LF Networking
AI Task force

Update - July 2024

The challenge of PII in Telco data sets

Good AI models require high quality network data

Training Telco AI models has to be performed on actual Telco data

Raw Telco data sets contain personally identifiable Information (PII)

- Names (Systems, Domain, Individuals, Organizations, Places, etc.)
- Address (IP and MAC)
- Telco Fields - IMSI, IMEI, MSIN, MSISDN, MCC+MNC
- Location Data (GPS, Cell-ID, Count, etc.)
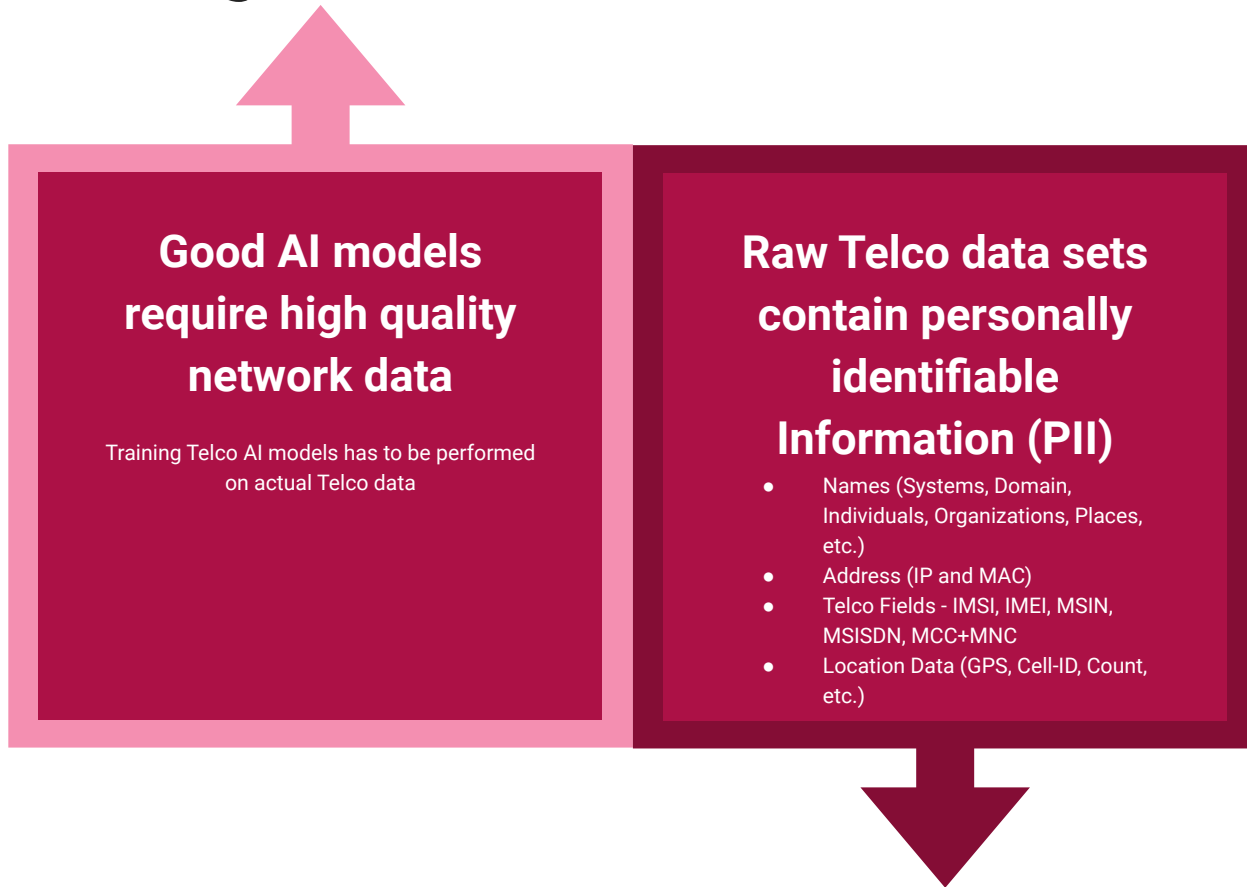
What does the Anuket/Thoth project do?

Agree on what constitute the 'sensitive' data. Agree on the problem set (questions we would want to answer)

Try available tools (Libraries) and techniques (implementations) on the available datasets.

Find the gaps in datasets, tools and techniques.

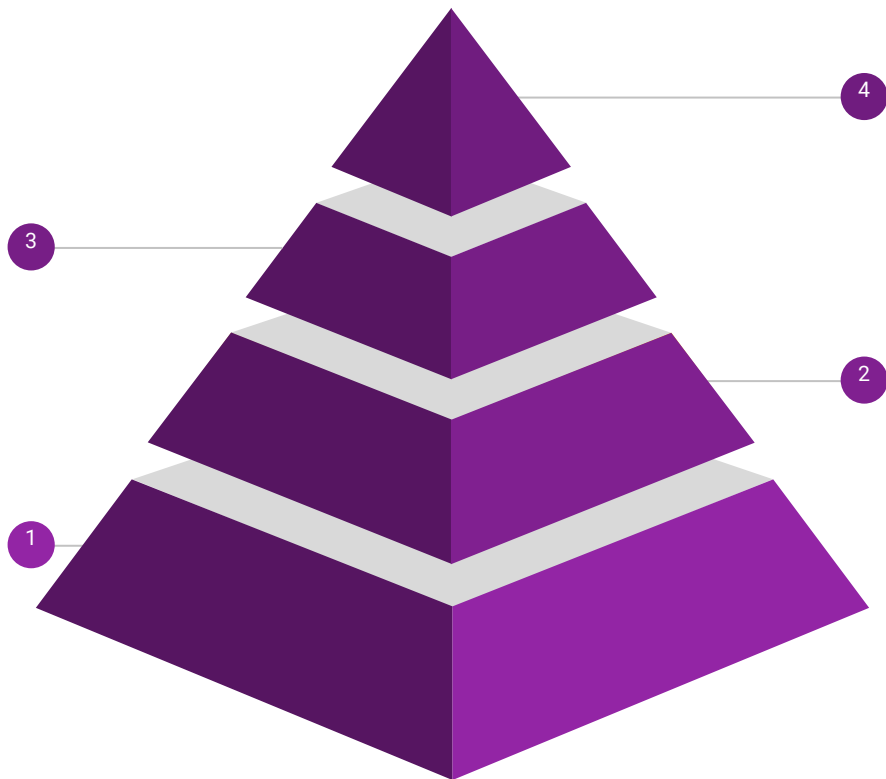**Fill** those gaps considering the problem-set.

Publish the results.

What are the techniques we are trying?

GANs
Synthetic data generation as a perfect anonymization solution.

4

Autoencoders
Unsupervised techniques for the anonymization

3

Natural Language Processing
NLP techniques for the Logs.

2

Classic Techniques
K-Anonymity, L-Diversity, T-Closeness, Differential Privacy

1

Why do we need this project?

- There are many available tools and open source projects that provide the generic capabilities required for data anonymization

- However, up until now, there has not been a solution to address the unique data of Telecom network, and the generic techniques do not just work out of the box.

- The Thoth project is focusing on developing this Networking domain specific solution on top of the generic tools

Focus Area 2: Intent based network automation using LLMs
A blueprint using ONAP and Nephio

Blueprint Goals

Req 01
Intent Processing Enhancement

Enhancing intent processing capabilities by introducing large models

Req 02
Large Model Convergence

Provide a convergence platform for large models to integrate various capabilities of multiple large models

Req 03
Agent Construction

Building intelligent agents based on large models, and providing the ability to handle cross-domain complex tasks
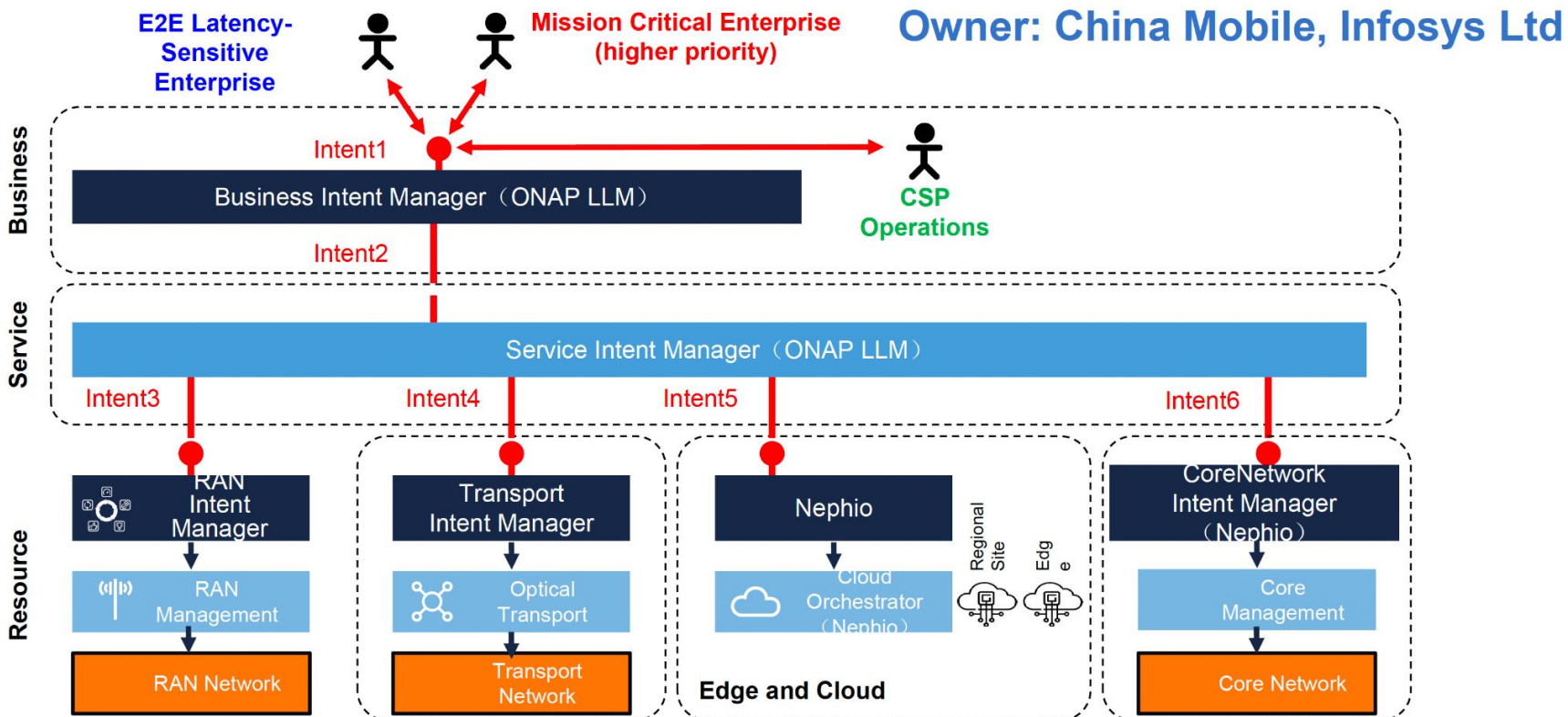
Req 04
Multi-domain Support

End-to-end intent-driven autonomous network, supporting intent processing across different domains such as RAN, Transport, Cloud, and Core.

Project Overview

Focus Area 3: Leveraging AI to augment project functionality
Nephio GenAI example

What GenAI means for Nephio?

Nephio User Automation Simplification

Templates generation
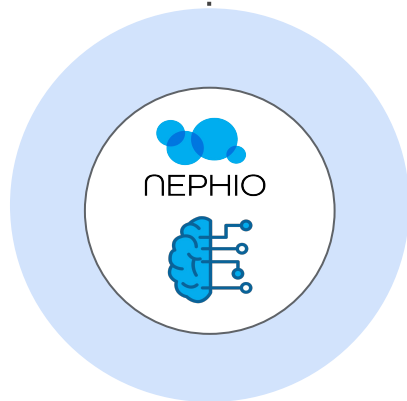GenAI based Operators, CRDs, TOSCA, KRM etc.

Templates hydration
GenAI based data-fill for various environments and context

SDKs and APIs
GenAI based SDKs and APIs creation

Nephio Services Closed Loop

Cloud Optimization
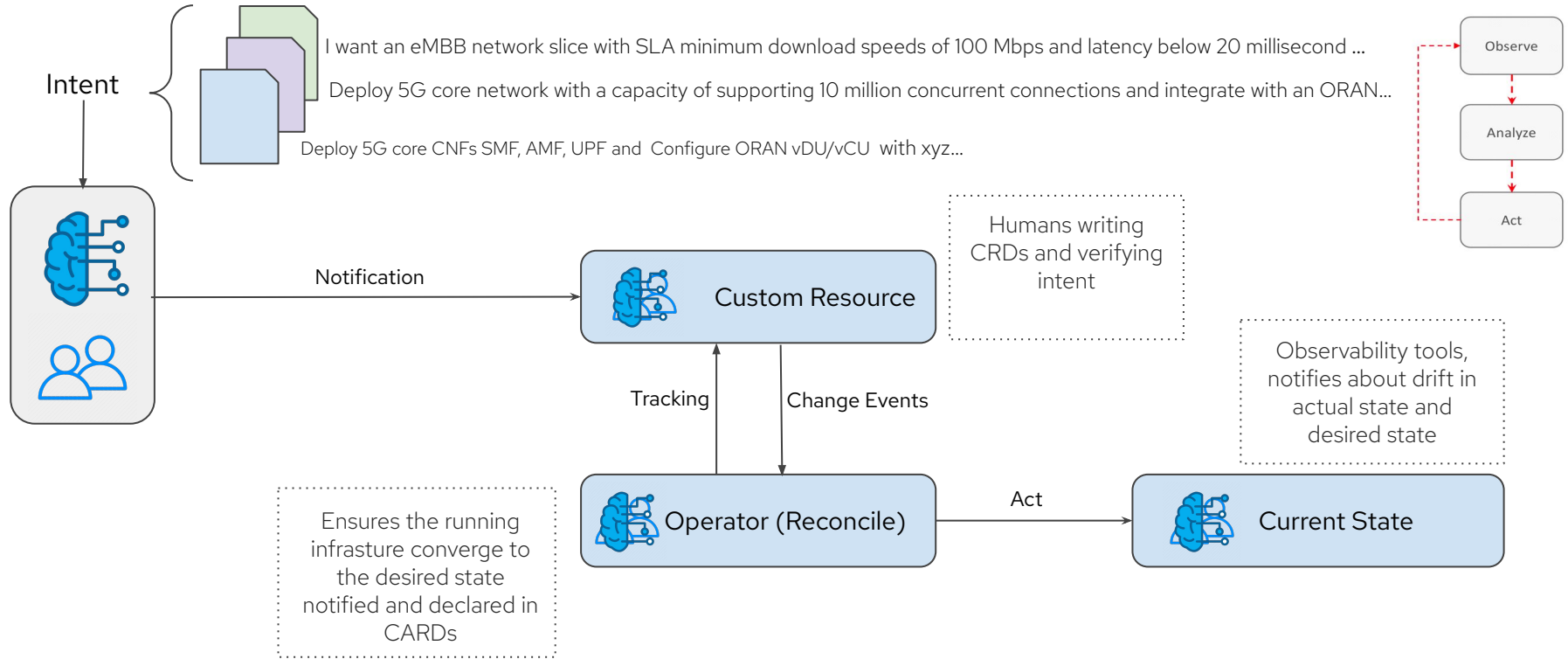AI optimizing cloud capacity energy and cost

5GC, RAN and Edge
AI optimizing performance, efficiency and reliability
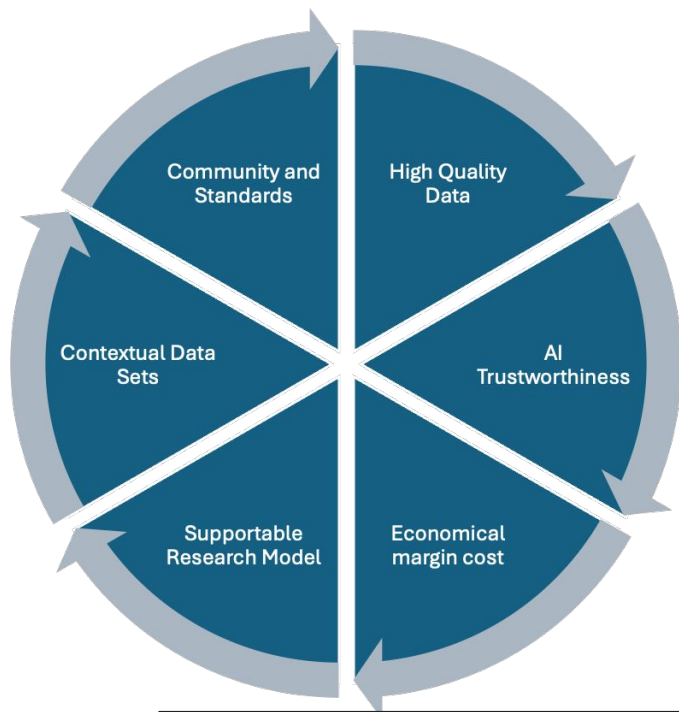
Network Operations
GenAI provisioning and troubleshooting agents

NEPHIO

Nephio GitOps and GenAI - A match made in Heaven!

Focus area 4: Strategy

Whitepaper published during ONE Summit

- **High Quality Structured Data** - Avoiding "information islands" that cannot be interpreted
- **AI Trustworthiness** - To enable full automation and taking humans out of the equation
- **Economical marginal cost** - The cost for any single organization to build models is too high
- **Supportable Research Models** - Resources must be pools to become cost effective
- **Contextual Data Sets** - Coming from all layers - Application, Security, OSS/BSS, etc.
- **Community Unity and Standards** - To avoid limited "field of view" of a single vendor solution

Open Source Collaboration is the only way to address these challenges

LFN AI Taskforce

- There is a consensus that work should continue at the projects level, with the taskforce serving as a coordinator

- We have participation from several of the key initiatives (Thoth, Nephio) but struggling with others (ONAP)

- Meeting cadence is now monthly