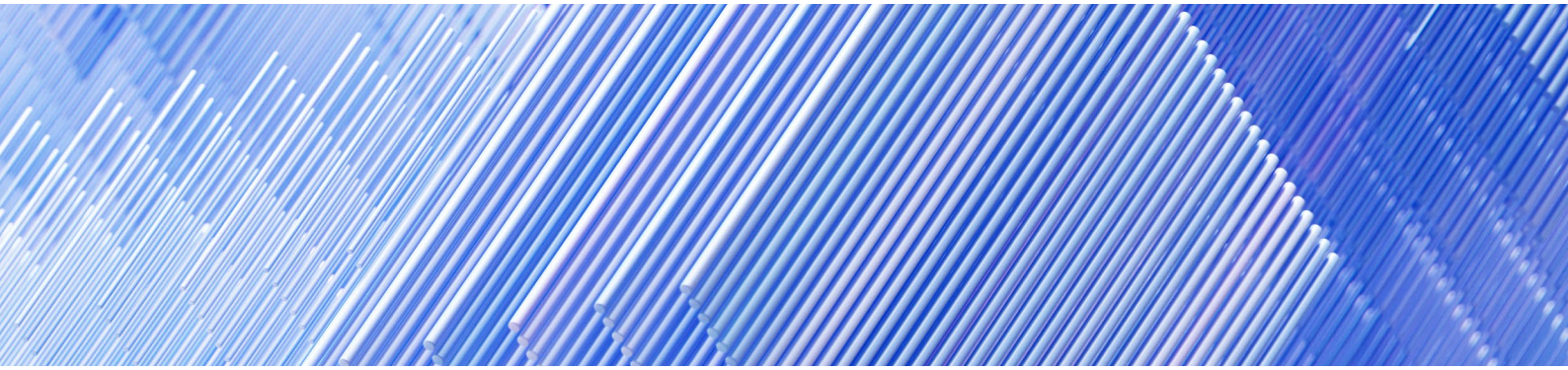
Data foundation for Autonomous Networks

Nokia Data Suite

White paper

.



The biggest obstacle on the way to Autonomous Networks

Autonomous networks are networks that, under most conditions, can self-configure, self-monitor, self-optimize, and self-heal. TM Forum's autonomous network framework defines the different levels of network automation. Level 4 (highly autonomous) and Level 5 (fully autonomous) represent the goal for many CSPs, where their network anticipates the needs and self-configures for optimal performance. This is done through the increasing use of analytics and AI (Artificial Intelligence) that allows the network to take certain low-risk actions proactively or prescribe a series of actions to a human operator. These include activities such as optimizing network resources, automating service operations, and ensuring the customer experience.

The emergence of **Agentic AI**, a new paradigm in AI development, further underscores the fast transition to more autonomous operations. Agentic AI systems are designed to be more autonomous and self-directed, capable of learning and adapting to new situations without constant human intervention.

While operators are investing in AI and GenAI (Generative AI) to gain a competitive edge in driving toward Level 5 autonomous networks, solving the top challenges hindering autonomy's progress requires a foundational change. Read more about Nokia's Sense, Think, Act strategy for Autonomous Networks here.

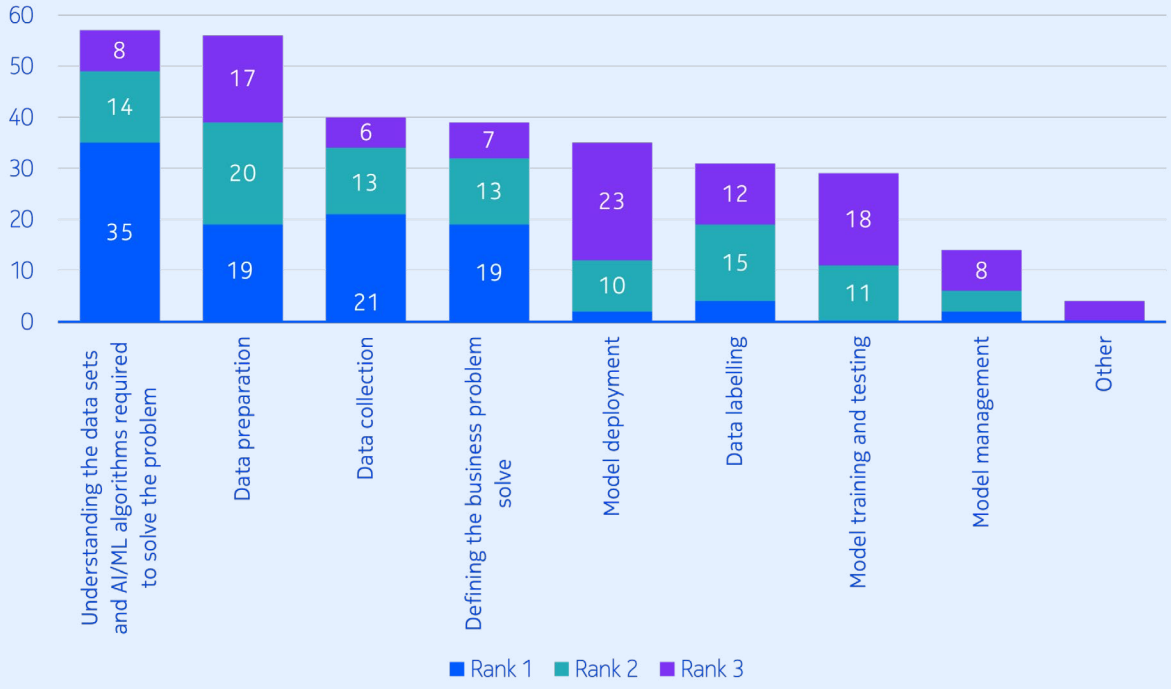
The biggest obstacle – missing the solid data foundation

The ability to effectively harness the vast amount of raw network data and leverage it for AI use cases remains a significant obstacle for CSPs. TM Forum's latest AI research highlights this challenge, stating that access to clean, high-quality, and usable data is a major obstacle to deploying AI at scale. This lack of data quality can lead to a "garbage in, garbage out" scenario, where AI models produce unreliable results, undermining trust in their outcomes.

The implications for Generative AI (GenAI) are particularly noteworthy. Infusing relevant telco data into foundational models and employing prompt engineering and fine-tuning techniques is crucial to ensure accurate responses and prevent "hallucinations" – instances where GenAI models generate incorrect or misleading information. For CSPs to fully embrace AI-driven automation, it is essential that the outcomes of AI use cases are demonstrably accurate and reliable.

Analysys Mason's research reveals this data challenge as well—a primary challenge for CSPs in using AI is limited access to high-quality data, cited by 21% of respondents. Existing data infrastructures exacerbate the issue with siloed systems and proprietary interfaces. These data hurdles impact CSPs' AI talent efforts, with data scientists spending 70% of their time on access and preparation, limiting innovation.

Moreover, the inefficiency in utilizing data scientists' time is compounded by a major shortage of skilled professionals in the field. TM Forum's AI research indicates that the demand for AI experts is second only to cybersecurity.



Stages considered or expected to be most challenging by CSPs (Analysis Mason 2023)

Here are some essential characteristics of AI-ready data:

- **Clean, prepared data** – Ensure data cleanliness, consistency, and usability across network and customer experience datasets. Techniques like enrichment, feature engineering, normalization, handling missing values, data labeling, and annotation contribute to data readiness.
- **Sufficient data volume** – AI and machine learning thrive on vast amounts of data for model training. CSPs need diverse domain-specific datasets to ensure unbiased and accurate AI results.
- **Real-time data** – Business critical use cases often demand real-time data.
- **Common data model** – A unified data model is essential for integrating structured and unstructured data from various vendors. It streamlines data processing and facilitates cross-domain insights.
- **Data Governance** – Implementing policies for data quality, access control, privacy and security.

Data Architecture matters

Today, the major adoption of data architectures is a data lake or a data warehouse, with real-time data availability and stream processing, with the goal of ingesting, enriching, transforming, and serving data from a centralized data platform. However, the popular data lake architecture falls short in a few ways:

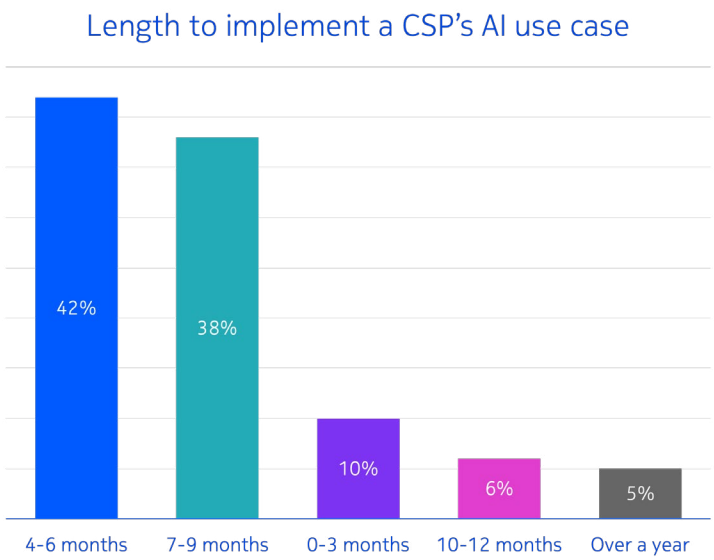
- A central ETL/ELT pipeline gives teams less control over increasing volumes of data
- Different data use cases require different types of transformations, putting a heavy load on the central platform

While data lakes work well with batch-oriented workloads, their monolithic storage is not well suited to low latency, highly demanding, real-time business applications. Furthermore, data lakes are hard to scale. Some CSPs add resources to stretch legacy data capacity as demand grows, but eventually, a breaking point will be reached with a backlogged data team struggling to keep pace with the business's demands.

A more modern and agile approach is the data mesh, which fundamentally differs from this approach - instead of consolidating data into a monolith, it relies on data federation and data abstraction to virtually combine data, called data products, from dispersed sources for increased agility. Ready-to-use data products are designed to serve specific user or business needs, providing data in a consumable and meaningful way. Data mesh is a revolutionary approach that decentralized data architecture, data ownership, and management, breaking down data silos and empowering teams to work more collaboratively and efficiently with their data. This allows the facilitation of greater data experimentation and innovation while lessening the burden on data owners to field the needs of every data consumer through a single pipeline.

Scaling AI further with MLOps

However, to fully realize the potential of these data products, a robust MLOps (machine learning operations) platform is essential. The average duration to implement a telco AI use case is 6–7 months due to building models from scratch, data access constraints, and skillset acquisition. An MLOps platform streamlines the AI development lifecycle, enhances collaboration and efficiency, automates repetitive tasks, ensures scalability and reliability, facilitates continuous integration and deployment, and enhances model governance and compliance. By integrating data mesh architecture with a great MLOps platform, CSPs can accelerate AI use case development and drive significant business value.



Length to implement a CSP's AI use case (Analysis Mason 2023)

Nokia Data Suite – Data products & AI studio to solve the data challenge and accelerate CSPs AI innovation

Nokia Data Suite is a product designed to address telecommunications data challenges by leveraging Nokia's extensive expertise in the field. As the No. 1 provider in the Network Automation Software segment for 2023, according to Appledore Research*, Nokia brings years of experience and a wealth of AI use cases in telco networks to the table. This expertise is further enhanced through intensive research on AI/ML in collaboration with Bell Labs.

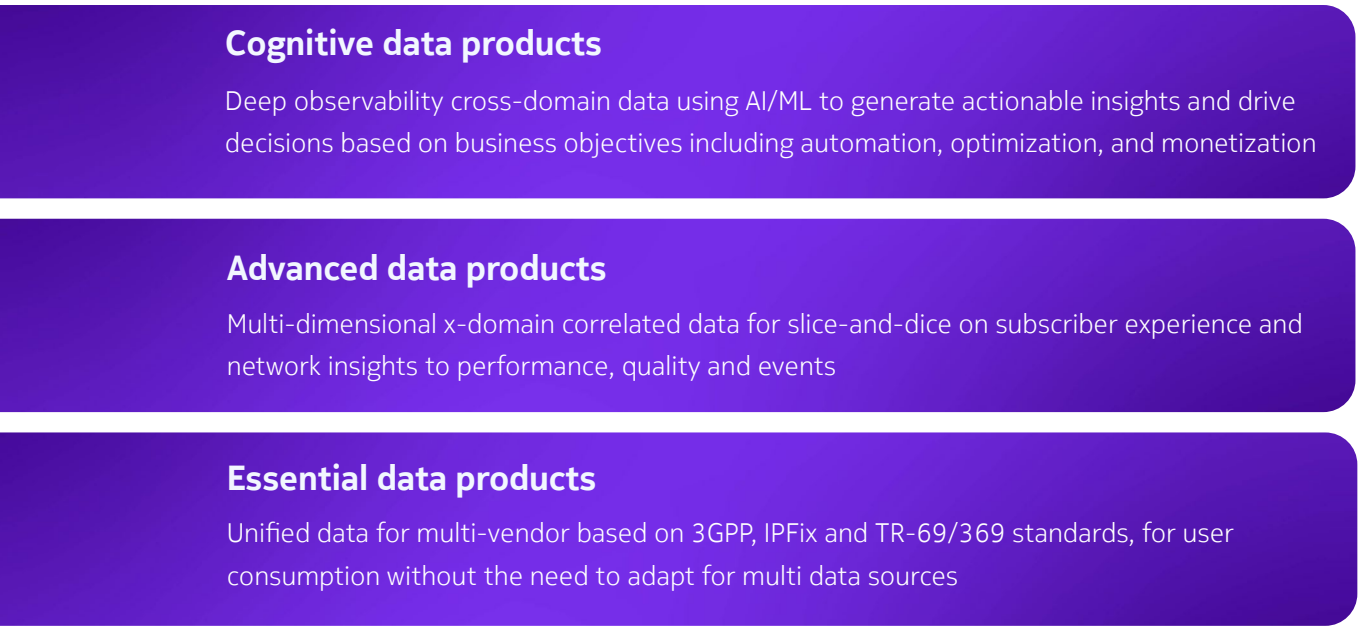
The cloud-native Nokia Data Suite empowers CSPs by integrating expertly curated telecom data products with a robust AI studio. This platform accelerates the development of AI use cases, paving the way for Autonomous Networks.

Nokia Data Suite's data products

Nokia Data Suite offers standardized and ready-to-use data products that transform raw data into advanced and cognitive assets by incorporating domain knowledge and intelligence. These reusable data products provide high-value actionable insights to drive customer business objectives, such as automating operational workflows, improving quality of experience (QoE), reducing energy consumption, and optimizing network performance. They offer a comprehensive end-to-end view of subscriber experience and network observability, and they also enable the drive towards network autonomy.

By eliminating complexity and reducing the time required for data preparation, these data products facilitate data scientists in accelerating the development of new AI or GenAI applications. Leveraging Data Suite's data products can make the MLOps lifecycle up to 70% faster.

Nokia Data Suite classifies data products into three categories: essential data products, advanced data products, and cognitive data products. Data Suite also allows users to create their own data products through the AI Studio from existing data products or external data sources and publish them for external consumers.



The four network domains of data products in Nokia Data Suite include:

- Network Domain: Mobile Networks Data, including OSS performance counters, fault management, configuration data, RAN Cell Traces data, Control Signaling and User Plane data from Core interfaces.
- Fixed Domain: Fixed Networks Data, including Access and Home performance counters.
- Subscriber Domain: Applies to both Mobile and Fixed Networks, with the difference on the data containing subscriber identity linkage to the applications (e.g. Voice/Video/Gaming/OTT) used on KQIs (e.g. Quality/Usage/Failures/Signaling/Speedtest).
- Energy Domain: Data focusing on Sustainability (e.g. Site infra, RAN active, UE subscribers, Fixed domain, Core domain).

The data products are built by ingesting and correlating data from data sources across multi-vendor, multi-domain networks. Using our deep telco domain knowledge and AI/ML capabilities, we transform the raw data from the network into high-value advanced and cognitive data products. The underlying architecture is designed for scale and resilience.

This streamlined process empowers CSP's data scientists, enabling them to swiftly develop and deploy new AI/ML applications based on the unified data model while eliminating the time to adapt to different vendor-specific data sources & data fields.

Nokia Data Suite's AI studio

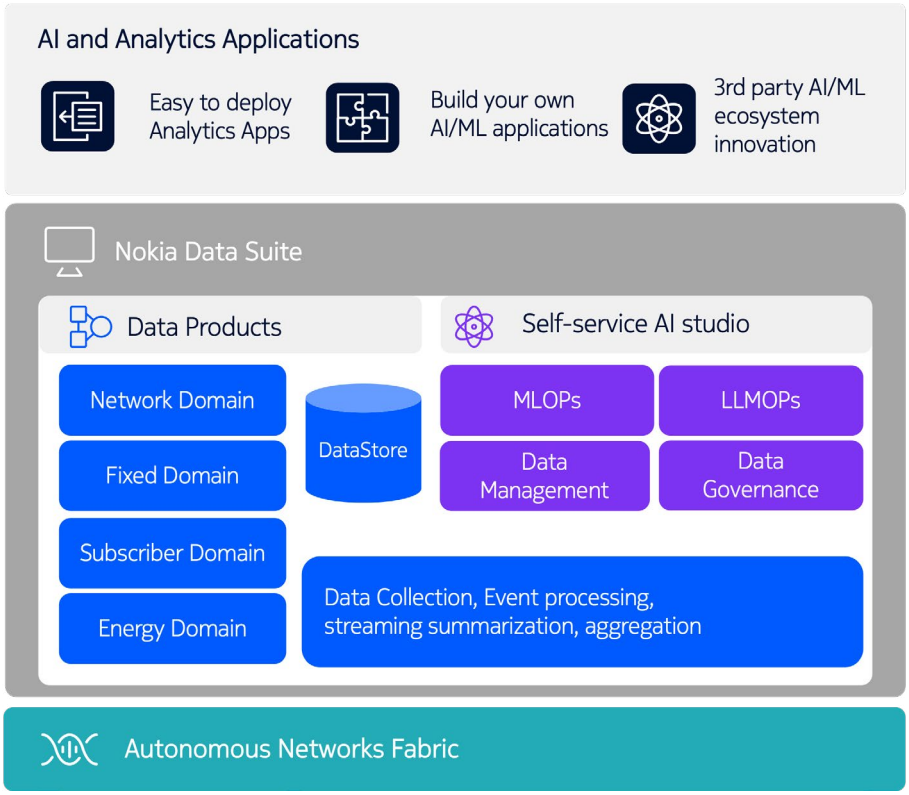
Nokia Data Suite is also equipped with a telco-centric self-service AI studio, providing an MLOps and LLMOps framework, including pre-packaged AI models for CSPs to build their own AI and GenAI use cases with ease. LLM libraries are available for users to integrate and develop GenAI use cases on existing or custom-built data products .

The AI studio offers comprehensive data governance and security mechanisms to ensure accurate data availability, synchronization, reliability, and relevance throughout the data lifecycle. Key features include a data catalog for managing metadata, data lineage for understanding data dependencies, and a data quality view for maintaining data health. Additionally, the suite provides advanced data processing capabilities, data engineering for building and managing data pipelines, and data access features for optimizing API consumption and data integration operations. The data observability view ensures a holistic perspective of data management processes to maintain trustworthiness.

The AI studio component allows users to create new data products from external data sources via inbound APIs and by transforming existing data products. These custom-built data products are managed in a single pane for quality and accuracy across varied ecosystems, including public cloud and on-premises/edge-based environments. The data products can be published through a data catalog for external access using outbound APIs to end consumers. The output of the data products may include root cause analysis (RCA) or next best actions (NBA) from AI models, which can be shared with an external orchestration engine for consumption. This enables closed-loop actions on the network, contributing to the autonomy of network operations.

Together with the AI studio, Data Suite's data products provide CSPs with a powerful, scalable solution to build AI use cases. Designed to be cloud-agnostic, Data Suite seamlessly integrates with any existing ecosystems and environments. Data Suite is designed to blend into customers' existing ecosystem and environment (e.g., cloud or on-prem) and data products to be consumed by customers' own systems, Nokia, or external applications. This is to minimize data movement and reduce costs for different consumers on same data products.

Nokia Data Suite, as part of Nokia Autonomous Networks applications, is enabled by the underlying Nokia Autonomous Networks Fabric, providing seamless data integration, automation, and security across all network domains, including core, mobile, and transport.



“Nokia has evolved its Nokia Data Suite to address CSPs’ data and AI challenges. The suite offers a unified solution that includes curated network data products, backed by robust data management and governance capabilities, and a self-service AI studio for Non GenAI and GenAI use case development. In addition, the suite has a modular architecture that allows operators to buy components of their choice and integrate them with existing CSP technology. These features make Nokia Data Suite an attractive solution for the telecom market.”

Adaora Okeleke, Principal analyst, Analysys Mason

GenAI-powered Autonomous Agents enhance Data Suite

Nokia Data Suite can be enhanced further with GenAI-enabled Autonomous Agents. With natural language processing, CSP employees—whether network engineers or customer care representatives—can easily retrieve insights and generate reports without requiring specialized technical skills. Based on Data Suite’s data, the Autonomous Agents can automatically fetch the insights needed – truly fast-tracking the journey to Autonomous Networks.

Why Nokia AI and Analytics?

Nokia AI and Analytics software are proven to help CSPs boost productivity, enhance telecom customer satisfaction and reputation, and grow new revenues.

Over 150 CSPs globally rely on Nokia AI and Analytics solutions to unlock the intelligence in their 5G, 4G and fixed broadband networks.

Read more about AI and Analytics solutions.

Abbreviations:

AI=Artificial intelligence

CSP=Communications service provider

GenAI=Generative AI

LLM=Large language models

ML = Machine Learning

QoE = Quality of Experience

RAN = Radio access network

MLOps= Machine Learning Operations

LLMOps= Large Language Models Operations

About Nokia

At Nokia, we create technology that helps the world act together.

As a B2B technology innovation leader, we are pioneering networks that sense, think and act by leveraging our work across mobile, fixed and cloud networks. In addition, we create value with intellectual property and long-term research, led by the award-winning Nokia Bell Labs.

Service providers, enterprises and partners worldwide trust Nokia to deliver secure, reliable and sustainable networks today – and work with us to create the digital services and applications of the future.

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: SRxxxxxxEN (Month) CID214204