard University repository. Many datasets used in the writing of articles are not in the public domain. Research on data analytics in telecommunications usually requires support from an MNO, WiFi provider, etc.

It is clear that the telco data and related datasets are very fragmented and heterogeneous efforts by different players. It requires a standardization roadmap for secure data interoperable access, and it requires a stable business model that will bind the telco data generators, operators, and service providers. This value chain is required for successful AI deployment in the context of telco data. One can cite the effort from the ITU Focus group on data processing and management[20] that describes the multi-dimensions related to secure trusted interoperable data.

VI. Conclusions

This article introduces three dimensions related to the rise of AI in telecommunication: network architecture evolution with cloud oriented approaches, such as telco cloud, AI models and potential use in the telco industry, data space and related limits of dataset availability, and secure data interoperability access issues. This includes citations of some of the latest bibliographies on the subject. This points to current limitations on the progress of the subject and the drawbacks of the technologies involved. Currently, it is a hot subject, with great expansion in the short and medium terms, with pair evolution with Artificial Intelligence application to network telco architectures to reduce complexity and open new avenues for innovations based on telco data and AI. Highly qualified human resources are needed not only because of the expertise required in each domain but also because the subject spans several domains.

[18] https://dataverse.harvard.edu/

[19] https://ieee-dataport.org/

[20] https://www.digitaldubai.ae/docs/default-source/publications/itu-t-focus-group-on-data-processing-and-management-framework_-19-07-2019.pdf

References

[1] L. M. P. Larsen, A. Checko, and H. L. Christiansen, "A Survey of the Functional Splits Proposed for 5G Mobile Crosshaul Networks," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 146–172, 2019. https://doi.org/10.1109/COMST.2018.2868805.

[2] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1376–1411, 2023. https://doi.org/10.1109/COMST.2023.3239220.

[3] S. Ebrahimi, F. Bouali, and O. C. L. Haas, "Resource Management From Single-Domain 5G to End-to-End 6G Network Slicing: A Survey," *IEEE Communications Surveys & Tutorials*, 2024. https://doi.org/10.1109/COMST.2024.3390613.

[4] E. Letouzé and J. Jütting, "Official Statistics, Big Data and Human Development," *Data-Pop Alliance*, March 2015. https://paris21.org/sites/default/files/WPS_OfficialStatistics_June2015.pdf.

[5] H. Liang, Z. Zhang, C. Hu, Y. Gong, and D. Cheng, "A Survey on Spatio-temporal Big Data Analytics Ecosystem: Resource Management, Processing Platform, and Applications," *IEEE Transactions on Big Data*, pp. 1–20, 2023. https://doi.org/10.1109/TBDATA.2023.3342619.

[6] D. S. Linthicum, "Cloud-Native Applications and Cloud Migration: The Good, the Bad, and the Points Between," *IEEE Cloud Computing*, vol. 4, no. 5, pp. 12–14, 2017. https://doi.org/10.1109/MCC.2017.4250932.

[7] P. Song, H. Peng, and X. Zhang, "A Micro-Service Approach to Cloud Native RAN for 5G and Beyond," *IEEE Access*, vol. 11, pp. 130257–130271, 2023. https://doi.org/10.1109/ACCESS.2023.3332964.

[8] F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber, "Bigtable: A Distributed Storage System for Structured Data," *ACM Transactions on Computer Systems*, vol. 26, no. 2, pp. 4:1–26, 2008. https://doi.org/10.1145/1365815.1365816.

[9] A. Okeleke, "CSPs can futureproof their analytics strategies by taking a platform-based approach to NWDAF implementation," May 2022. https://www.analysysmason.com/contentassets/131001974d2544089211c40025aae96f/analysys_mason_nwdaf_platform_implementation_may2022_rma14.pdf.

[10] UN Global Working Group on Big Data for Official Statistics, "Handbook on the Use of Mobile Phone Data for Official Statistics," September 2019. https://unstats.un.org/bigdata/task-teams/mobile-phone/MPD%20Handbook%2020191004.pdf.

[11] K. L.-M. Ang, F. L. Ge, and K. P. Seng, "Big Educational Data & Analytics: Survey, Architecture and Challenges," *IEEE Access*, vol. 8, pp. 116392–116414, 2020. https://doi.org/10.1109/ACCESS.2020.2994561.

[12] H. Zahid, T. Mahmood, A. Morshed, and T. Sellis, "Big Data Analytics in Telecommunications: Literature Review and Architecture Recommendations," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 1, pp. 18–38, 2020. https://doi.org/10.1109/JAS.2019.1911795.

[13] "DIN 27070 IDSA framework," https://internationaldataspaces.org/ids-is-officially-a-standard-din-spec-27070-is-published/.

[14] "Industry 4.0 the OPC-UA framework," https://opcfoundation.org/wp-content/uploads/2014/03/OPC_UA_I_4.0_Pioneer_US_v2.pdf.

[15] "ETSI related to 3GPP.," https://www.etsi.org/deliver/etsi_ts/129200_129299/129222/16.08.00_60/ts_129222v160800p.pdf.

[16] "SimpleRan project page.," https://www.nexedi.com/web_page_module/12634/view.

[17] S. Duan, D. Wang, J. Ren, F. Lyu, Y. Zhang, H. Wu, and X. Shen, "Distributed Artificial Intelligence Empowered by End-Edge-Cloud Computing: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 591–624, 2023. https://doi.org/10.1109/COMST.2022.3218527.

[18] S. Ma, H. Wang, L. Ma, L. Wang, W. Wang, S. Huang, L. Dong, R. Wang, J. Xue, and F. Wei, "The Era of 1-bit LLMs: All Large Language Models are in 1.58 Bits," *arXiv:2402.17764*, 17th February 2024. https://doi.org/10.48550/arXiv.2402.17764.

[19] C. Zhu, S. Han, H. Mao, and W. J. Dally, "Trained Ternary Quantization," *arXiv:1612.01064*, 23rd February 2017. https://doi.org/10.48550/arXiv.1612.01064.

[20] L. Du, Y. Du, Y. Li, J. Su, Y.-C. Kuan, C.-C. Liu, and M.-C. F. Chang, "A Reconfigurable Streaming Deep Convolutional Neural Network Accelerator for Internet of Things," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 1, pp. 198–208, 2018. https://doi.org/10.1109/TCSI.2017.2735490.

[21] N. Saha, M. Zangooei, M. Golkarifard, and R. Boutaba, "Deep Reinforcement Learning Approaches to Network Slice Scaling and Placement: A Survey," *IEEE Communications Magazine*, vol. 61, no. 2, pp. 82–87, 2023. https://doi.org/10.1109/MCOM.006.2200534.

[22] R. Han, D. Chen, S. Guo, J. Wang, Q. Qi, L. Lu, and J. Liao, "Multi-SP Network Slicing Parallel Relieving Edge Network Conflict," *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, no. 11, pp. 2860–2875, 2023. https://doi.org/10.1109/TPDS.2023.3310013.

[23] A. Feriani and E. Hossain, "Single and Multi-Agent Deep Reinforcement Learning for AI-Enabled Wireless Networks: A Tutorial," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1226–1252, 2021. https://doi.org/10.1109/COMST.2021.3063822.

[24] C. Ssengonzi, O. P. Kogeda, and T. O. Olwal, "A survey of deep reinforcement learning application in 5G and beyond network slicing and virtualization," *Array*, vol. 14, p. 100142, 2022. https://doi.org/10.1016/j.array.2022.100142.

[25] X. Lin, J. Wu, J. Li, W. Yang, and M. Guizani, "Stochastic Digital-Twin Service Demand With Edge Response: An Incentive-Based Congestion Control Approach," *IEEE Transactions on Mobile Computing*, vol. 22, no. 4, pp. 2402–2416, 2023. https://doi.org/10.1109/TMC.2021.3122013.

[26] D. Soldani, "eBPF for Modern Telco Infrastructure – A New Approach to Observability, Networking and Security Tooling," February 24, 2024. https://cdn.prod.website-files.com/6317e170a9eabbe0fbbf4519/65d9ab15c056d9ec592b919e_01%20RakutenSymphony_EBF-for-Telco_Whitepaper_R6_02182024_FINAL.pdf.

[27] Y. Li, J. Zheng, Z. Li, Y. Liu, F. Qian, S. Bai, Y. Liu, and X. Xin, "Understanding the Ecosystem and Addressing the Fundamental Concerns of Commercial MVNO," *IEEE/ACM Transactions on Networking*, vol. 28, no. 3, pp. 1364–1377, 2020. https://doi.org/10.1109/TNET.2020.2981514.

[28] A. Banchs, D. M. Gutierrez-Estevez, M. Fuentes, M. Boldi, and S. Provvedi, "A 5G Mobile Network Architecture to Support Vertical Industries," *IEEE Communications Magazine*, vol. 57, no. 12, pp. 38–44, 2019. https://doi.org/10.1109/MCOM.001.1900258.

[29] D. Bega, M. Gramaglia, M. Fiore, A. Banchs, and X. Costa-Pérez, "DeepCog: Optimizing Resource Provisioning in Network Slicing With AI-Based Capacity Forecasting," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 2, pp. 361–376, 2020. https://doi.org/10.1109/JSAC.2019.2959245.

[30] "Options of AI integration in telecommunications.," https://www.openpr.com/news/3295387/ai-in-telecommunication-market-may-see-a-big-move-google.

[31] Y. Sanjalawe, S. Fraihat, S. Al-E'mari, M. Abualhaj, S. Makhadmeh, and E. Alzubi, "A Review of 6G and AI Convergence: Enhancing Communication Networks With Artificial Intelligence," *IEEE Open Journal of the Communications Society*, 2025. https://doi.org/10.1109/OJCOMS.2025.3553302.

[32] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. Vincent Poor, "Less Data, More Knowledge: Building Next-Generation Semantic Communication Networks," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 1, pp. 37–76, 2025. https://doi.org/10.1109/COMST.2024.3412852.

[33] "AI native Ericsson white paper," https://www.ericsson.com/en/reports-and-papers/white-papers/ai-native.

[34] S. Tarkoma, R. Morabito, and J. Sauvola, "AI-native Interconnect Framework for Integration of Large Language Model Technologies in 6G Systems," 2023. https://arxiv.org/abs/2311.05842.

[35] L. Bonati, S. D'Oro, M. Polese, S. Basagni, and T. Melodia, "Intelligence and Learning in O-RAN for Data-Driven NextG Cellular Networks," *IEEE Communications Magazine*, vol. 59, no. 10, pp. 21–27, 2021. https://doi.org/10.1109/MCOM.101.2001120.

[36] D. Villa, M. Tehrani-Moayyed, C. P. Robinson, L. Bonati, P. Johari, M. Polese, and T. Melodia, "Colosseum as a Digital Twin: Bridging Real-World Experimentation and Wireless Network Emulation," 2024. https://doi.org/10.48550/arXiv.2303.17063.

[37] L. Bonati, P. Johari, M. Polese, S. D'Oro, S. Mohanti, M. Tehrani-Moayyed, D. Villa, S. Shrivastava, C. Tassie, K. Yoder, A. Bagga, P. Patel, V. Petkov, M. Seltser, F. Restuccia, A. Gosain, K. R. Chowdhury, S. Basagni, and T. Melodia, "Colosseum: Large-Scale Wireless Experimentation Through Hardware-in-the-Loop Network Emulation," 2021. https://doi.org/10.48550/arXiv.2110.10617.

[38] A. Chaudhari and M. Braun, "A Scalable FPGA Architecture for Flexible, Large-Scale, Real-Time RF Channel Emulation," in *2018 13th International Symposium on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC).* https://doi.org/10.1109/ReCoSoC.2018.8449390.

[39] "Colosseum Dataset for Communications Magazine Article," https://github.com/wineslab/colosseum-oran-commag-dataset.

[40] C.-W. Huang and P.-C. Chen, "Joint Demand Forecasting and DQN-Based Control for Energy-Aware Mobile Traffic Offloading," *IEEE Access*, vol. 8, pp. 66588–66597, 2020. https://doi.org/10.1109/ACCESS.2020.2985679.

[41] G. Barlacchi, M. De Nadai, R. Larcher, A. Casella, C. Chitic, G. Torrisi, F. Antonelli, A. Vespignani, A. Pentland, and B. Lepri, "A multi-source dataset of urban life in the city of Milan and the Province of Trentino," *Scientific Data*, vol. 2, p. 150055, Oct 2015. https://doi.org/10.1038/sdata.2015.55.

[42] "OpenCellID," Constantly updated. https://opencellid.org/.