

# AI-O-RAN Convergence: From Foundational LLMs to Automated System Testing

**Vijay K. Shah, PhD**

Assistant Professor, ECE, NC State University

Director, NextG Wireless Lab (NextG Lab@NC State)

Community Outreach Director, AERPAW

Co-Founder, WiSights Lab and All Things Intelligence Inc.

OAI Workshop  
Nov. 17-19, 2025

**NC STATE**



NextG  
Wireless Lab



AERPAW

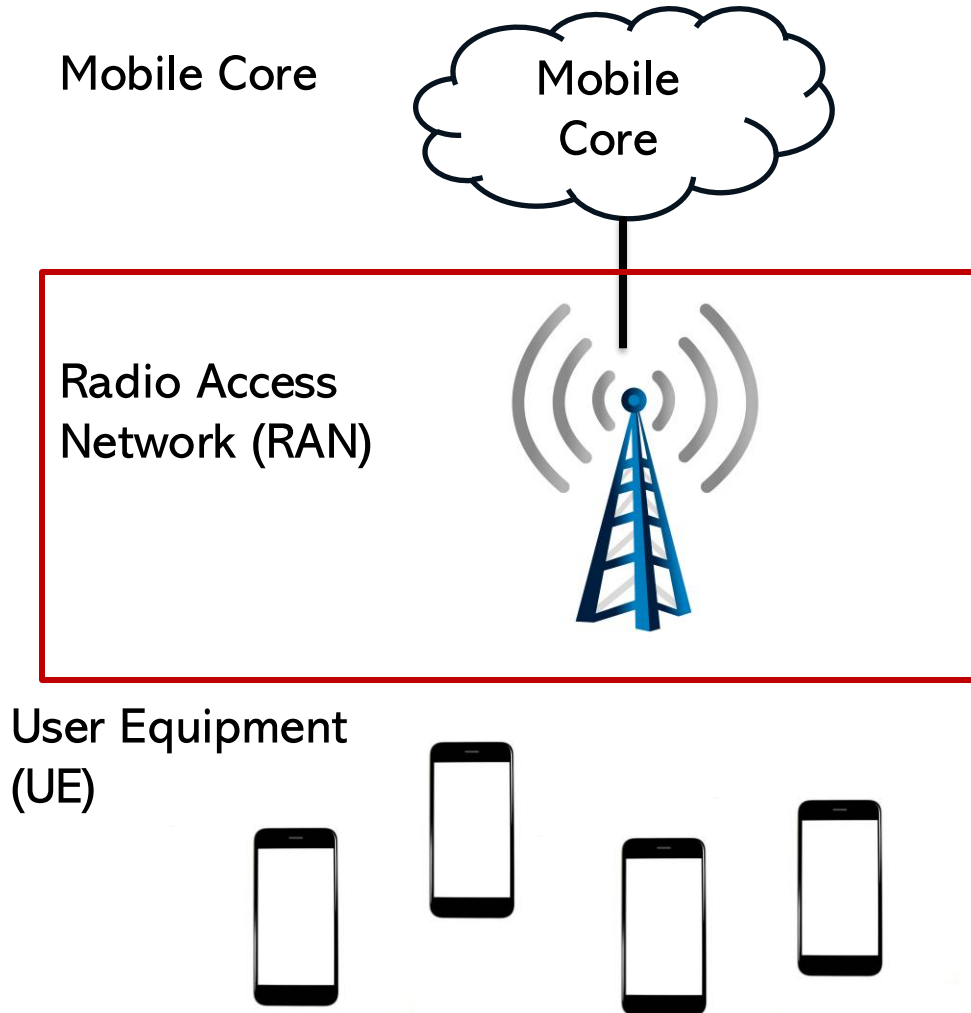


WiSights **Lab**



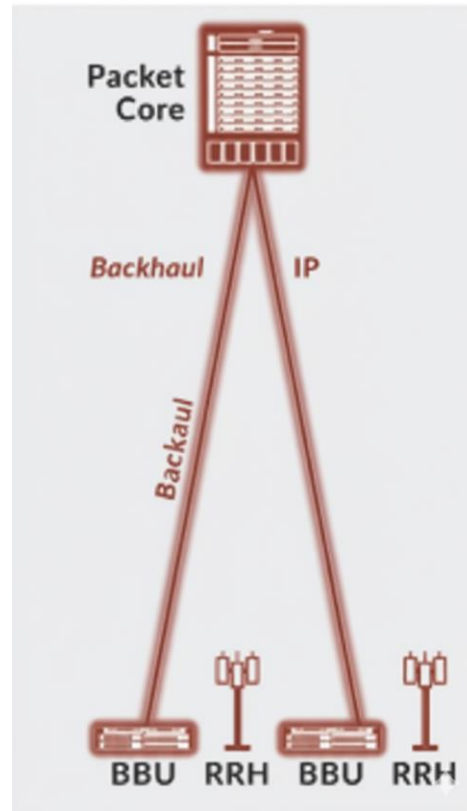
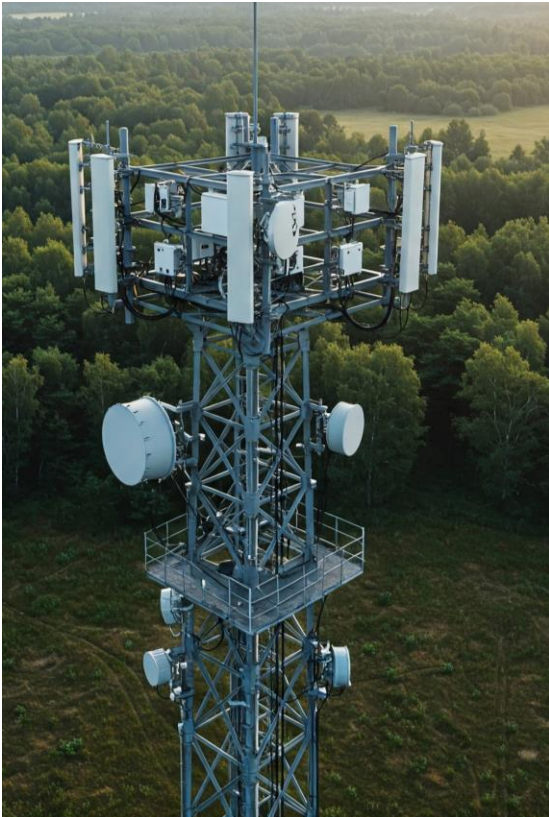
ALL THINGS  
INTELLIGENCE

# Reference Telecom Architecture



# Let's look the evolution of RAN...

## Legacy RAN (2G, 3G)



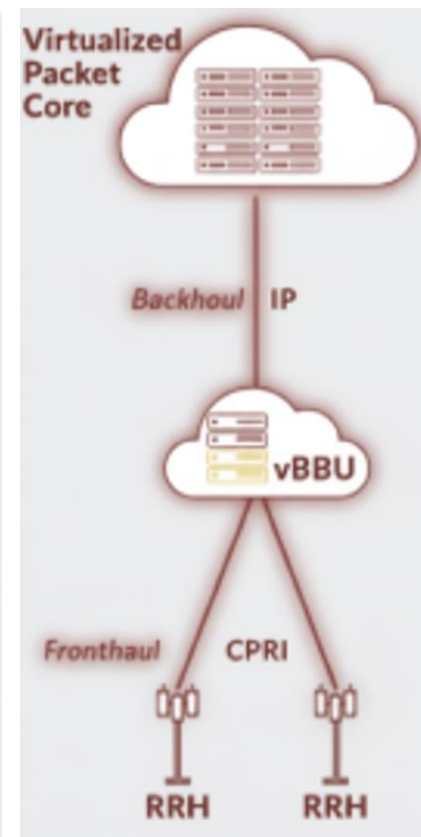
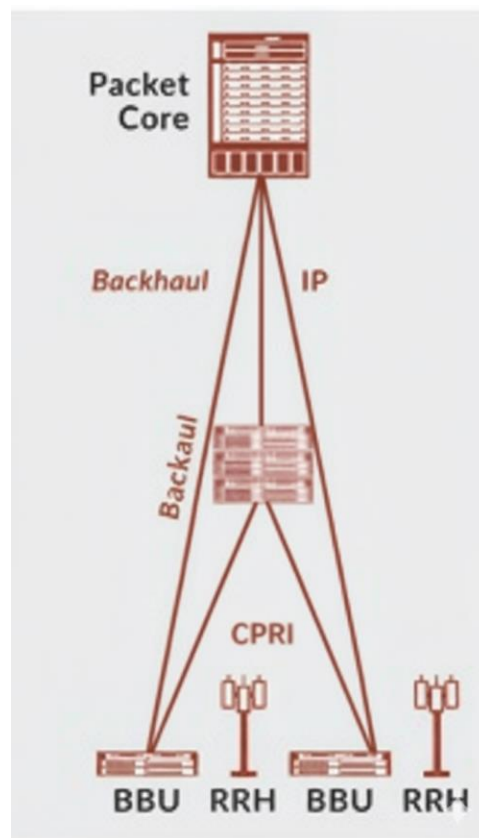
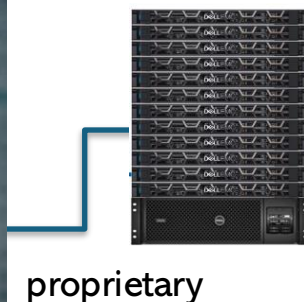
**Inflexible**

**Monolithic**

**Proprietary**

# Let's look the evolution of RAN...

C-RAN/V-RAN (Mostly the case today)



Limited  
Flexibility

Softwarization

Virtualization

Disaggregated  
(RRH+BBU)

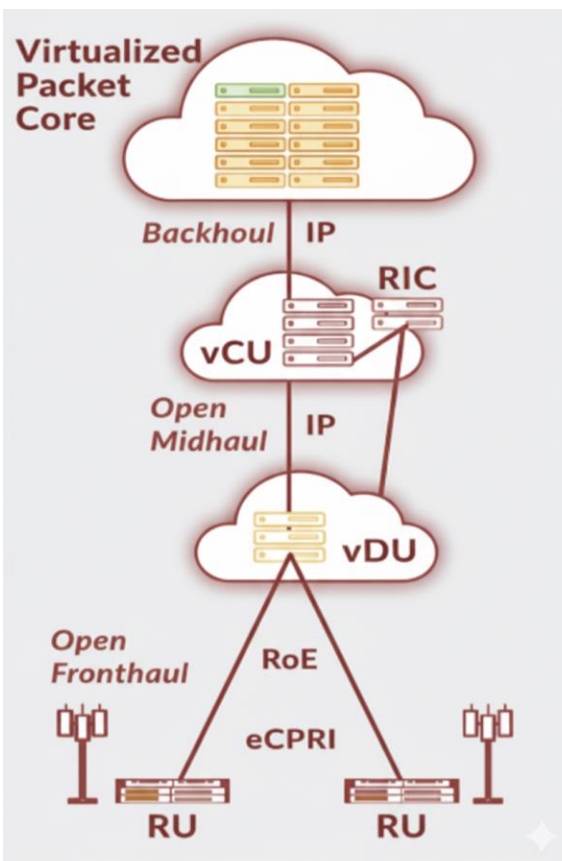
Proprietary

# Let's look the evolution of RAN...

## O-RAN (Early 5G trial deployments)



Open  
Interfaces



Intelligence

Flexibility

Softwarization

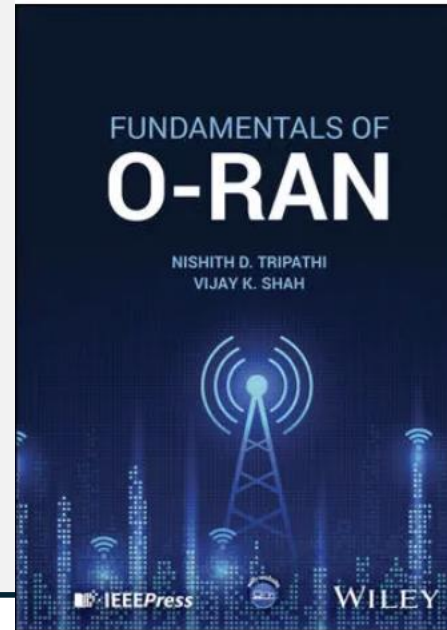
Virtualization

Disaggregated  
(O-RU+O-DU+O-CU)

Open

# Quick Peek: O-RAN Architecture

- O-RAN Alliance's mission is to re-shape the RAN industry towards more intelligent, open, virtualized interoperable mobile networks
- Based on two core principles
  - **Openness**
  - **Intelligence**
- Flexibility by design
  - **Open interfaces** and **APIs**
- 3GPP Split 7-2x



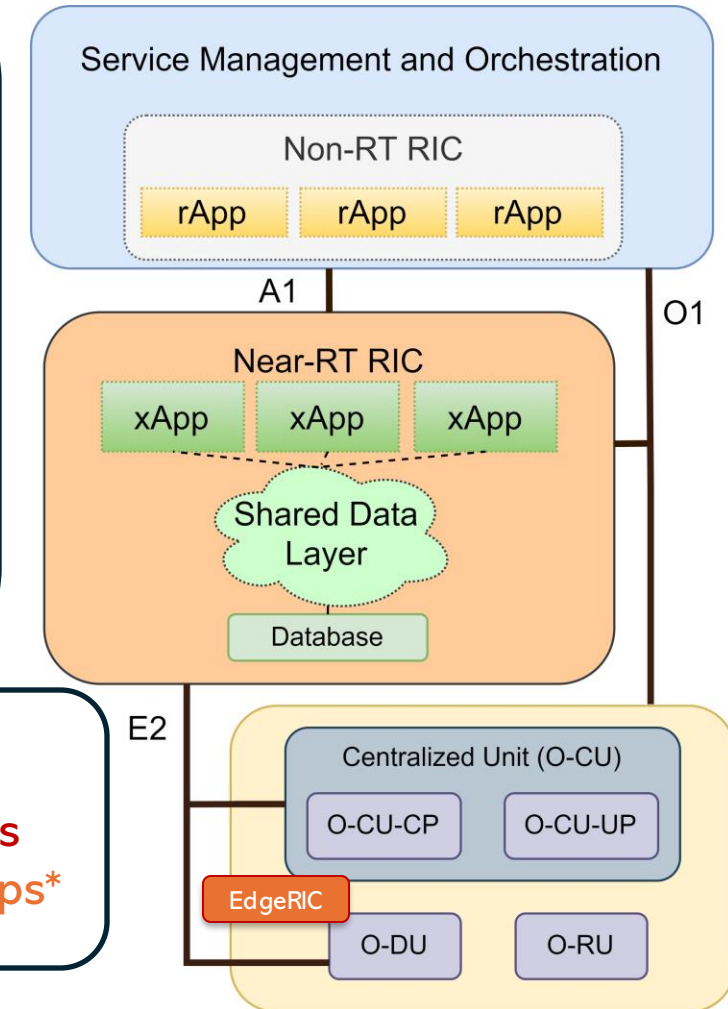
## RAN Intelligent Controllers (RICs)

- Abstract the networks
- Allow telecom operators to implement custom control logic



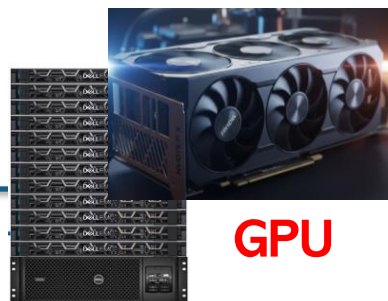
- Non-real-time RIC (> 1s) – **rApps**
- Near-RT RIC (10 ms – 1s) – **xApps**
- **Real-time RIC\*** (< 10 ms) – **d/μApps\***

\* Not standardized yet. Under review in nGRG RS-02 research item



# Big Push! AI-RAN – Nvidia/Nokia Effect!

AI-RAN (Likely to happen in 6G!)



GPU

Open  
Interfaces

Mini data  
centers

Versatility

New revenue?

- Run AI models
- Perform edge analytics

Physical AI

Flexibility

Softwarization

Virtualization

Disaggregated  
(O-RU+O-DU+O-CU)

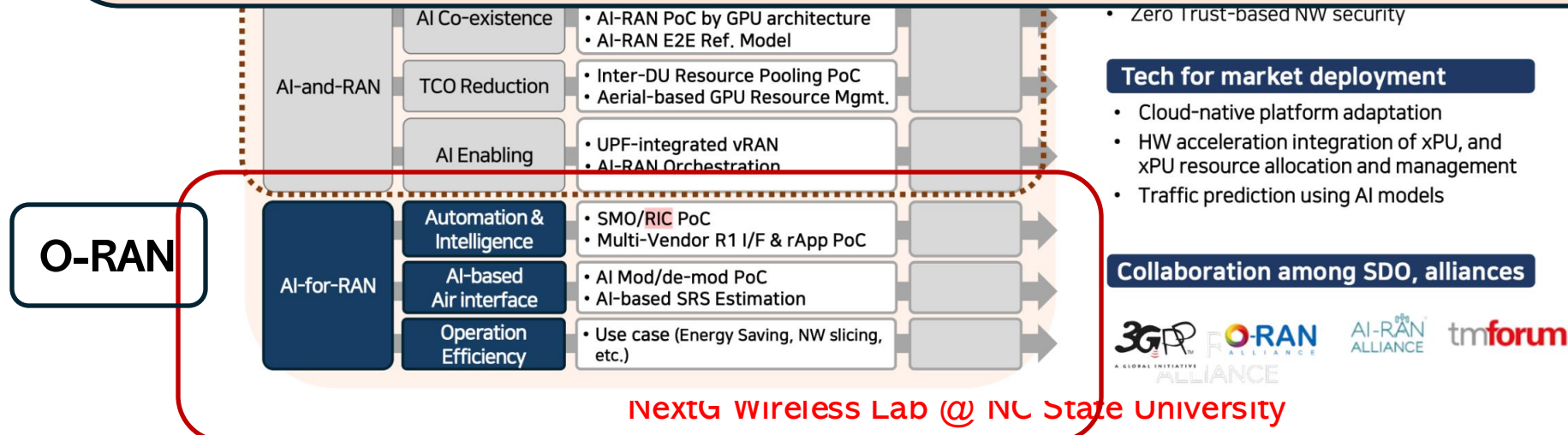
Open

# AI-RAN Expected to Evolve within O-RAN

## Network AI : AI-RAN

SKT defines the AI-RAN R&D roadmap across 3 areas, and both standardization and technology development for market deployment should progress concurrently

**So, AI-O-RAN convergence seems to be already happening!**



# AI-O-RAN R&D areas (from my lens!)

## AI and \*Apps

- AI-RAN, NVIDIA-NOKIA partnership!
- VINES – Open Network for verticals!

## Automated O-RAN Testing (particularly, IOT)

- NOFO1 served as a major catalyst!

## O-RU

- NOFO2

## O-RAN Integration Automation

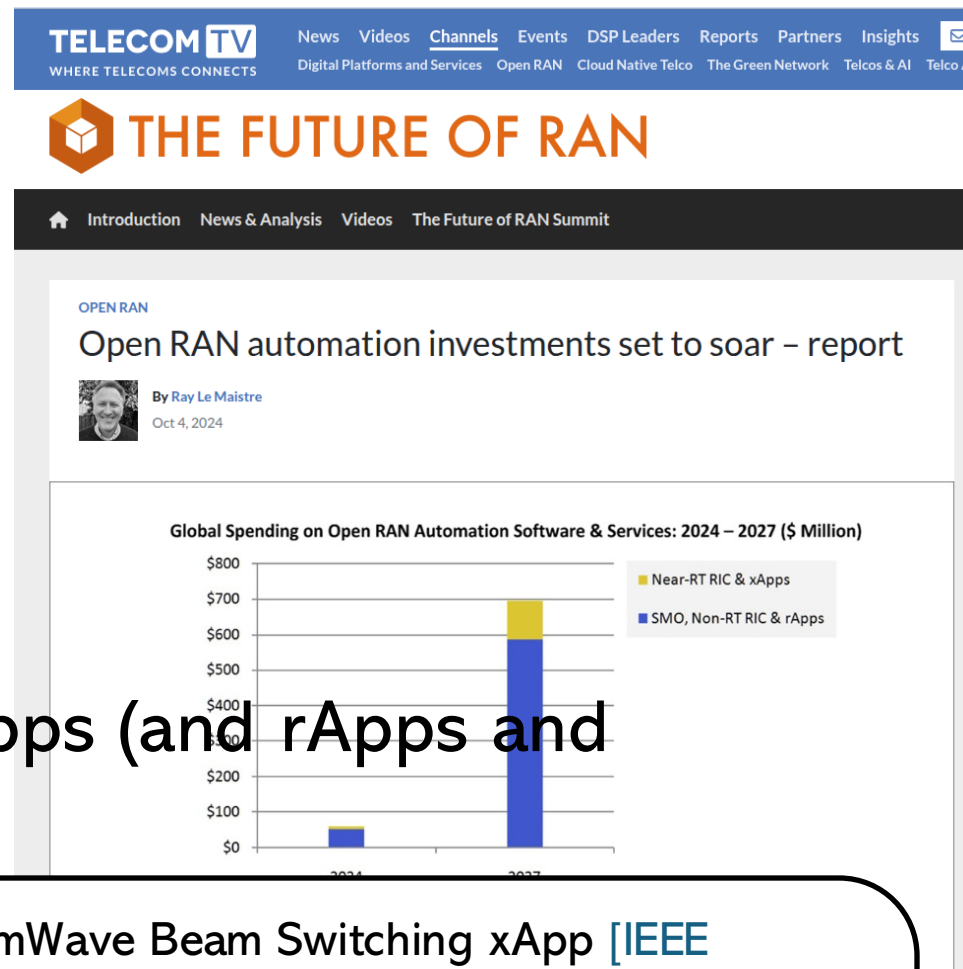
...

# (AI) \*Apps

- **rApps** – in Industry!
- **xApps** – (mostly) in Lab/Academia!
- We have developed several (ML) xApps (and rApps and  $\mu$ App/dApp )

- Interference Classification (IC) xApp [IEEE MILCOM 2023] \*Best demo award
- Radar Detection xApp [IEEE JSAC'23]
- SPARC – IC xApp + RB scheduling  $\mu$ App [ACM CoNext'24]
- IMPACT xApp [Demonstrated at NTIA RIC Forum in March 2024]

- mmWave Beam Switching xApp [IEEE DySPAN'24]
- UE Positioning xApp\*
- Handover x/rApp\*
- Energy Saving rApp\*
- ISAC dApp
- ...



# How to build best-fit AI/ML models?



AI/ML is **everywhere** in  
telecoms now!  
Network orchestration,  
management,  
PHY/MAC optimization!

We are **NOT** an AI/ML  
expert!  
Lots of Guesswork and  
Gruntwork!

ML Engineering is **time-  
consuming**  
Several weeks to  
months!



# NeuralSmith: Accuracy Performance on Representative Wireless Dataset

Dataset	ML Task	SOTA	NeuralSmith
<b>TRACTOR Data</b> Groen, Joshua, et al. "TRACTOR: Traffic analysis and classification tool for open RAN." ICC 2024-IEEE International Conference on Communications. IEEE, 2024.	C	95%	95.84%
<b>UCI Localization Dataset</b> Rohra, Jayant G., et al. "User localization in an indoor environment using fuzzy hybrid of particle swarm optimization & gravitational search algorithm with neural networks." Proceedings of Sixth International Conference on Soft Computing for Problem Solving: SocProS 2016, Volume 1. Singapore: Springer Singapore, 2017.	C	95.16%	97.00%
<b>Path Loss Dataset</b> Popoola, Segun I., et al. "Path loss dataset for modeling radio wave propagation in smart campus environment." Data in brief 17 (2018): 1062-1073.	R	1.275dB	2.79dB
<b>Interference Classification Dataset</b> Chiejina, Azuka, et al. "System-level analysis of adversarial attacks and defenses on intelligence in O-RAN based cellular networks." Proceedings of the 17th ACM Conference on Security and Privacy in Wireless and Mobile Networks. 2024.	C	98%	99.33%

# NeuralSmith: Accuracy Performance on Representative Wireless Dataset

Dataset	Temporal Characteristics		
	Features	Compute	Time
<b>TRACTOR Dataset</b> Groen, Joshua, et al. "TRACTOR: Traffic analysis and classification tool for open RAN." ICC 2024-IEEE International Conference on Communications. IEEE, 2024.	<b>44,984 Samples, 288 Features</b>	<b>16GB RAM, CPU</b>	<b>2.57 minutes</b>
<b>User Localization Dataset</b> Rohra, Jayant G., et al. "User localization in an indoor environment using fuzzy hybrid of particle swarm optimization & gravitational search algorithm with neural networks." Proceedings of Sixth International Conference on Soft Computing for Problem Solving: SocProS 2016, Volume 1. Singapore: Springer Singapore, 2017.	<b>2000 Samples, 7 Features</b>	<b>16GB RAM, CPU</b>	<b>13.801 seconds</b>
<b>Path Loss Dataset</b> Popoola, Segun I., et al. "Path loss dataset for modeling radio wave propagation in smart campus environment." Data in brief 17 (2018): 1062-1073.	<b>3616 Samples, 8 Features</b>	<b>16GB RAM, CPU</b>	<b>16.73 seconds</b>
<b>Interference Classification Dataset</b> Chiejina, Azuka, et al. "System-level analysis of adversarial attacks and defenses on intelligence in O-RAN based cellular networks." Proceedings of the 17th ACM Conference on Security and Privacy in Wireless and Mobile Networks. 2024.	<b>4,500 Samples, 32x32x3 Images</b>	<b>24GB VRAM, GPU: RTX-4090</b>  <b>16GB RAM, CPU</b>	<b>10.51 minutes,</b>  <b>1.671 hours</b>

# NeuralSmith: Key TakeAways!

- **Simplified process:** Upload your dataset and get a production-ready ML model-fast, simple and hassle-free. **No Guesswork, No Gruntwork!**
- **Fully interactive:** users can interact via user-friendly conversation!
- **Local-first, Privacy-First:** User's data never leaves their machine!
- **Top 2% ML model on MLEBench-Lite**
  - Performs **consistently better** than SOTA models, and in a **fraction of time!**

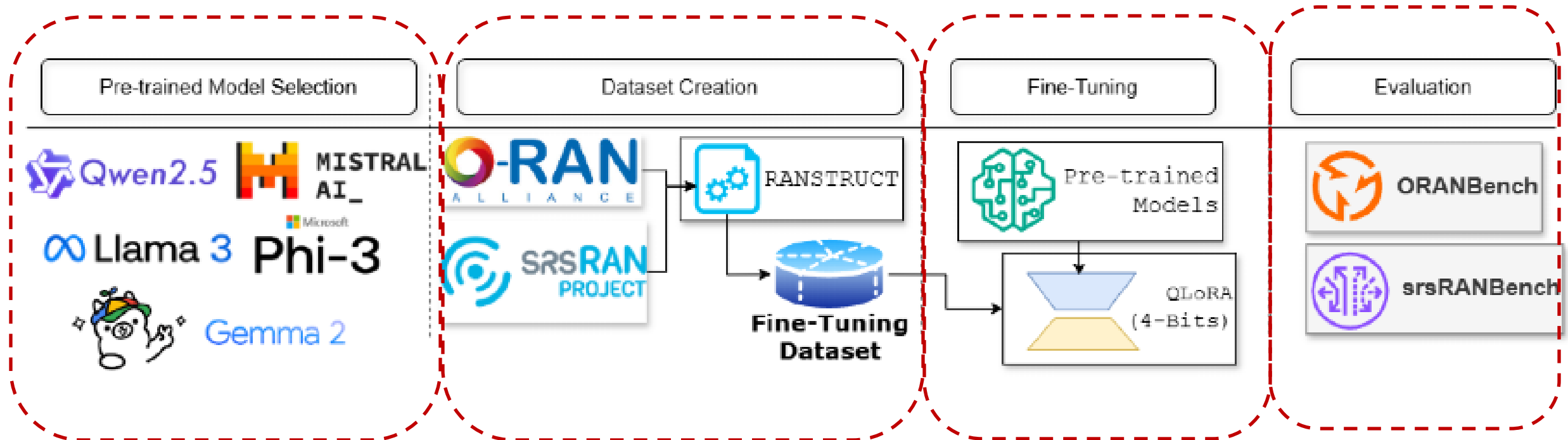
# ORANSight – Foundational LLMs for O-RAN

- Publicly available on HuggingFace: <https://huggingface.co/NextGLab>
- P. Gajjar and V. K. Shah, ORANSight-2.0: Foundational LLMs for O-RAN. In IEEE TMLCN 2025.
- P. Gajjar, V. K. Shah. ORAN-Bench-13K: An open-source benchmark for assessing LLMs in open radio access networks. in IEEE CCNC 2025.

# ORANSight-2.0

- ORANSight-2.0 (**O-RAN Insights**) aims at developing specialized foundational LLMs tailored for O-RAN.
- Built on **18** models spanning – **Mistral, Qwen, Llama, Phi** and **Gemma** – ranging from **1B** to **70B** parameters
- **RANSTRUCT** – a novel RAG-based instruction-tuning framework to generate high-quality instruction-tuning datasets.
- Key Techniques: **Finetuning + RAG**
  - **O-RAN**: 116 Specification Documents\*
  - **3GPP**: 14560 Specification Documents\*
  - The **srsRAN Project** (and **OAI**\*)
- ORANSight-2.0 models are available on HuggingFace: <https://huggingface.co/NextGLab>
- Dataset available on GitHub: <https://github.com/prnshv/ORAN-Bench-13K>
- ORANSight-R will be available soon!

# ORANSight-2.0 Pipeline



1. Pre-trained LLM Model Selection

2. RANSTRUCT Framework for Dataset Creation

3. Fine-tuning with QLoRA

4. Evaluation with ORANBench and srsRANBench

# ORANSight-2.0 against LLM baselines

Model Family	Parameters (B)	ORANBench				srsRANBench	Cumulative Score
		Easy	Medium	Difficult	Average		
ORANSight-Gemma-2	2	0.698	0.618	0.584	0.633	<b>0.848</b>	0.74
	<b>9</b>	0.822	0.718	0.680	<b>0.740</b>	<b>0.834</b>	<b>0.787</b>
	<b>27</b>	0.818	0.722	0.702	<b>0.747</b>	<b>0.911</b>	<b>0.829</b>
ORANSight-Mistral	7 (v3)	0.730	0.618	0.612	0.650	<b>0.863</b>	0.756
	<b>12 (Nemo)</b>	0.784	0.710	0.650	0.715	<b>0.828</b>	<b>0.772</b>
	<b>22 (Small)</b>	0.802	0.700	0.658	0.720	<b>0.860</b>	<b>0.79</b>
	8x7 (AQLM)	0.780	0.654	0.626	0.687	0.753	0.72
ORANSight-Llama	3.2: 1	0.438	0.350	0.364	0.384	0.565	0.474
	3.2: 3	0.708	0.582	0.540	0.610	0.768	0.689
	3.1: 8	0.728	0.730	0.618	0.692	0.728	0.71
	3.1: 70 (AQLM)	0.820	0.680	0.652	0.717	0.761	0.739
ORANSight-Phi	3.5: Mini	0.716	0.670	0.658	0.681	0.754	0.718
	3: Medium	0.754	0.710	0.682	0.715	<b>0.783</b>	0.749
ORANSight-Qwen-2.5	1.5	0.660	0.586	0.606	0.617	<b>0.804</b>	0.71
	3	0.704	0.642	0.628	0.658	<b>0.857</b>	0.758
	<b>7</b>	0.788	0.720	0.696	0.735	<b>0.837</b>	<b>0.786</b>
	14	0.804	0.730	0.684	0.739	0.761	0.75
	<b>32</b>	0.856	0.784	0.738	<b>0.793</b>	<b>0.796</b>	<b>0.794</b>
(Baseline) ChatGPT	4o-mini	0.766	0.727	0.677	0.723	0.755	0.739
	4o	0.792	0.760	0.693	0.752	0.769	0.76
(Baseline) Gemini	1.5:8B	0.723	0.665	0.631	0.673	0.767	0.72
	1.5	0.743	0.707	0.669	0.706	0.775	0.741

# ORANSight-2.0 against Telecom-specific LLM models

Model	ORANBench				srsRANBench	Cumulative Score
	Easy	Medium	Difficult	Average		
<b>ORANSight-2.0 (Best)</b>	0.904	0.884	0.878	0.889	0.911	0.900
ORANSight-Qwen-2.5: 32B (RAG)	0.904	0.884	0.878	0.889	0.818	0.854
Tele-LLM	0.630	0.556	0.534	0.573	0.770	0.672
TSpec-LLM	0.748	0.668	0.526	0.647	0.714	0.681
NVIDIA-O-RAN	0.826	0.816	0.836	0.826	0.688	0.757

ORANSight-2.0 (best) outperforms state-of-the-art by 27.96%

ORANSight-Qwen-2.5: 32B for ORANBench

ORANSight-Gemma-2: 27B for srsRANBench

# Sample Generated Questions

**Instruction:** What is the significance of the `byte_buffer` and `byte_buffer_chain` classes?

**Answer:** The `byte_buffer` class represents a sequence of bytes in memory, while the `byte_buffer_chain` class manages multiple `byte_buffers` together, allowing for efficient storage and manipulation of large amounts of data. It provides methods to append new buffers to its end and to perform operations like copying itself or creating new instances based on existing ones. These classes are fundamental for handling and manipulating binary data structures within the context of software systems designed around O-RAN technology. They enable efficient management of buffer pools and help ensure proper allocation and deallocation of resources, which is crucial for applications running on the radio access network infrastructure. The design of these classes reflects the need for scalable and flexible memory management strategies in such environments.

**Instruction:** Generate code for a function that simulates a failure in the underlying F1AP interface and tests the robustness of the DU's handling of such failures.

**Answer:** To simulate a failure in the underlying F1AP interface and test the robustness of the DU's handling of such failures, you can follow these steps:

1. **Simulate Failure:** Create a function that generates a `flap_message` representing a failure scenario, such as timeout or other errors.
2. **Handle Message:** Implement logic within the DU to handle this simulated failure message by checking its type and reacting accordingly.
3. **Test Robustness:** Ensure the DU correctly identifies and responds to various types of failure messages, including timeouts, invalid requests, or other potential issues.
4. **Assertions:** Use assertions to verify that the DU correctly handles different types of failure responses, ensuring the system remains stable under varying conditions.
5. **Validation:** Validate that the DU's behavior aligns with expected outcomes based on the simulation results.

Here is a simplified example of how this might look in C++ using the provided framework and technology:

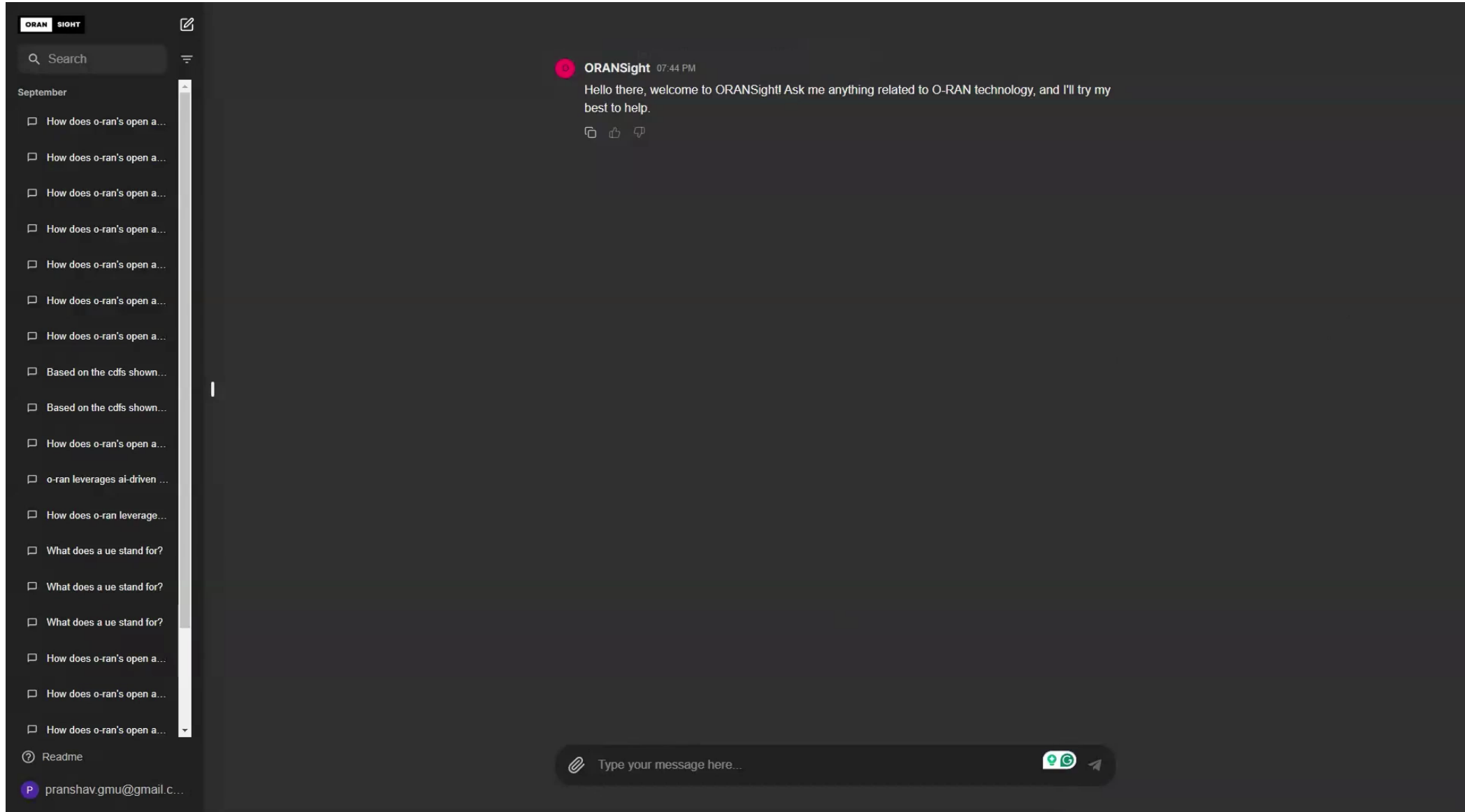
```
void simulate_and_handle_failure(  
    ↳ du_processor_obj* du_processor_obj)  
    ↳ {  
    // Simulate failure: Timeout  
    flap_message failureMsg =  
        ↳ generate_ue_context_setup_failure  
        ↳ (  
            int_to_gnb_cu_ue_flap_id(0),  
            ↳ int_to_gnb_du_ue_flap_id(  

```

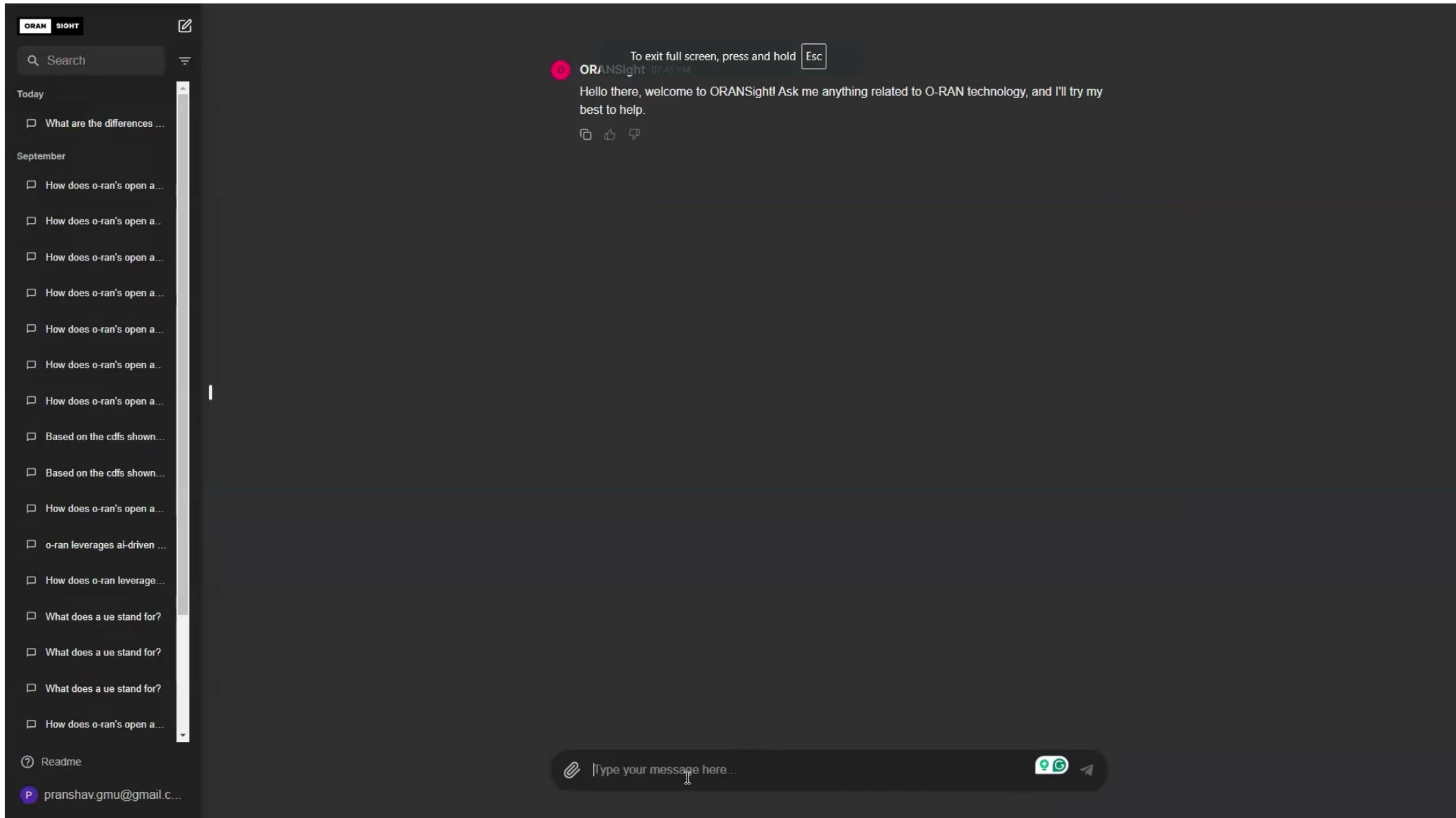
```
        ↳ test_rgen::uniform_int<  
        ↳ uint32_t>());  
  
    // Handle the failure message  
    du_processor_obj->handle_message(  
        ↳ failureMsg);  
  
    // Verify that the DU has handled the  
    ↳ failure properly  
    // This could involve checking  
    ↳ specific fields in the DU's  
    ↳ state or asserting certain  
    ↳ behaviors  
  
    // Additional checks could include  
    ↳ verifying responses to other  
    ↳ types of messages  
}
```

This approach allows you to systematically test the DU's resilience against various failure scenarios without relying on external figures or sections. The focus is on implementing and validating the DU's internal mechanisms to ensure it functions reliably even when faced with unexpected or simulated failures. Please note that the actual implementation would require careful consideration of the specific requirements and constraints of the DU and the F1AP protocol. The above example provides a basic structure for testing failure scenarios.

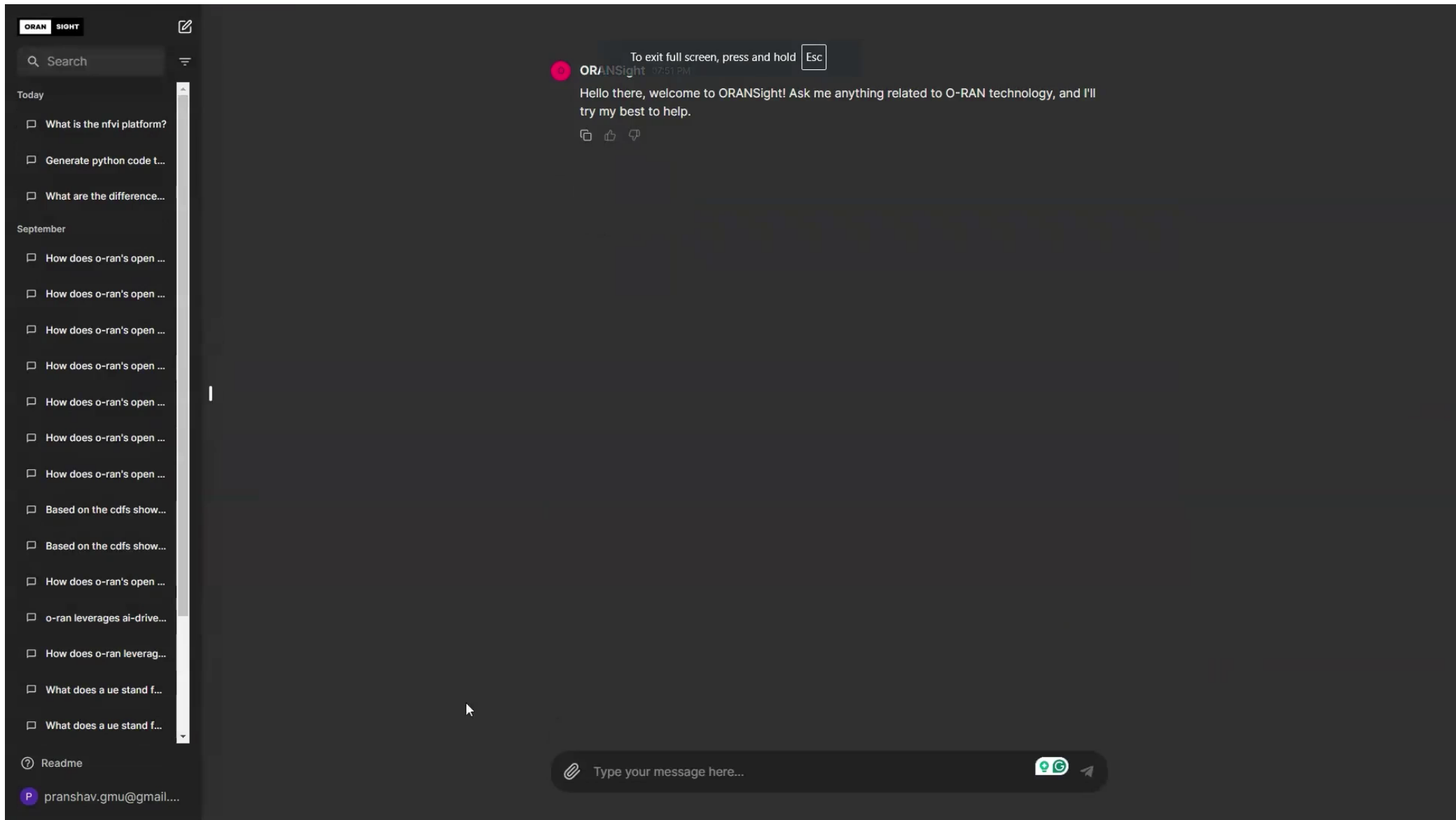
ORANSight currently has a maximum context window of 128k tokens, allowing us to work with a maximum of roughly 96k words.



ORANSight supports code generation with high fidelity in multiple languages, with an emphasis on C++, and Python



ORANSight combines the underlying Mistral-7B with a Phi-3.5 Vision, to create a dual LLM setup making the entire process substantially resource efficient.



# AI5GTest

AI-Driven Specification-Aware Automated Testing and Validation of 5G O-RAN Components

- A. Ganiyu, P. Gajjar, and V. K. Shah, AI5GTest: AI-Driven Specification-Aware Automated Testing and Validation of 5G O-RAN Components. In ACM WiSec 2025.
- A. Ganiyu, P. Gajjar, V. K. Shah, DEMO: AI5GTest: LLM based Automation for 5G O-RAN Testing. In ACM WiSec 2025.

# Background and Motivation

- O-RAN Testing, particularly IOT, is a MUST!
- O-RAN ALLIANCE WGs (TIFG and WG5) have documented various test cases for IOT and other E2E testing, these documents often
  - **cross-reference multiple O-RAN ALLIANCE, 3GPP** (and sometimes, ETSI and ITU-T) specifications
  - **are fragmented, incomplete and prone to human interpretations**
- Current approach (including, OTICs) involves **manually running tests**, capturing packet files from the 5G network, and checking if all signaling messages conform to the specified specification standards.
  - Take **several weeks to months** to conduct such manual O-RAN testing approach

Open Testing and Integration Centres (OTICs) – collaborative hubs for verification, integration, testing, badging and certification of O-RAN components

# Background and Motivation

State-of-the-art manual approach is time-consuming and prone to human error.

- Under NOFO1 award, we aim to develop AI5GTest - AI-assisted testing framework for O-RAN components.
- Current AI5GTest capabilities
  - Automated generation of expected procedural flows for test cases
  - Automated validation of signaling messages against established standards, reducing manual overhead and enhancing testing consistency and scalability
  - Automated identification of root causes of test failures by analyzing deviations in execution flows and providing insights for debugging.

# Overview of Manual Testing Process

## 2.3.1.15 F1 Setup for NR Stand-Alone

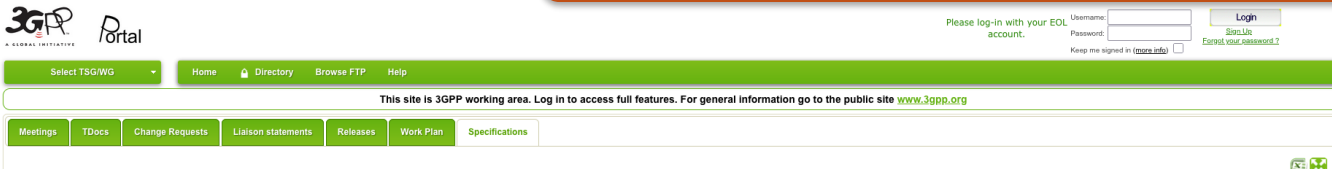
### 2.3.1.15.1 Test Purpose

The purpose of this test case is to verify that the F1 Setup procedure between gNB-CU and gNB-DU from different vendors can be successfully completed, conforming to the NR C-Plane profile specification [2] Section 4.2.3.1.

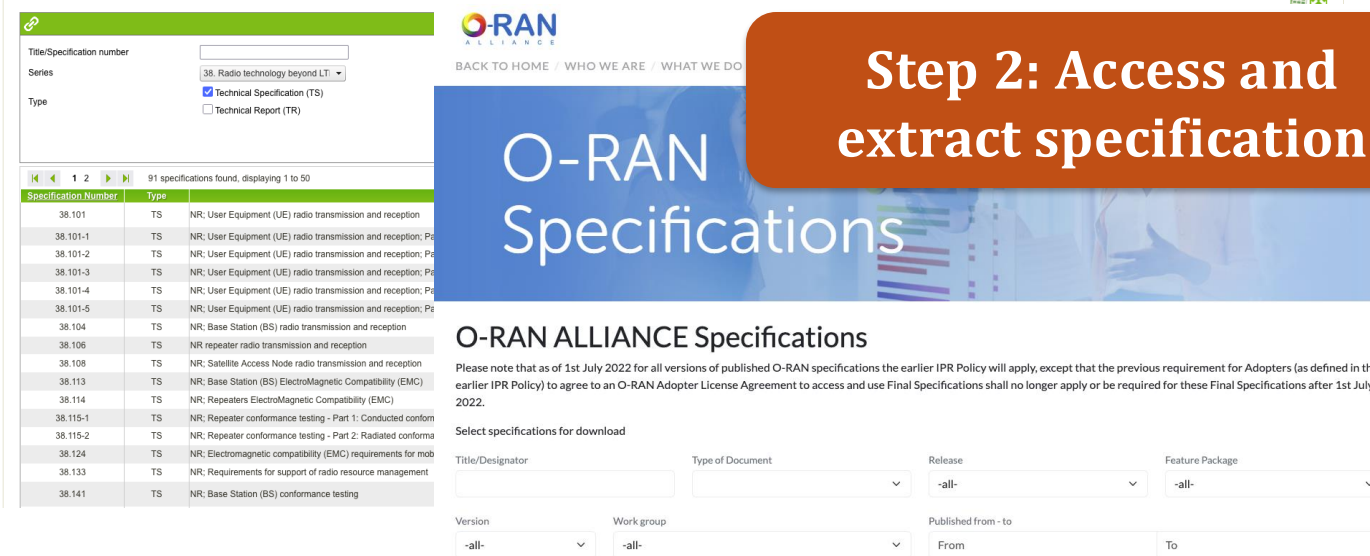
### 2.3.1.15.2 Minimum Requirements

DUTs: Single gNB-CU and single gNB-DU;

**Step 1: Load/Read a test case**

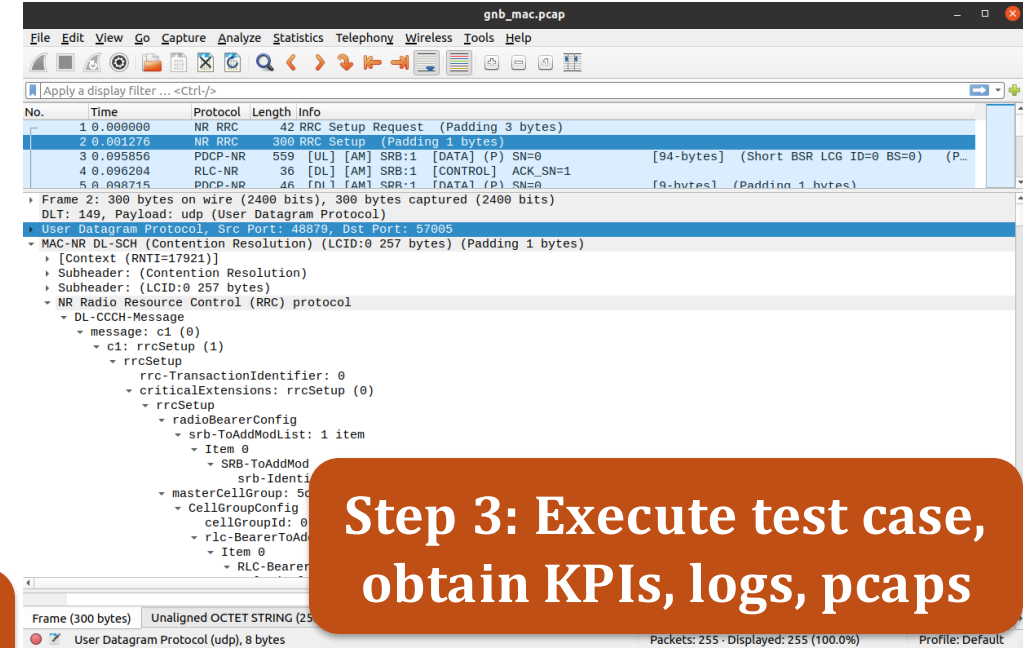


**Step 2: Access and extract specification**



Other standard specifications – ETSI, ITU-T

NextG Wireless Lab @ NC State University

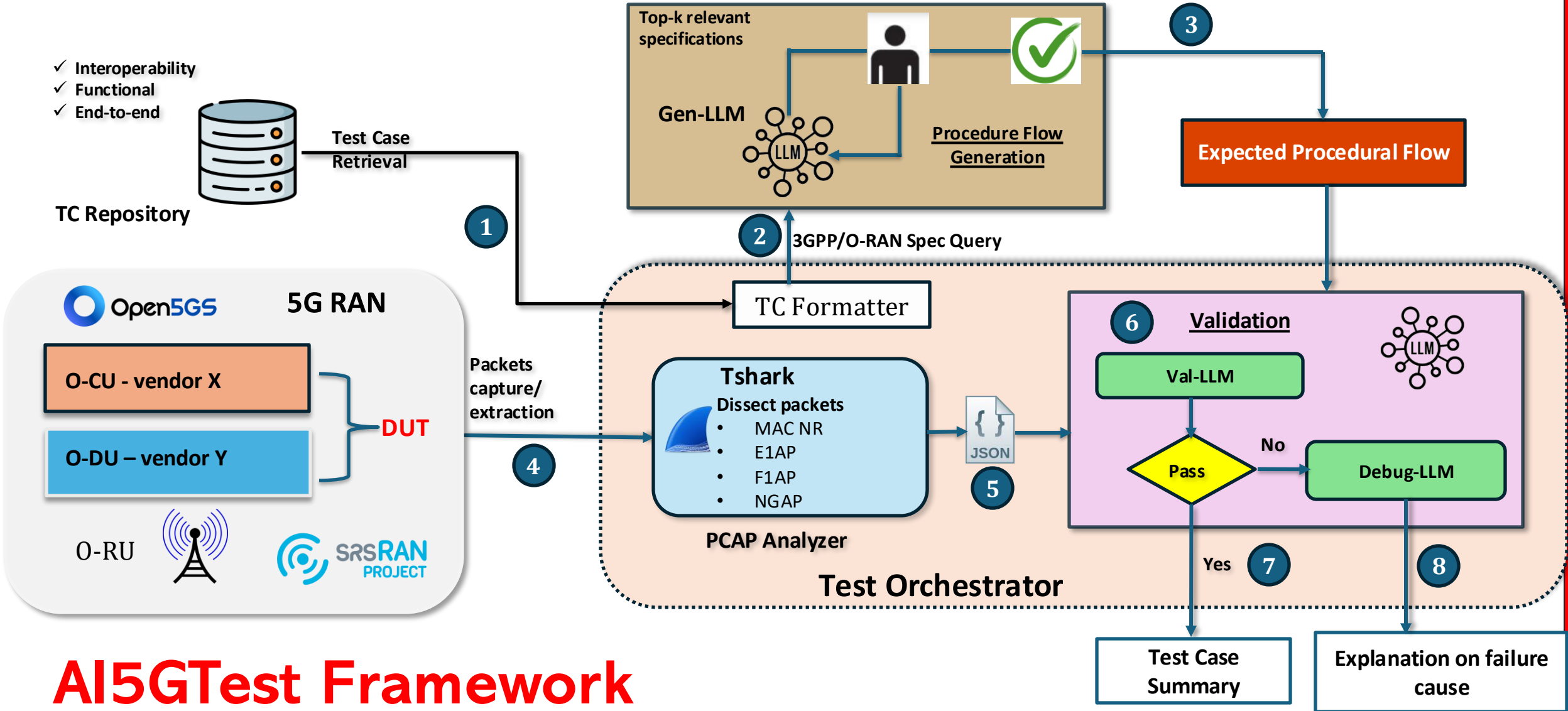


**Step 3: Execute test case, obtain KPIs, logs, pcaps**



**Step 4: Compare and verify**

- ✓ Interoperability
- ✓ Functional
- ✓ End-to-end



# AI5GTest Framework

# Example Test Case: UE Initial Attach/5G SA Registration of a single UE

- The purpose of the test is to verify the full registration procedure of a single UE.
- The test focuses on the procedure of 'Initial registration' as defined in 3GPP TS 38.401 Clause 4.2.2.2.

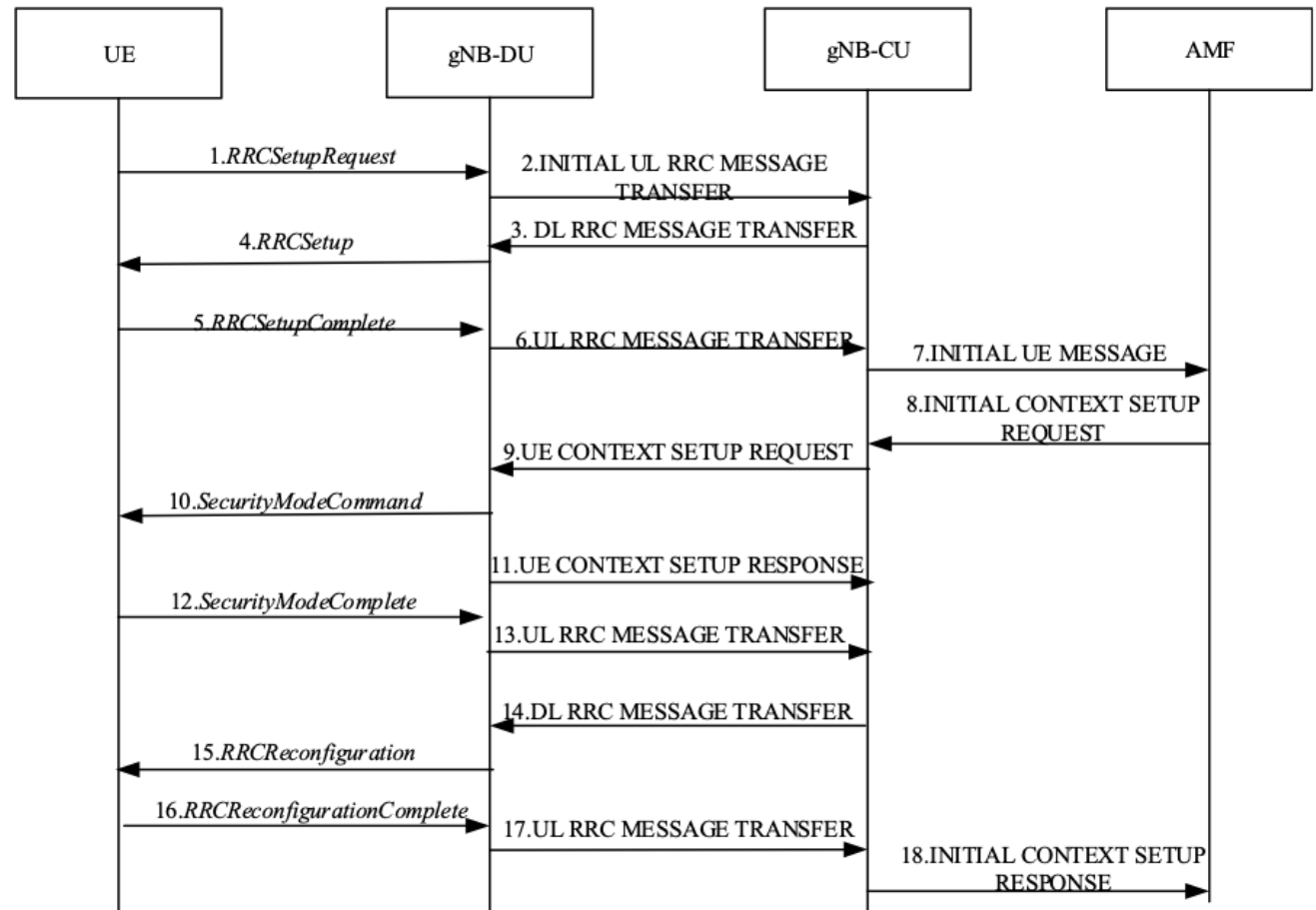


Fig. 3GPP TS 38.401: Procedural flow for successful UE Attach

# Eg 1: UE Initial Access procedure

- Retrieved Document Rank: **1**
- Specification Location:  
*docs/38401-gb0.docx*

Figure 8.1-1: UE Initial Access procedure

1\ The UE sends an \*RRCSetupRequest\* message to the gNB-DU.

2\ The gNB-DU includes the RRC message and, if the UE is admitted, the corresponding low layer configuration for the UE in the INITIAL UL RRC MESSAGE TRANSFER message and transfers to the gNB-CU. The INITIAL UL RRC MESSAGE TRANSFER message includes the C-RNTI allocated by the gNB-DU.

3\ The gNB-CU allocates a gNB-CU UE F1AP ID for the UE and generates a \*RRCSetup\* message towards UE. The RRC message is encapsulated in the DL RRC MESSAGE TRANSFER message.

4\ The gNB-DU sends the \*RRCSetup\* message to the UE.

5\ The UE sends the RRC CONNECTION SETUP COMPLETE message to the gNB-DU.

6\ The gNB-DU encapsulates the RRC message in the UL RRC MESSAGE TRANSFER message and sends it to the gNB-CU.

7\ The gNB-CU sends the INITIAL UE MESSAGE message to the AMF.

8\ The AMF sends the INITIAL CONTEXT SETUP REQUEST message to the gNB-CU.

...

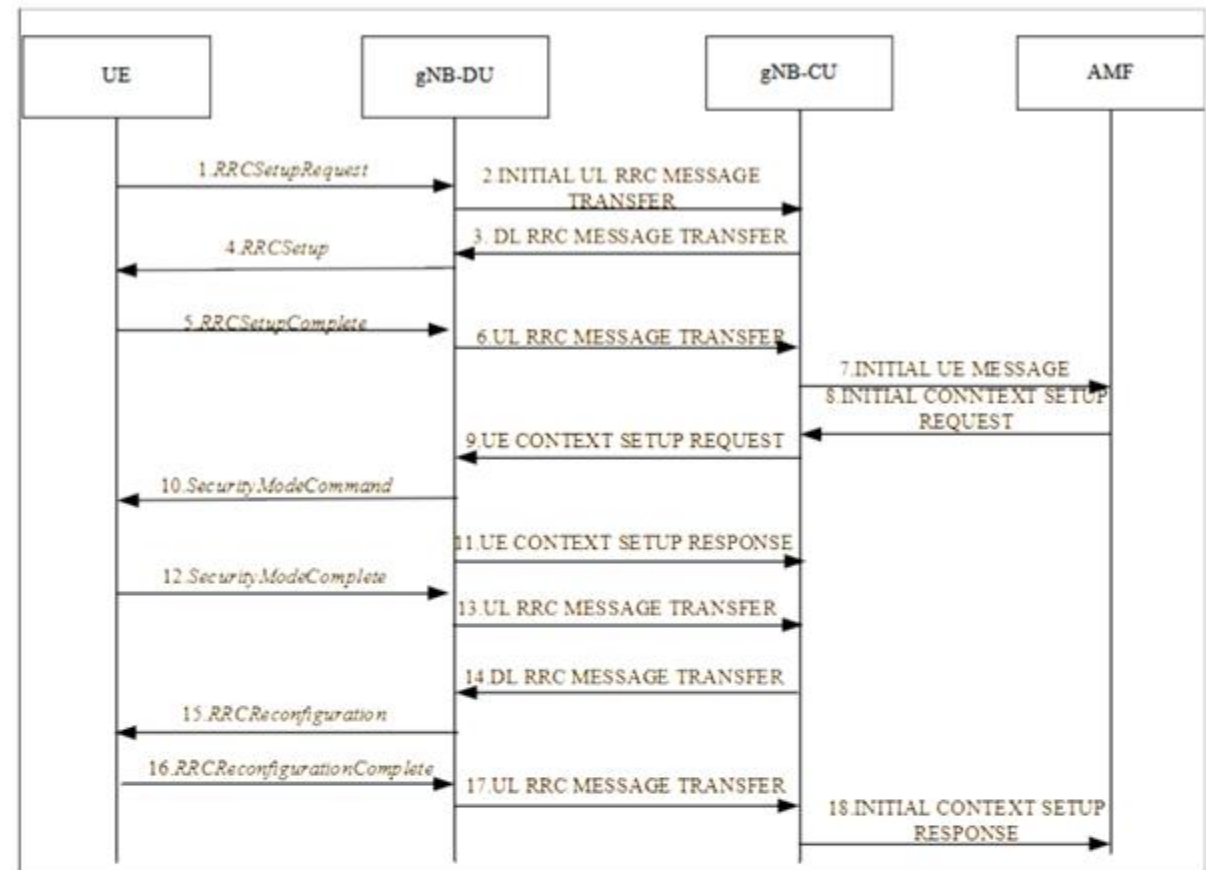
4\ The gNB-DU sends the \*RRCSetup\* message to the UE.

5\ The UE sends the RRC CONNECTION SETUP COMPLETE message to the gNB-DU.

**Document:**

*docs/38401-gb0.docx*

**Page:** 20, 21



**Figure 8.1-1: UE Initial Access procedure**

1. The UE sends an *RRCSetupRequest* message to the gNB-DU.
2. The gNB-DU includes the RRC message and, if the UE is admitted, the corresponding low layer configuration for the UE in the INITIAL UL RRC MESSAGE TRANSFER message and transfers to the gNB-CU. The INITIAL UL RRC MESSAGE TRANSFER message includes the C-RNTI allocated by the gNB-DU.
3. The gNB-CU allocates a gNB-CU UE F1AP ID for the UE and generates a *RRCSetup* message towards UE. The RRC message is encapsulated in the DL RRC MESSAGE TRANSFER message.
4. The gNB-DU sends the *RRCSetup* message to the UE.
5. The UE sends the RRC CONNECTION SETUP COMPLETE message to the gNB-DU.
6. The gNB-DU encapsulates the RRC message in the UL RRC MESSAGE TRANSFER message and sends it to the gNB-CU.
7. The gNB-CU sends the INITIAL UE MESSAGE message to the AMF.

# Gen-LLM Prompt: Give the procedural flow for initial UE attach in gNB-CU-DU architecture?

## LLM Generation:

The UE Initial Access procedure between gNB-DU, gNB-CU, and AMF is as follows:

1. The UE sends an RRCSetupRequest message to the gNB-DU.
2. The gNB-DU includes the RRC message and, if the UE is admitted, the corresponding low layer configuration for the UE in the INITIAL UL RRC MESSAGE TRANSFER message and transfers it to the gNB-CU. The INITIAL UL RRC MESSAGE TRANSFER message includes the C-RNTI allocated by the gNB-DU. If the gNB-DU identifies the UE as a Reduced Capability UE during the random access procedure, a NR RedCap UE Indication is provided in the INITIAL UL RRC MESSAGE TRANSFER message.
3. The gNB-CU allocates a gNB-CU UE F1AP ID for the UE and generates a RRCSetup message towards the UE. The RRC message is encapsulated in the DL RRC MESSAGE TRANSFER message.
4. The gNB-DU sends the RRCSetup message to the UE.
5. The UE sends the RRC CONNECTION SETUP COMPLETE message to the gNB-DU.
6. The gNB-DU encapsulates the RRC message in the UL RRC MESSAGE TRANSFER message and sends it to the gNB-CU.
7. The gNB-CU sends the INITIAL UE MESSAGE message to the AMF.
8. The AMF sends the INITIAL CONTEXT SETUP REQUEST message to the gNB-CU.
9. The gNB-CU sends the UE CONTEXT SETUP REQUEST message to establish the UE context in the gNB-DU. In this message, it may also encapsulate the SecurityModeCommand message.

10. The gNB-DU sends the SecurityModeCommand message to the UE.
11. The gNB-DU sends the UE CONTEXT SETUP RESPONSE message to the gNB-CU.
12. The UE responds with the SecurityModeComplete message.
13. The gNB-DU encapsulates the RRC message in the UL RRC MESSAGE TRANSFER message and sends it to the gNB-CU.
14. The gNB-CU generates the RRCReconfiguration message and encapsulates it in the DL RRC MESSAGE TRANSFER message.
15. The gNB-DU sends the RRCReconfiguration message to the UE.
16. The UE sends the RRCReconfigurationComplete message to the gNB-DU.
17. The gNB-DU encapsulates the RRC message in the UL RRC MESSAGE TRANSFER message and sends it to the gNB-CU.
18. The gNB-CU sends the INITIAL CONTEXT SETUP RESPONSE message to the AMF.

## Other relevant information:

The UE Initial Access procedure is the process by which a UE establishes an initial connection with the gNB-DU and gNB-CU and obtains the necessary resources for data transmission. The procedure involves the exchange of several messages, including RRCSetupRequest, RRCSetup, RRC CONNECTION SETUP COMPLETE, and RRCReconfiguration. The gNB-DU and gNB-CU also communicate with each other during this procedure, and the AMF is informed once the UE has been successfully established.

Confidence Score: 95

# Eg 2: Registration of a single eMBB network slice in the 5G SA network

- Retrieved Document Rank: **1**
- Specification Location:  
*docs/O-RAN.TIFG.E2E-Test.0-R003-v05.00.pdf*

```
requested PDU Session Release for Non-Roaming and Roaming with Local Breakout case' as defined in 3GPP 23.501
4.3.4.2.
Table 5-7 defines the verification steps along with validation of mandatory IEs for 5G SA registration procedure
eMBB network slice.
Table 5-7 5G SA Registration verification with mandatory IEs in single eMBB network slice
St. Procedure
Msg
Flow
Expected Output
1
RRCSetupComplete

UE →
gNB
Verify that UE sends only one S-NSSAI in IE s-nssai-List and SST = '0x01' in the
S-NSSAI to gNB.
2
Registration Request
UE →
AMF

Verify that UE sends only one S-NSSAI in IE Requested NSSAI to 5G AMF, and
IE 5GS registration type = 'initial registration', SST = '0x01' in the S-NSSAI.

© 2023 by the O-RAN ALLIANCE e.V. Your use is subject to the copyright statement on the cover page of this spec
35
O-RAN.TIFG.E2E-Test.0-R003-v05.00
```

Table 5-7 5G SA Registration verification with mandatory IEs in single eMBB network slice

St.	Procedure	Msg Flow	Expected Output
1	RRCSetupComplete	UE → gNB	Verify that UE sends only one S-NSSAI in IE <i>s-nssai-List</i> and <i>SST</i> = '0x01' in the <i>S-NSSAI</i> to gNB.
2	Registration Request	UE → AMF	Verify that UE sends only one S-NSSAI in IE <i>Requested NSSAI</i> to 5G AMF, and IE <i>5GS registration type</i> = 'initial registration', <i>SST</i> = '0x01' in the <i>S-NSSAI</i> .

**Document:** docs/O-RAN.TIFG.E2E-Test.0-R003-v05.00.pdf

**Page:** 34,35

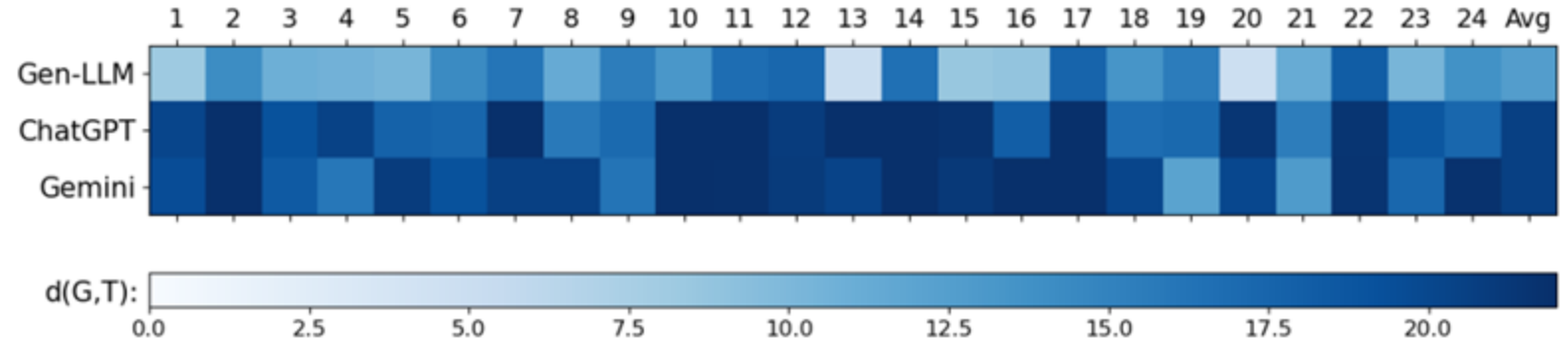


St.	Procedure	Msg Flow	Expected Output
3	Initial Context Setup Request	AMF → gNB	Verify that 5G AMF sends only one S-NSSAI in IE <i>Allowed NSSAI</i> and <i>S-NSSAI</i> (under parent IE <i>PDU Session Resource Setup Request Item</i> ) to gNB, and <i>SST</i> = '0x01' in the <i>S-NSSAI</i> .
4	Registration Accept	AMF → UE	Verify that 5G AMF sends only one S-NSSAI in IE <i>Allowed NSSAI</i> and <i>Configured NSSAI</i> to UE, <i>SST</i> = '0x01' in the <i>S-NSSAI</i> and IE <i>Rejected NSSAI</i> shall not exist.

# Executed Test Cases

Test Category	Test Case Title	DUT	Referenced Documents
Interoperability, E2E	Initial UE access – UE Context Creation, Service Request.	O-CU, O-DU	O-RAN NR C-Plane Profile Spec - Clause 6.1.1
Interoperability, E2E	Initial access – UE Context Creation for Initial Registration.	O-CU, O-DU	O-RAN NR C-Plane Profile Spec - Clause 6.1.2
Interoperability, E2E	Registration Update without Follow-on Request.	O-CU, O-DU	O-RAN NR C-Profile Spec v13 - Clause 6.1.3
Interoperability, E2E	gNB-CU Initiated UE Context Modification.	O-CU, O-DU	O-RAN NR C-Profile Spec v13 – Clause 6.3.1.1
Interoperability, E2E	gNB-DU Initiated UE Context Release.	O-CU, O-DU	O-RAN NR C-Profile Spec v13 – Clause 6.2.2
Interoperability	FI Setup for NR SA.	O-CU, O-DU	O-RAN NR C-Profile Spec v13 – Clause 4.2.3.1
Interoperability, E2E	UE Initial Access over F1.	O-CU, O-DU	3GPP TS-Spec 38.401 v18 – Clause 8.1
Interoperability, E2E	Bearer Context Setup over F1-U.	O-CU-CP, O-CU-UP, O-DU	3GPP 38.401 v18 – Clause 8.9.2
Interoperability, E2E	RRC Connected to RRC Inactive.	O-CU, O-DU	O-RAN NR C-Profile Spec v13 - Clause 6.9.1
Interoperability, E2E	PDU Session Establishment.	O-CU, O-DU	O-RAN NR C-Profile Spec v13 - Clause 6.3.3.1
Interoperability, E2E	UE Initial Access over E1 and F1.	O-CU-CP, O-CU-UP, O-DU	3GPP TS-Spec 38.401 v18 – Clause 8.9.1
Interoperability, E2E	gNB-CU-UP Initiated Bearer Context Release over F1-U.	O-CU-CP, O-CU-UP, O-DU	3GPP TS-Spec 38.401 v18 – Clause 8.9.3.1
E2E - Functional	LTE/5G NSA Attach of a Single UE	EPC, gNB	3GPP TS 23.401 - Clause 5.3.2.1
E2E - Functional	LTE/5G NSA Detach of a Single UE	EPC, gNB	3GPP TS 23.401 - Clause 5.3.8.2.1
E2E - Functional	LTE/5G NSA Attach of Multiple UEs	EPC, gNB	3GPP TS 23.401 - Clause 5.3.2.1
E2E - Functional	LTE/5G NSA Detach of Multiple UEs	EPC, gNB	3GPP TS 23.401 - Clause 5.3.8.2.1
E2E - Functional	General Registration of UE in 5G SA	5GC, gNB	3GPP TS 23.502 - Clause 4.2.2.2.2
Interoperability, E2E	Intra O-DU Mobility for 5G NSA and SA	O-CU, O-DU	3GPP TS 38.401 - Clause 8.2.1
Interoperability, E2E	Inter O-DU Mobility for 5G NSA and SA	O-CU, O-DU	3GPP TS 38.401 - Clause 8.2.1
Interoperability, E2E	Inter O-CU Mobility for 5G NSA and SA	O-CU-CP, O-CU-UP, O-DU	3GPP TS 38.401 - Clause 8.9.4 (SA); 3GPP TS 37.340 - Clause 10.5.1 (NSA)
E2E - Functional	Registration to a Single emBB Network Slice in 5G SA	5GC, gNB	3GPP TS 23.502 - Clause 4.2.2.2.2; TIFG.E2E-Test Spec-Table 5.7
E2E - Functional	De-registration from a Single emBB Network Slice in 5G SA	5GC, gNB	3GPP TS 23.502 - Clause 4.2.2.3.2; TIFG.E2E-Test Spec-Table 5.8
E2E - Functional	Registration to Multiple Network Slices in 5G SA	5GC, gNB	3GPP TS 23.502 - Clause 4.2.2.2.2; TIFG.E2E-Test Spec-Table 5.10
E2E - Functional	De-registration from Multiple Network Slices	5GC, gNB	3GPP TS 23.502 - Clause 4.2.2.3.2; TIFG.E2E-Test Spec-Table 5.11

# Evaluation: Gen-LLM



- Gemma Score (based on Bert Score) for evaluation:
  - Uses a Gemma-2B instance (context length: 8k tokens) and L2 Norm.
- Gemma Score Range:
  - 8.33% of samples fall in the 0-5 category,
  - 12.5% in the 5-10 category,
  - 45.83% in the 10-15 category,
  - and the remaining 8 samples are in the 15-20 category.

**Table 1: Comparison of LLM Performance on Test Case Procedural Flow Generation.**

Model	Avg. Gemma Score
Gen-LLM (Ours)	12.581
GPT-4o	20.559
Gemini	20.579

# Gen-LLM Demo

The image shows a web browser window displaying a page titled "AI5GTest: LLM based Automation for Testing". The page includes a section for "Acknowledgement" and a list of "Ground Truths" with test cases like "TC-11: UE-attach over e1-f1". A "Screen Recorder - Scre.io" window is overlaid on the page, showing options for "What do you want to capture?" (Only Screen, Screen + Camera, Only Camera) and "Record audio?" (None, Microphone, System + Mic., System). A "Start Recording" button is visible at the bottom of the recorder window. The browser's address bar shows "localhost:8501".

## AI5GTest: LLM based Automation for Testing

### Acknowledgement

We acknowledge the funding support from the Public Wireless Supply Chain Innovation Fund (PWS) shepherd and reviewers for their constructive feedback, which greatly contributed to improving the

#### Ground Truths

- TC-11: UE-attach over e1-f1 [Flowchart View](#)
- TC-07: UE Initial Access
- TC-05: gNB-DU Initiated UE Context Release
- TC-03: E2E Registration Update

#### Interactive Flow Diagram

UE	gNB-DU	gNB-CU-CP	gNB-CU-UP	AMF
----	--------	-----------	-----------	-----

Screen Recorder - Scre.io

What do you want to capture?

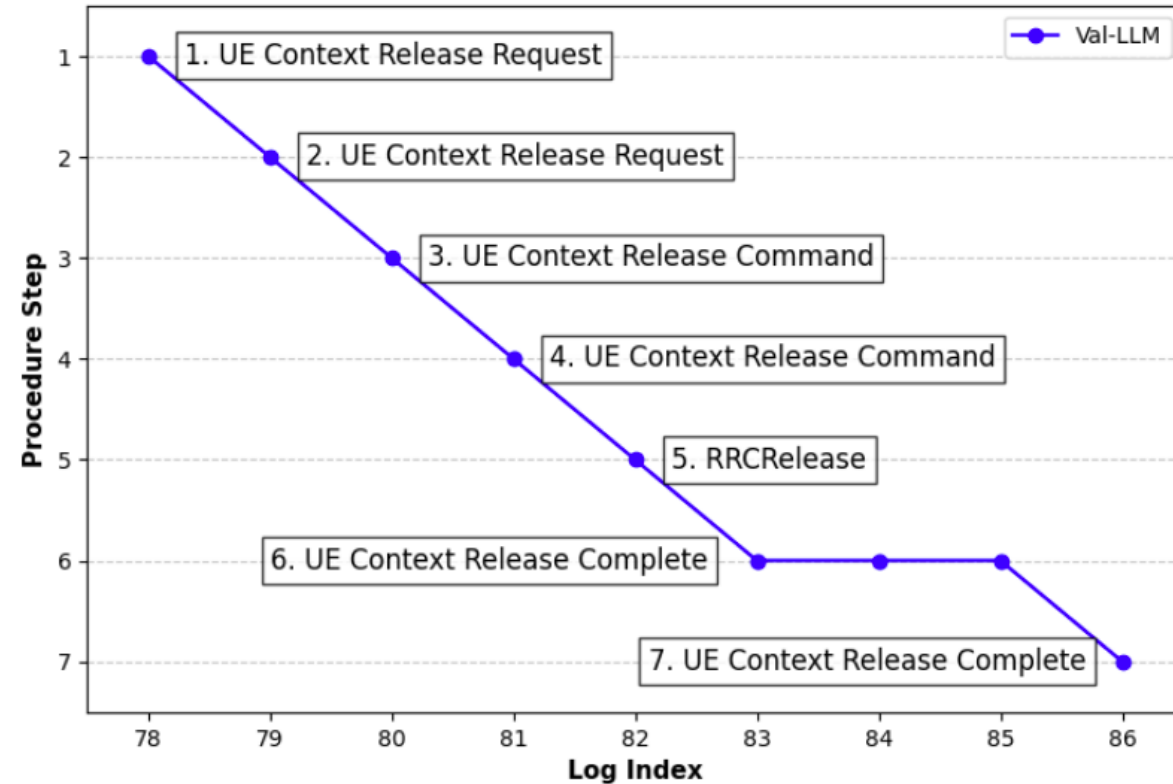
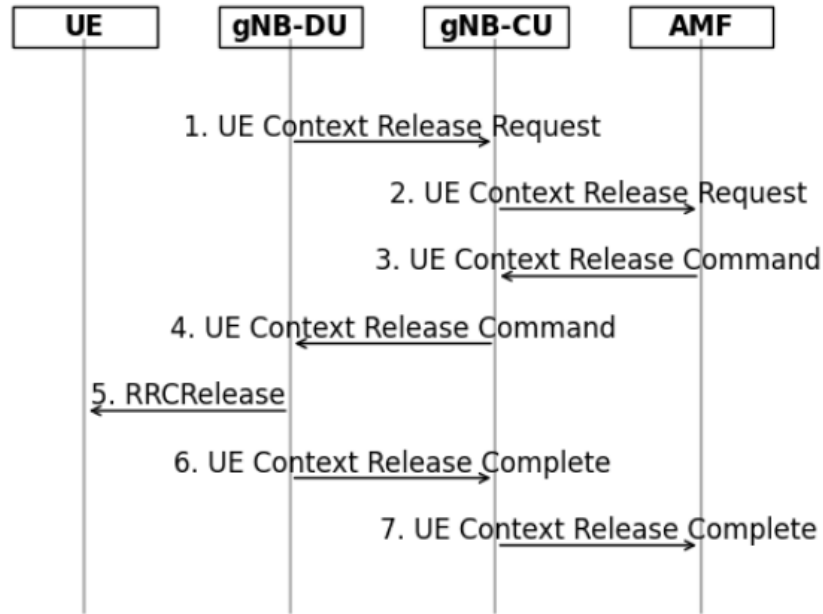
- Only Screen
- Screen + Camera **Coming Soon**
- Only Camera

Record audio?

- None
- Microphone
- System + Mic.
- System

Start Recording

# Result of Val-LLM – TP



**Label:**

Yes

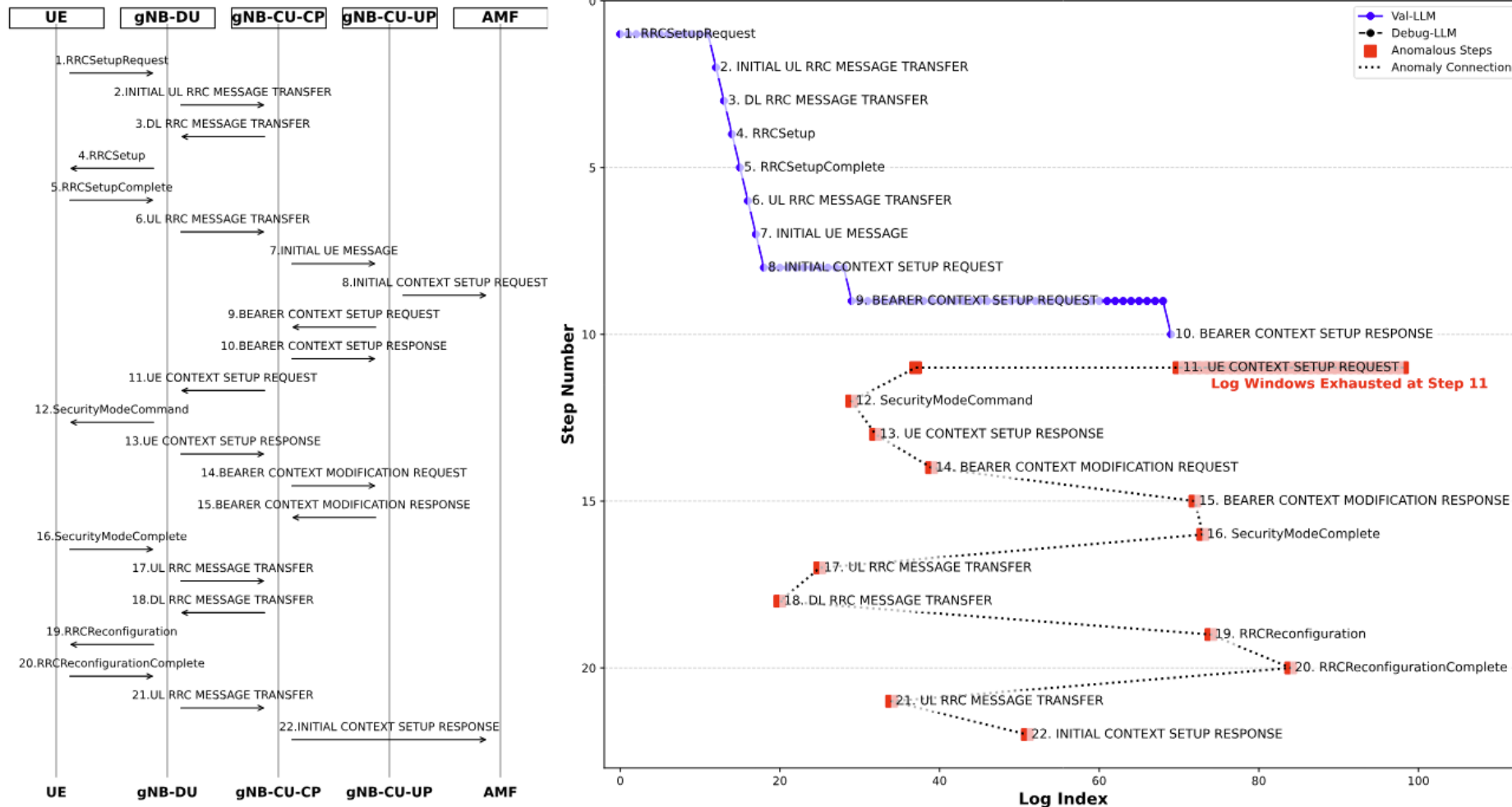
**Confidence Score:**

100%

**Explanation:**

The log file contains the exact message name "UEContextReleaseRequest" which matches the message name in the test case step. Additionally, the log file indicates that the message is sent from the gNB-DU to the gNB-CU, as it mentions "f1ap" which corresponds to communication between the gNB-DU and gNB-CU. The presence of "GNB-CU-UE-F1AP-ID" and "GNB-DU-UE-F1AP-ID" in the log file further validates that the communication occurs between the correct entities. Therefore, the step has been executed as described in the test case.

# Result of Val-LLM and Debug-LLM – TN



TC: UE Initial Attach over E1 and F1 Interface

# Validation (Val-LLM and Debug-LLM)

Test Case Title	Ground Truth	Val-LLM	Debug-LLM	Inference
Initial UE access – UE Context Creation, Service Request.	Partial Pass	Fail	Partial Pass	Incorrect chronology in signaling sequence (The signalling message where gNB-CU should send an INITIAL CONTEXT SETUP RESPONSE to the AMF is executed prematurely)
Initial access – UE Context Creation for Initial Registration.	Partial Pass	Fail	Partial Pass	Incorrect chronology in signaling sequence (The signalling message where gNB-CU should send a UL NAS TRANSPORT (Registration Complete) to the AMF is executed prematurely)
Registration Update without Follow-on Request.	Pass	Pass	-	-
gNB-CU Initiated UE Context Modification.	Pass	Pass	-	-
gNB-DU Initiated UE Context Release.	Pass	Pass	-	-
F1 Setup for NR.	Pass	Pass	-	-
UE Initial Access over F1	Partial Pass	Fail	Partial Pass	Incorrect chronology in signaling sequence (The signalling message where gNB-CU should send an INITIAL CONTEXT SETUP RESPONSE to the AMF is executed prematurely)
Bearer Context Setup over F1-U.	Pass	Pass	-	-
RRC Connected to RRC Inactive.	Pass	Pass	-	-
PDU Session Establishment.	Fail	Fail	Fail	Incorrect message name (PDU SESSION RESOURCE SETUP REQUEST instead of PDU SESSION RESOURCE REQUEST)
UE Initial Access over E1 and F1.	Partial Pass	Fail	Partial Pass	Incorrect chronology in signaling sequence (Multiple discrepancies in the expected chronology)
gNB-CU-UP Initiated Bearer Context Release over F1-U.	Pass	Pass	-	-
UE Initial Access over F1: Simulated Failure (5GC Crash)	Fail	Fail	Fail	Both algorithms converge early due to the absence of the INITIAL CONTEXT SETUP REQUEST message from the AMF to the gNB-CU, leading to an incomplete signaling procedure
UE Initial Access over F1: Simulated Failure (IMSI Mismatch)	Fail	Fail	Fail	Both algorithms converge early due to the absence of the INITIAL CONTEXT SETUP REQUEST message from the AMF to the gNB-CU, leading to an incomplete signaling procedure
UE Initial Access over F1: Simulated Failure (USIM Algo mismatch)	Fail	Fail	Fail	Both algorithms converge early due to the absence of the INITIAL CONTEXT SETUP REQUEST message from the AMF to the gNB-CU, leading to an incomplete signaling procedure

# Val-LLM Demo

localhost:8501

## AI5GTest: LLM based Automation RAN Testing

### Validation

TC-05: gNB-DU Initiated UE Context Release

DUT: O-CU, O-DU

Specification: O-RAN NR C-Profile Spec v13 - Clause 6.2.2

### Controls & Status

Resume / Start | Pause | Restart | PCAP Load ed

### Hierarchical Log

Step Number	Log Entry
0	
1	
2	

Screen Recorder - Scre.io

SCRE.IO

What do you want to capture?

Only Screen | Screen + Camera | Only Camera

Record audio?

None | Microphone | System + Mic. | System

Start Recording

Very hot weather Now

6:21 PM 6/29/2025

# AI5GTest Timing Evaluation

On average for each test case,

- Gen-LLM: ~33.44 seconds
- Val-LLM: ~3.06 minutes
- Debug-LLM: ~41.873 minutes
  
- AI5GTest total testing time: **~3.5 minutes**, and **~45 minutes** (+ debugging)
- Manual approach, as per [1], **~15 hours per test case.**

[1] Aspire Technology, “Stimulating open ran adoption: Accelerating validation, optimization and deployment,” Aspire Technology, Tech. Rep., 2021, Accessed: March 5, 2025.

# AI5GTest: From Lab to Field

- AI5GTest – an AI-powered, specification-aware testing framework that **automates** procedural flow generation, signaling validation, and root cause analysis.
- AI5GTest significantly reduces **manual effort** and enhances **test consistency** and **scalable validation** of O-RAN implementations.
- Currently being integrated in **ORCID T&E center (and VALOR\*)!**

# Thanks to My Awesome NextG Lab Team!

(and Collaborators)

We are grateful to our sponsors



NextG Wireless Lab @ NC State



Dara Ron, PhD  
(Spring 2024 -)  
POSTDOC



Sayanta Seth, PhD  
(Spring 2024 -)  
POSTDOC



Ta-seen Niloy  
(Spring 2022 -)  
PHD STUDENT



Azuka Chiejina  
(Summer 2023 -)  
PHD STUDENT



Pranshav Gajjar  
(Fall 2023 -)  
PHD STUDENT



Abiodun Ganlyu  
(Spring 2024 -)  
PHD STUDENT



Farzad Mehri  
(Fall 2024 -)  
PHD STUDENT



Osman Mamudu  
(Fall 2024 -)  
PHD STUDENT



Ali Abbas  
(Fall 2024 -)  
PHD STUDENT



Md. Sharif Hossen  
(Fall 2023 -)  
PHD STUDENT (CO-ADVISED WITH DR. ISMAIL OUVENC)



Bibhor Kumar  
(Fall 2025 -)  
PHD STUDENT



Corbin Kim  
(Fall 2025 -)  
VISITING PHD STUDENT



Emmanuel Ojo  
(Fall 2025 -)  
PHD STUDENT



Vaibhav Dandala  
(Fall 2025 -)  
UNDERGRADUATE STUDENT

# Concluding Remarks

**AI-O-RAN is already happening** – C-RAN, V-RAN, O-RAN, AI-RAN ...

- Things look uncertain/chaotic, but, I think it's **good chaos!**
- From **lab validations** to **field trials** all will contribute to the success of AI-O-RAN.

**Similar to Linux and Android, AI-O-RAN will thrive!**

- **University labs, small vendors, hobbyists, and governments!** Niche market, value proposition

**The way OS transformed the entire computer world, and AI-O-RAN will transform the cellular networks!**

- Once AI-O-RAN is deployed, going from 6G to 7G in the future, for example, will probably just be a **software update\*\*!**

**AI-O-RAN is eating the (Telecom) world!**

It just seems to be a matter of time!

# Questions

## Contact

Vijay K. Shah, Ph.D.

[vijay.shah@ncsu.edu](mailto:vijay.shah@ncsu.edu)

[vijay.shah@wisightslab.com](mailto:vijay.shah@wisightslab.com)

[info@neuralsmith.com](mailto:info@neuralsmith.com)