Toward Experimentation-as-a-Service in 5G/6G: The Plaza6G Prototype for AI-Assisted Trials

Sergio Barrachina-Muñoz, Marc Carrascosa-Zamacois, Horacio Bleda, Umair Riaz,
Yasir Maqsood, Xavier Calle, Selva Vía, Miquel Payaró, Josep Mangues-Bafalluy
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), Barcelona, Spain
{sbarrachina, mcarrascosa, hbleda, uriaz, ymaqsood, xcalle, svia, mpayaro, jmangues}@cttc.cat

***Abstract*—This paper presents Plaza6G, the first operational *Experiment-as-a-Service* (ExaS) platform unifying cloud resources with next-generation wireless infrastructure. Developed at CTTC in Barcelona, Plaza6G integrates GPU-accelerated compute clusters, multiple 5G cores, both open-source (e.g., Free5GC) and commercial (e.g., Cumucore), programmable RANs, and physical or emulated user equipment under unified orchestration. In Plaza6G, the experiment design requires minimal expertise as it is expressed in natural language via a web portal or a REST API. The web portal and REST API are enhanced with a Large Language Model (LLM)-based assistant, which employs retrieval-augmented generation (RAG) for up-to-date experiment knowledge and Low-Rank Adaptation (LoRA) for continuous domain fine-tuning. Over-the-air (OTA) trials leverage a four-chamber anechoic facility and a dual-site outdoor 5G network operating in sub-6 GHz and mmWave bands. Demonstrations include automated CI/CD integration with sub-ten-minute setup and interactive OTA testing under programmable propagation conditions. Machine-readable experiment descriptors ensure reproducibility, while future work targets policy-aware orchestration, safety validation, and federated testbed integration toward open, reproducible wireless experimentation.**

***Index Terms*—Experiment-as-a-service (ExaS), network automation, wireless testbeds, 5G, 6G, LLM**

I. Introduction

The increasing complexity of next-generation (xG) wireless networks necessitates experimentation environments that transcend simulation and theoretical modeling. Cloud-native architectures, software-defined networking, and programmable radio interfaces are fundamentally reshaping wireless service design and validation. However, integrating realistic radio conditions with scalable computing resources remains challenging. Our previous work [1] built upon the *Experimentation-as-a-Service* (ExaS) paradigm [2], [3] enabling on-demand, reproducible, and automated experimentation by embedding xG testbeds directly into Continuous Integration and Continuous Deployment (CI/CD) pipelines. While that work established the vision and architectural principles, this paper presents the first operational realization of ExaS via the *Plaza6G* platform.

Developed at CTTC in Barcelona, *Plaza6G* implements ExaS by unifying cloud computing flexibility with real 5G and 6G experimental infrastructure (Fig. 1). The platform provisions bare-metal servers, virtual machines/instances, Kubernetes clusters, GPUs, and xG radio and core networks through unified orchestration. Users interact via a web portal or API enhanced with a large-language-model (LLM)-based interface that interprets natural-language requests and automatically configures compute and wireless resources. This design enables users with minimal domain expertise to deploy complex experiments immediately, eliminating manual scripting and configuration barriers.

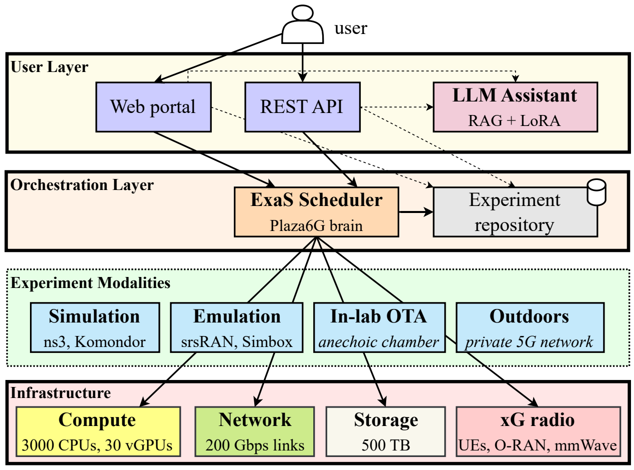

Fig. 1: Conceptual architecture of Plaza6G.

Plaza6G represents, to our knowledge, the first platform merging cloud-scale resources with real wireless infrastructure under an AI-assisted ExaS model. Existing 5G/6G testbeds typically require manual intervention, present fragmented interfaces, and demand specialized expertise, constraining accessibility and hindering reproducibility. Plaza6G addresses these limitations through integrated automation and intelligent orchestration.

This paper demonstrates Plaza6G's current capabilities through two representative use cases. First, the API is used to launch in parallel three emulated 5G networks, enabling automated CI/CD network acceptance testing as part of a software developer's pipeline, where applications under development are validated concurrently across different 5G core implementations. Second, a real over-the-air (OTA) experiment is executed via the web portal, in which a user requests the deployment of a complete 5G network with a physical user equipment (UE) and a gNB. These concurrent experiments highlight Plaza6G's scalability, reproducibility, and multi-tenant isolation capabilities.

This work was partially funded by Spanish MINECO grants TSI-064100-2022-16/-2023-26 (Plaza6G/Plaza6G+) and Grant PID2021-126431OB-I00 (ANEMONE) funded by MCIN/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe"

II. The Plaza6G Platform and Resources

A. Architecture and Experiment Modalities

Building upon the *Experimentation-as-a-Service* (ExaS) model introduced in [1], *Plaza6G* represents its first operational instantiation—a multi-domain environment enabling automated wireless experimentation across four distinct modalities with varying fidelity levels.

The platform adopts a three-layer architecture: the *user layer* provides web portal and REST API access augmented by an LLM-based interface interpreting intents into executable experiment graphs; the *orchestration layer* coordinates resource allocation through policy-driven scheduling; and the *infrastructure layer* exposes heterogeneous compute, network, and radio assets as composable services.

User interaction combines graphical workflow composition with natural-language processing via an LLM-based assistant. The LLM backend, deployed locally at CTTC, leverages retrieval-augmented generation (RAG) and Low-Rank Adaptation (LoRA) [4] to improve technical dialogue accuracy and orchestration safety while maintaining low latency. Fig. 2 shows the portal where natural-language intents are transformed into executable workflows.

Plaza6G supports four experimentation modalities under unified orchestration: **(i) Simulation** employs discrete-event simulators (*ns-3* [5], *Komondor* [6]) for large-scale protocol evaluation; **(ii) Emulation** instantiates virtualized protocol stacks (*UERANSIM* or srsRAN UE/gNB) for rapid multi-configuration benchmarking; **(iii) In-lab** integrates physical equipment (e.g., *Amarisoft Callbox*, commercial UEs) within a four-chamber anechoic facility for repeatable OTA validation; and **(iv) Outdoors** leverages a dual-site outdoor 5G network (sub-6 GHz, mmWave) for end-to-end trials under realistic conditions. These modalities enable progressive validation from simulation through field deployment within the same framework.

B. Infrastructure and Resources

Plaza6G infrastructure spans three integrated technological domains supporting all experiment modalities. The *compute domain* provides GPU-accelerated clusters hosting virtual machines and Kubernetes workloads with support for edge-cloud continuum deployment. Current capacity exceeds 3,000 CPU cores, 30 vGPUs (NVIDIA L40S), and approximately 500 TB of storage, enabling concurrent execution of simulation, emulation, and virtualized network function workloads. The *network domain* offers multiple 5G core implementations such as *Free5GC*, *Open5GS*, or *Cumucore*, supporting end-to-end network slicing and service isolation across all experiment types. The *radio domain* comprises programmable RAN platforms including *Amarisoft Callbox*, O-RAN, and srsRAN, alongside both emulated user equipment (*UERANSIM*, *Amarisoft Simbox*) and physical devices (commercial Android smartphones). Controlled OTA testing leverages the four-chamber anechoic facility ensuring repeatable radio propagation conditions for in-lab experiments, while outdoor infrastructure supports field trials as described below.

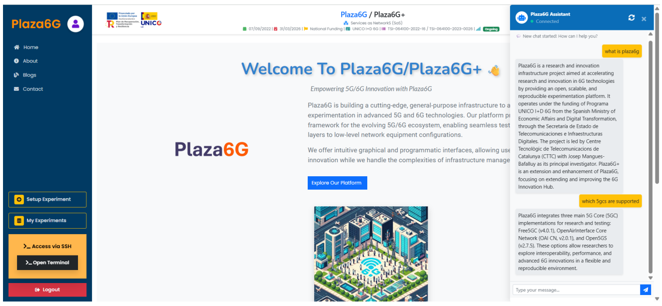

Fig. 2: Plaza6G web portal with the LLM assistant.

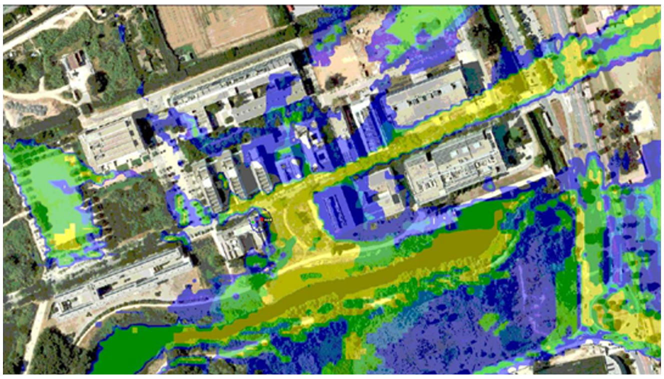

Fig. 3: Coverage of the Plaza6G outdoor private 5G network at the PMT campus, operating across sub-6 GHz and mmWave bands. The color scale indicates received signal power, from yellow (strong) to dark blue (weak).

Plaza6G extends laboratory capabilities through a dual-site outdoor 5G network deployed across rooftop installations at the *Parc Mediterrani de la Tecnologia* campus. Operating in both sub-6 GHz and mmWave bands, this network supports configuration with open-source or commercial 5G cores, enabling field experimentation under realistic propagation, interference, and mobility conditions. In particular, the two outdoor radio sites can also be re-configured so that they belong to two independent 5G networks to test, e.g., roaming scenarios. Fig. 3 presents measured coverage, demonstrating stable connectivity across the campus area served by the rooftop antenna deployment.

C. Positioning Against Existing Testbeds

Large-scale European initiatives including 5GENESIS [7], 5G-VINNI [8], VITAL-5G [9], and 5G-EVE [10] have established comprehensive validation infrastructures for 5G technologies. Recent efforts such as 6G-SANDBOX [11] target early 6G experimentation, while *Joiner* [12] federates 11 UK testbeds with automation capabilities, and the IEEE 5G/6G Innovation Testbed pursues end-to-end 3GPP-compliant CI/CD integration (currently under development). However, these platforms predominantly rely on predefined workflows requiring manual configuration, constraining both accessibility and automation potential.

Plaza6G differentiates itself through three key innovations: (i) unified orchestration spanning simulation, emulation, controlled in-lab, and field experiment modalities within a single platform, (ii) AI-assisted zero-touch experimentation accessible to users without domain-specific scripting expertise, and (iii) native support for the *Experiment-as-Code* (ExaC) paradigm [2], [3], [13], enabling fully reproducible, version-controlled experiment definitions deployable via natural language or programmatic APIs. This combination of elements positions Plaza6G as a next-generation platform bridging the gap between cloud elasticity and realistic wireless experimentation across the complete validation spectrum.

III. Demonstration Use Cases

To validate Plaza6G's operational capabilities, two representative scenarios are presented: (i) automated API-driven validation integrated with external CI/CD pipelines, and (ii) controlled OTA experimentation with physical 5G equipment. These use cases highlight Plaza6G's ability to embed network acceptance testing within software workflows while also supporting reproducible OTA and field experiments across sub-6 GHz and mmWave environments.

A. Emulated Use Case: CI/CD Network Acceptance via API

This scenario demonstrates Plaza6G as an automated validation stage within continuous integration workflows. A CI/CD pipeline triggers experimentation via a REST API call expressed in natural language. For instance, to validate application performance across different 5G core implementations, a user can submit:

```
{
  "user_request": "Deploy <my_app> across
  ↪  three 5G cores (Open5GS, Free5GC,
  ↪  OAI-CN) and verify <my_kpi> exceeds
  ↪  threshold for test approval."
}
```

For the sake of simplicity, in this demonstration, `my_app` is instantiated as *iperf3* and `my_kpi` is mean throughput with an acceptance test threshold of 50 Mbps. The LLM backend interprets the request, identifies the application under test, target cores, and success criteria, then generates an experiment plan. The system returns one of three responses (*approved*, *clarification required*, or *denied*) based on resource availability and policy constraints. CI/CD pipelines may proceed automatically upon approval or incorporate human-in-the-loop review, depending on organizational trust policies.

Each experiment instantiates a complete emulated 5G system comprising *UERANSIM*-based UE and gNB connected to one of the three 5G core implementations. A dedicated Data Network Name (DNN) virtual machine hosts the application server (*iperf3 -s*), while the emulated UE executes *iperf3 -c* for 2 minutes, generating TCP and UDP traffic for benchmarking. The ExaS scheduler manages the complete lifecycle (resource allocation, UE/gNB/5GC/DNN instantiation, measurement collection, and teardown) through an asynchronous scheduler. All three experiments execute concurrently on isolated compute pools, with comprehensive telemetry such as throughput, latency, and CPU utilization automatically archived.

Figure 4 depicts representative throughput distributions across the three 5G cores. All measured mean values exceed

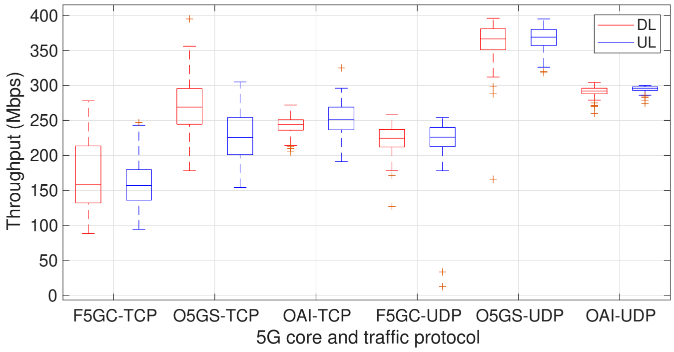

Fig. 4: Throughput performance comparison for both UDP and TCP traffic across three 5G core implementations executed concurrently via Plaza6G API. Tests duration of 120 seconds, with throughput logged every second.

the specified 50 Mbps threshold, demonstrating consistent performance for user's app under test. From a DevOps perspective, this enables "network acceptance testing" where CI/CD pipelines advance to staging or production only after minimum KPI thresholds are satisfied. Measured setup time remains below ten minutes per experiment, reducing configuration effort by over an order of magnitude compared to manual procedures.

B. In-lab use case: Over-the-Air Controlled Experiment

The second scenario illustrates Plaza6G's capability to automate the provisioning of complex physical experimentation environments while deliberately supporting human-in-the-loop experimentation. Unlike the CI/CD use case we discussed before, which targets fully automated and script-driven validation, this scenario is designed for exploratory and interactive experimentation, where users manually access network elements and conduct measurements without predefined execution scripts.

The experiment provisioning phase is fully automated. Using the Plaza6G web portal and its LLM assistant, the user selects an experiment template that deploys a *Free5GC* core, an *Amarisoft Callbox* gNB, and a commercial Android smartphone acting as UE. Templates expose a curated set of commonly used 5G parameters (e.g., 100 MHz bandwidth, MIMO) to simplify initial configuration, while allowing users to override or refine parameters either manually or through the LLM assistant. The gNB and UE are placed in separate chambers of the four-chamber anechoic facility, enabling programmable control of path loss, attenuation profiles, and interference conditions. Figure 5 shows the physical OTA experiment setup. The scheduler automatically provisions compute, core, and radio resources and establishes end-to-end connectivity, after which the environment is handed over to the user for interactive experimentation.

Once the setup is complete, the user remotely accesses the UE via *Vysor* to manually execute application-level tests, install software, or explore network behavior under controlled radio conditions. By progressively adjusting inter-chamber

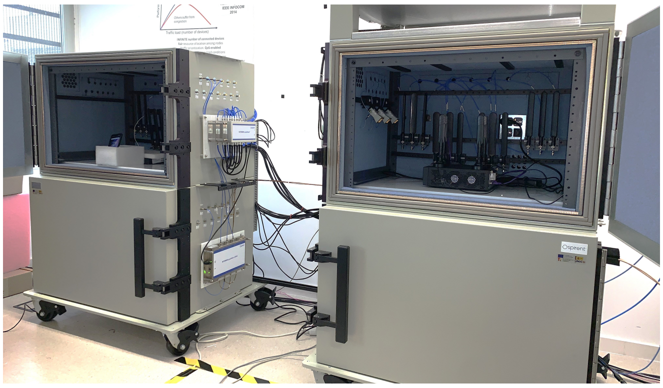

Fig. 5: Anechoic chamber with an Android UE and an Amarisoft Callbox for controlled OTA experimentation.

attenuation, users can interactively study the impact of channel degradation on throughput, latency, and perceived quality of experience. This mode of operation supports exploratory studies such as adaptive streaming behavior, application robustness to radio impairments, and edge–cloud service performance under varying link quality, without constraining the experiment to predefined workflows.

Throughout the session, the LLM assistant can provide contextual information on current signal conditions, active configurations, and runtime statistics in natural language, assisting users in interpreting observations while retaining full control over experiment execution.

IV. Discussion and Outlook

A. Reproducibility and Workflow Traceability

All Plaza6G experiments follow a unified orchestration workflow in which every action, from initial resource provisioning to final teardown, is logged as a machine-readable descriptor capturing software versions, network topology, hardware identifiers, and configuration parameters. These descriptors are archived in a searchable experiment repository that supports version control and longitudinal performance tracking across multiple runs.

In addition to local traceability, Plaza6G maintains a consistent naming and indexing scheme that links experiment descriptors, orchestration logs, and collected metrics. This structured organization enhances repeatability while allowing users to audit the full lifecycle of an experiment, from creation to completion, through a unified web interface. To foster transparency and community reproducibility, the detailed procedure for replicating the emulated use case presented in Section III-A has been published on the *protocols.io* platform. This companion protocol demonstrates the same workflow executed via the Plaza6G web portal rather than the API, providing additional insight into the graphical interface and user interaction process.[1]

[1]The complete experimental protocol, including setup, execution, and data collection steps, is available at https://www.protocols.io/view/plaza6g-experiment-reproduction-protocol-use-case-a-dm6gpm6pjgzp/v1.

B. Conclusions and future work

The demonstrations in Section III validate *Plaza6G* as a practical realization of the ExaS paradigm, unifying radio, core, and compute infrastructure under automated orchestration. LLM-assisted interfaces and programmable APIs reduce the expertise required for wireless experimentation, enabling reproducible, concurrent trials with setup times below ten minutes. By coupling data-center automation with xG infrastructure, Plaza6G turns network experimentation into an on-demand cloud service for developers, researchers, and vendors.

Future work will extend Plaza6G along several directions. (1) Policy-aware orchestration will optimize scheduling for cost, energy, and radio resource usage. (2) Safety and validation mechanisms will verify LLM-generated actions to ensure correctness and reproducibility. (3) Planned LoRA-based fine-tuning will incrementally refine the LLM using selected orchestration logs and user dialogues, improving technical accuracy without full-model retraining. (4) Federated operation with external testbeds is envisioned to enable geographically distributed, multi-domain experiments; integration with ETSI OpenSlice [14] is under study to align Plaza6G with emerging open orchestration standards.

References

[1] S. Barrachina-Muñoz, H. Bleda, M. Requena, S. Vía, M. Payaró, and J. Mangues-Bafalluy, "Experiment-as-a-Service in the Pipeline: Empowering CI/CD with xG Acceptance Testing," in *Wireless On-Demand Netw. Syst. and Serv. Conf. (WONS)*. IEEE, 2025, pp. 1–4.

[2] T. W. Edgar and T. R. Rice, "Experiment as a service," in *2017 IEEE Int. Symp. on Tech. for Homeland Security (HST)*. IEEE, 2017.

[3] M. Boniface *et al.*, "BonFIRE: A Multi-Cloud Experimentation-as-a-Service Ecosystem," in *Building the Future Internet through FIRE*. River Publishers, 2022, pp. 243–266.

[4] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen *et al.*, "LoRA: Low-rank adaptation of large language models." *ICLR*, vol. 1, no. 2, p. 3, 2022.

[5] A. Larrañaga *et al.*, "An open-source implementation and validation of 5G NR configured grant for URLLC in ns-3 5G LENA: A scheduling case study in industry 4.0 scenarios," *Journal of Network and Computer Applications*, vol. 215, p. 103638, 2023.

[6] S. Barrachina-Muñoz, F. Wilhelmi, I. Selinis, and B. Bellalta, "Komondor: A wireless network simulator for next-generation high-density WLANs," in *2019 Wireless Days (WD)*. IEEE, 2019, pp. 1–8.

[7] G. Xylouris *et al.*, "Experimentation and 5G KPI measurements in the 5GENESIS platforms," in *Proceedings of the 1st Workshop on 5G Measurements, Modeling, and Use Cases*, 2021, pp. 1–7.

[8] A. J. Gonzalez, M. Xie, P. H. Lehne, and P. Grønsund, "Achieving high throughput and low latency with 5G: A real implementation experience," *IEEE Communications Magazine*, vol. 59, no. 10, pp. 84–90, 2021.

[9] V. Charpentier *et al.*, "Utilizing the VITAL-5G platform to advance 5G standalone integration with vertical industries," *EURASIP Journal on Wireless Comm. and Networking*, vol. 2025, no. 1, p. 27, 2025.

[10] M. Gupta *et al.*, "The 5G EVE end-to-end 5G facility for extensive trials," in *2019 ICC workshops*. IEEE, 2019, pp. 1–5.

[11] P. Merino-Gomez, B. Garcia, C. Andreo, D. Artuñedo, and J. Macias, "On-demand Trial Networks over 6G-SANDBOX infrastructure," in *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2024, pp. 1021–1026.

[12] S. Saunders, O. Dilley, and D. Simeonidou, "JOINER-NSF: UK's national facility for spectrum access innovation," in *IEEE Intl. Symp. on Dynamic Spectrum Access Netw. (DySPAN)*. IEEE, 2025, pp. 77–78.

[13] L. Aguilar *et al.*, "Experiments as Code and its application to VR studies in human-building interaction," *Scientific Reports (Nature Portfolio)*, vol. 14, no. 1, p. 9883, 2024.

[14] ETSI, "ETSI OpenSlice: Open Source Platform for Service Orchestration and Management," https://osl.etsi.org/, accessed: Oct. 24, 2025.