

## **AINA AI-RAN 기술위원회 Kick-Off 회의 개최 결과**

( '26.07.31.(금), AINA 사무국 )

### **I. 회의 개요**

- (일시/장소) '26.07.01.(수), 14:00~17:00/HJ비즈니스센터광화문 세미나룸A
- (논의내용) 위원회 및 WG 소개, 전문가 세미나 등
- (참 석 자) AI-RAN 기술위원회 및 산하 WG, 사무국 등 29명
  - 정희상 실장, 나지현 실장, 문정모 기술총괄, 최성우 책임(이상 ETRI), 김동구 교수(연세대), 전형준 연구소장(유캐스트), 전병민 책임, 최태영 수석, 배성포 수석, 박종원 팀장, 전보라 책임(이상 TTA), 강문모 팀장(솔리드), 지승환 CTO(LIG Accuver), 정상민 매니저, 이재현 매니저(이상 SKT), 송준혁 수석(삼성전자), 육영수 상무(노키아), 이창수 책임(LG전자), 박병성 디렉터(에릭슨), 차귀춘 부장, 김찬기 차장, 유광열 이사(이상 키사이트), 박재우 수석(KT), 김창훈 책임(LG유플러스), 박규태 센터장(HFR), 손세일 책임(KCA), 김용필 상무(RedHat), 박은복 상무(진인프라), 김현수 과장(KANI)

### **II. 회의 결과**

#### **□ AI-RAN 지능화 WG 활동계획 : 정희상 실장(ETRI) (발표자료 참고)**

- 위원회 구성 및 AI-for-RAN 정의
  - AI-RAN 기술위원회는 인프라(AI-and-RAN)·컴퓨팅(AI-on-RAN)·지능화(AI-for-RAN) 3개 WG로 구성되고, 이번 회의는 컴퓨팅·지능화 WG 공동 주관으로 개최함
  - AI-for-RAN의 지향점은 ① AI를 통한 RAN 성능 개선, ② 완전 자율운영 실현, ③ OPEX 절감이며, 지난 10여 년간 룰 기반 알고리즘을 학습 기반으로 대체·보완하는 연구가 전방위적으로 진행됨
  - 빔 매니지먼트, CSI 피드백/예측, 포지셔닝 등 일부 유스케이스는 3GPP에서 표준화가 진행 중이며, 표준 반영 후 수년 내 상용망 적용이 예상됨

○ 최근 산업 동향 공유

- MWC Shanghai 2026과 DTW Ignite 2026 등 최근 산업 동향을 공유, DTW 기조연설에서는 통신산업의 위기와 AI라는 기회, AI 네이티브·자율기업으로의 전환 필요성, 레거시(특히 마인드셋) 청산, AI Washing 경계, 신뢰 확보의 중요성이 강조됨

○ 위원회 활동계획 및 산출물

- 올해 4월 정부 과제 'AI-RAN 글로벌 선도 프로젝트'(ETRI 주관, 통신 3사·대학 4곳·중소중견기업 3곳·TTA 등 총 13개 기관 참여)를 수주하였고, 참여기관에는 '국가지정 AI-RAN 전문연구소' 현판이 수여됨
- 주요 산출물로 ①기술백서/이슈리포트 발간, ②기술 로드맵(10월 말 목표, 정부 예산 반영 일정 고려), ③기술교류 워크숍(연 1회 이상)을 제시

○ 주요 질의응답(Q&A) 정리

- Q) 기술 로드맵 발간(10월 말)을 9월 'Korea Real-World AI Network 2026', 11월 AI-RAN Summit/AI-RAN Alliance F2F 미팅 등과 연계할 수 있을지, 특히 NIA 실증 과제 결과를 AI-RAN Alliance F2F에 기고하는 데 기술위원회가 지원할 수 있는지?
- A) 실증 결과를 11월 F2F 미팅에서 기고하기로 함, 기술위원회 차원에서 실증팀의 기고 준비(발표자료·KPI 정리 등)를 지원하는 방향으로 산업융합위원회와 협력하기로 함

□ AI-RAN 컴퓨팅 WG 활동계획 : 나지현 실장(ETRI) (발표자료 참고)

○ AI-RAN 컴퓨팅 WG 개요

- AI-RAN 컴퓨팅 WG는 AI-RAN Alliance의 WG3(AI on RAN)에 대응하고, 인프라(AI-and-RAN)·지능화(AI-for-RAN)와도 영역이 상당 부분 겹쳐 있어 향후 조인트 회의 형태로 함께 운영할 예정

○ AI-RAN Alliance 구조 및 AI-on-RAN 백서 내용

- AI-RAN Alliance는 WG1(AI for RAN)·WG2(AI and RAN)·WG3(AI on RAN) 3개 워킹그룹과 Data4AI·Test Methodology·Agentic AI·AI-RAN Commercialization 4개 테스트그룹으로 구성, Agentic AI·Commercialization TG는 올해 초 신설됨
- Alliance가 발간한 AI-on-RAN 백서에 따르면 모바일 네이티브 AI가 부상 중이며(유럽 모바일 GenAI 활성사용 약 20%, 2030년 50% 이상 전망), 다운로드·버퍼링 중심의 기존 서비스에서 업링크 중심·지연 보장 중심으로 패러다임이 전환되고 있음
- 'Best-effort' 방식으로는 AI 서비스 품질을 보장할 수 없고, AI 서비스는 품질을 즉시 체감하게 만들기 때문에 3GPP·O-RAN·GSMA/CAMARA 등 기존 산업계 기술 요소를 연계한 공동 전략이 필요하다는 것이 백서의 핵심 주장임
- Alliance WG3의 목표는 ① 네트워크 요구사항 정의, ② 실측 기반 벤치마킹, ③ 6G AI-네이티브 네트워크 요구사항 정의 3가지이며, Uplink Optimization부터 Phased Deployment까지 6단계 배치 전략과 Consumer·Enterprise·Government 3개 분야 유스케이스를 제시, MWC 2026 데모(로봇 내비게이션 등)와 NEU 5G Glasses·SKT V2N 등 Work Item이 진행 중임

#### ○ AI-RAN 컴퓨팅 WG 활동계획

- 컴퓨팅 WG는 Alliance WG3 논의를 참고하여 ① 국내에서 기지국이 AI/Gen-AI/LLM/VLM을 구동하기 위한 네트워크 요구사항 파악·정리, ② 실시간 추론·엣지 AI 서비스 제공 방법(시나리오·유스케이스 발굴), ③ 표준화는 각 기관 소관이므로 관련 동향 공유 위주로 활동 방향을 잡고 있음
- 모네타이제이션을 중요한 과제로 보고 있으며, 산출물은 지능화 WG와 별도로 내지 않고 하나로 통합하여 발간할 예정임

#### ○ 주요 질의응답(Q&A) 정리

- Q) 하이퍼스케일러 등 제3자 사업자의 요구사항도 컴퓨팅 WG가 수용하는 절차가 있는지?

- A) 5G 버티컬 요구사항 수용 사례처럼 향후 폭넓게 수용할 과제로 보고 있음, 특정 반도체·플랫폼 업체(NVIDIA 등) 영향력이 커 공정성 문제도 제기되고 있음

□ O-RAN, AI-RAN 핵심 이슈 : 송준혁 수석(삼성전자) (발표자료 참고)

○ 3GPP AI/ML 표준화 현황(5G·6G)

- Release 19: UE-side AI/ML 모델 지원을 규격화하였으며 대상 유스케이스는 빔 매니지먼트·CSI 예측·포지셔닝 3가지로 한정됨, 규격은 기본적으로 UE 관점에서 기술되며, 데이터 수집·학습 개시/중단 등은 UE Assistance 메시지 절차로, 최종 승인권은 gNB가 보유
- TS 28.105에 ML 모델 라이프사이클(학습-테스트-추론 애플리케이션-배포-추론)이 정의되어 관리(LCM) 프레임워크가 마련되었으며, 이를 통해 단말 지원 시 네트워크 측(AI-for-RAN)까지 폐루프(Full circle) 지원이 가능해짐
- 6G 아키텍처 핵심 고려사항은 ① Unified Data Framework, ② ISAC(통신·센싱 통합), ③ AI 기반 Management Agent 3가지임, 코어-RAN 인터페이스의 P2P/SBI 방식, CU/DU 및 CP/UP 분리(Disaggregated RAN, HLS) 등은 아직 열려 있는 이슈임

○ O-RAN의 AI/ML 지원 한계

- O-RAN의 핵심은 SMO와 프론트홀(O-RU)이며, SMO 내 'AI/ML Workflow' 기능은 정의만 되어 있고 실제 데이터 전처리, O1 인터페이스를 통한 파라미터 교환 등은 규정되지 않아 현재 구조로는 개방형 AI/ML 지원이 사실상 불가능함
- AI/ML·라이프사이클 관련 용어(ML 워크플로, MLOps, LCM 등)가 3GPP·O-RAN·업체마다 혼용되고 있어 3GPP 정의(TR 21.905 기반)로 통일할 필요가 있음

○ OCUDU 등 오픈소스 동향

- OCUDU(미국방부 FutureG 펀딩 기반 오픈소스 CU/DU, Linux Foundation 이관)는 NVIDIA 등 GPU 진영이 참여를 주도하고 있어 향후 기술 트렌드로서 주시가 필요함

○ 주요 질의응답(Q&A) 정리

- Q1) 6G AI-네이티브에서는 AI가 빌트인되는 구조인데, Massive MIMO·ISAC 등도 벤더 종속 없이 O-RAN 기반으로 구현 가능한가?
- A1) O-RAN 구조로는 어려움. 5-인터페이스(프론트홀) 규격이 아직 정의되지 않아, 6G 논의를 계기로 새로 정립해야 함
- Q2) 여러 오픈소스 중 OCUDU가 특히 주목해야 할 이유는? Massive MIMO 등 상용급 기능은 내년 내 구현 가능한가?
- A2) NVIDIA 참여로 기존 학술 수준 오픈소스를 실용 수준으로 끌어올릴 잠재력이 핵심, 구체 기능 구현 시기는 단정하기 어려움
- Q3) OCUDU 관련 메이저 벤더들의 대응 전략은? '오픈'이라면서 실제로는 벤더별 히든 요소가 있는 것 아닌지?
- A3) 현재 OCUDU 회의에서 메이저 벤더는 거의 발언하지 않고 모니터링 단계임, 결과물 완성도는 더 지켜봐야 함

□ AI-RAN Alliance F2F Meeting 결과 및 관련 기술 동향 : 정상민 매니저(SKT) (발표자료 참고)

○ AI-RAN Alliance F2F 미팅 개요 및 Alliance 조직 재편

- 2026 Q2 F2F(5.26~29, 美 Chandler, VIAVI 호스팅) 기준 회원사는 140개사 이상, 누적 데모는 33건, 참여 Tier-1 MNO는 11개(연말 20개 목표), Board F2F, TSC/MSC, WG1·2·3 병렬세션, Operator Panel, Valor Lab Tour로 세션이 구성됨
- Alliance는 '데모 공장'에서 탈피하여 신뢰성 있는 Benchmarking·Deployable Blueprint 중심 단체로 정체성을 재정의한다고 공식 선언, 자체 기준 발표인 33개 데모의 결과 신뢰성 부족 문제가 배경임

- 결과물 책임 강화를 위해 WG별 Work Item을 Cluster 단위로 재편 (Feature Lead 지정)하고, Task Group은 실행 중심으로 개편 중임, (TMTG 일시 중단, D4AI는 MNO 데이터 접근 모델 마련에 집중)
- 3대 Pillar(AI-for-RAN/AI-and-RAN/AI-on-RAN)별 핵심 어젠다 및 보안
  - 3대 Pillar의 핵심가치와 대표기술은 다음과 같음: WG1은 에너지 절감·운영자동화(RAN Perf Leaderboard, 5종 KPI: Throughput·P99 Latency·Joules/bit·컴퓨팅 소비·Real-time), WG2는 Inference-as-a-Service (Orchestrator 동적 배치, dUPF), WG3는 Latency-critical Physical AI(ISAC, 3대 비즈니스 모델).
  - 보안 측면 6개 Attack Surface(Supply Chain, Inference/Control Data Injection, RIC/xApp/rApp, Shared GPU, Telemetry Integrity, Cloud-native Platform)에 대한 Zero Trust 프레임워크를 WG2 주관으로 연말까지 수립 예정
  - Operator Panel(SKT·SoftBank·Vodafone·T-Mobile)에서는 에너지 소비 우려, Agentic RAN에 대한 신뢰(투명성·롤백·Human Oversight 전제), '수요 기반 유스케이스 없이는 모네타이제이션 시점 논하기 어렵다'는 점이 공통 논의됨
- 정부협력 및 벤더·생태계 동향(NVIDIA·OCUDU)
  - 美(NTIA·NIST·FCC·DoW FutureG)·亞(日 MIC, 대만 ITRI, 韓 AINA)·EU 등과 정부협력이 활발하며, DoW FutureG(3천만불, OCUDU 지원)·NTIA NOFO4(약 5천만불)·ISAC RPP 등 자금 프로그램이 진행 중임
  - NVIDIA는 AI Aerial·AI Grid·Genesis 등으로 생태계를 확장하며, 이를 통해 궁극적으로 컴퓨팅 수요를 선점하려는 전략으로 파악됨, 제조사별로 에릭슨(자체 실리콘), 노키아(NVIDIA GPU 풀가속), 삼성(vRAN·멀티실리콘)으로 전략이 분화되어 있음
  - OCUDU는 美 DoW FutureG 펀딩으로 출발, 올해 3월 Linux Foundation 산하 편입. O-CU/O-DU 풀스택을 BSD 라이선스로 제공하며, AI-RAN Alliance가 정의한 유스케이스를 재현 가능한 코드로 구현·검증하는 공통 기반 역할을 지향함

○ 주요 질의응답(Q&A) 정리

- Q1) OCUDU와 Alliance가 미국 정부 주도로 한 몸처럼 가는데, 다른 나라들이 이견 없이 지켜만 보는 게 이상하지 않은지?
- A1) 보드 미팅에서 공개적 이견은 없었음, 제조사 대부분은 OCUDU 자체를 반대하는 입장은 아닌 것으로 보임
- Q2) '데모 공장이 되지 않겠다'는 선언이 실무자 입장에서는 혼란스러울 수 있는데, 실제 의미는?
- A2) 자체 시험결과의 신뢰성 문제가 출발점, 공통 코드(OCUDU)와 공통 벤치마크(RAN Perf Leaderboard)로 검증받아야 공인 데모로 인정하는 방식으로 전환한다는 의미임
- Q3) 클러스터 구성과 워킹그룹별 참여 규모는? 서로 다른 AI-for-RAN 알고리즘 성능을 수치로 비교할 수 있다고 보는지?
- A3) WG1은 참여가 활발하나 WG2·3는 아직 체어 중심. 알고리즘 간 정량 비교는 쉽지 않을 것으로 예상됨
- Q4) 정부기관(NTIA 등)이 Board Member로 공식 편입되는지, FutureG와의 관계는?
- A4) 아직 논의 중이며 공식 보드는 아님, FutureG는 WG2·3와 사실상 결합되어 있음
- Q5) 국내 기업이 미국 주도 생태계(Alliance·OCUDU·FutureG)에 '올라타야' 하는지, 모니터링이 맞는지?
- A5) 의견(적극 참여 vs 모니터링)이 분분하여 지속 관찰이 필요함

□ 후속조치 계획

○ AI-RAN 기술위원회 & 산업융합위원회 공동회의(8월)

- (일시/장소) '26.08.06(목) 14:00~17:00/HJ비즈니스센터광화문
- (주요안건) 글로벌 기업 및 표준기술 동향, 오픈랜 실증 추진현황 등

### Ⅲ. 회의 전경



**인사 말씀**  
AINA 김동구 운영위원장(연세대, 교수)



**AI-RAN 지능화 WG 활동계획**  
AINA 정희상 AI-RAN 기술위원장(ETRI, 실장)



**AI-RAN 컴퓨팅 WG 활동계획**  
AINA 나지현 AI-RAN 컴퓨팅 WG장(ETRI, 실장)



**O-RAN, AI-RAN 핵심 이슈**  
송준혁 수석(삼성전자)



**AI-RAN Alliance F2F Meeting 결과 및 관련 기술 동향**  
정상민 매니저(SKT)



**단체 사진**

끝.