



ICON

HORIZON-HADEA

WP2 Use Cases, Physical Requirements and Data Format for Sensing Systems

Deliverable 2.2 Final Report on Data Presentation and Sharing Format

Main Author(s):

Kaida Kaeval, Tallinn University of Technology

Priit Roosipuu, Tallinn University of Technology

Andreas Papadopoulos, CyRIC | Cyprus Research & Innovation Center Ltd

Giorgos Gerontidis, CyRIC | Cyprus Research & Innovation Center Ltd

Aletia Trakakis, CyRIC | Cyprus Research & Innovation Center Ltd

Due date of deliverable: 30th of June 2026

Actual submission date: 30th of June 2026

Deliverable type: R — Document, report

Dissemination level: PU - Public

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them. The project is supported by the SNS-JU.



Funded by
the European Union

Project information

Project Name:	Intent and Context-Aware Optical Networks
Grant Agreement:	101189703
Project Duration:	1 January 2025 – 31 December 2027

List of Contributors

Partner number	Partner acronym	Partner name	Name of contact
7	TalTech	Tallinn University of Technology	Kaida Kaeval
7	TalTech	Tallinn University of Technology	Priit Roosipuu
4	CYRIC	Cyprus Research and Innovation Center Ltd	Andreas Papadopoulos
4	CYRIC	Cyprus Research and Innovation Center Ltd	Giorgos Gerontidis
4	CYRIC	Cyprus Research and Innovation Center Ltd	Aletia Trakakis

Document control

Document version	Date	Author / Organisation	Change
V1.0	03/04/2026	Kaida Kaeval/TalTech	First draft
V2.0	20/05/2026	Priit Roosipuu/TalTech	Second draft
V3.0	15/06/2026	Kaida Kaeval/TalTech	Third draft
V4.0	26/06/2026	Tomas Horvath/BUT	Fourth draft
V5.0	27/06/2026	Aleksandra Kaszubowska/TCD	Final draft

Table of contents

Project information	2
List of Contributors	2
Executive Summary	5
Introduction.....	6
1. Integrated communication and sensing	7
1.1. Background.....	7
1.2. Motivation.....	7
2. Use Cases and Applications	9
2.1. Real-time network sensing and monitoring	9
2.2. Long-term network paths availability evaluation	10
2.3. Future use cases	10
3. Device-specific telemetry and sensing data handling	12
3.1. Data management workflow for data presentation	12
3.2. Current methods of presenting/sharing the sensing data.....	13
4. Requirements of the cloud architecture	19
4.1. Functional requirements for the ICON cloud platform	19
4.1.1. Monitoring.....	20
4.1.2. Fibre sensing	20
4.1.3. Event correlation.....	21
4.1.4. Digital twin	21
4.1.5. AI explainability.....	22
4.1.6. Reporting.....	22
4.1.7. Intent-based control.....	22
4.2. Data format and volume	23
4.3. Data interpretation	23
4.4. Dual-tier storage	24
4.5. Data retention policy	24
4.6. Time synchronisation	24
4.7. Authentication and authorisation	25
4.8. API requirements	25
5. Proposed data flow and ingestion strategy for cloud	26
5.1. Data flow	26
5.2. Ingestion strategies.....	27
5.2.1. Unified ingestion data model for telemetry data	28
5.2.2. Unified ingestion data model for event data	28
5.2.3. Unified ingestion data model for object data	28
5.3. API interfaces in the cloud infrastructure.....	29

6.	Data presentation and sharing in the ICON project	30
6.1.	General presentation principles	31
6.2.	Application-specific presentation	31
6.3.	Reporting and data export	31
7.	Related standards and recommendations	33
7.1.	ITU-T Study Group 15 (SG15).....	33
7.1.1.	Recommendation G.681 (formerly G.681)	33
7.1.2.	Technical Report TR.100 — Feedforward sensing	34
7.1.3.	ION-2030 framework and the ION-aiDC report.....	34
7.1.4.	Related activities in other SG15 questions	34
7.2.	IEEE photonics and sensor standards	34
7.2.1.	IEEE 3101-2023 — DAS interrogator terminology	34
7.2.2.	IEEE 2067-2021 — FBG interrogator terminology	35
7.2.3.	IEEE 952 — interferometric fibre optic gyros	35
7.2.4.	IEEE-SA white paper on fibre optic sensors	35
7.2.5.	IEEE within the ION-2030 framework	35
	Conclusions	36
	References	37
	Annex A:.....	39
	Partner Questionnaire for Data Presentation and Sharing	39

List of Annexes

Executive Summary

This report investigates methods for telemetry and sensing data ingestion, storage, presentation, visualisation, and sharing within the Intent and Context-aware Optical Network (ICON) system, so that a single, network-wide dataset can be exploited across both telecommunication and non-telecommunication applications. The deliverable is the primary outcome of Task 2.3 and builds directly on the use cases and physical sensing specifications established in ICON Deliverable D2.1.

ICON treats sensing as an integral part of the optical network rather than an add-on, transforming the deployed fibre plant into a large-scale integrated communication and sensing (ICAS) manifold. The report shows that today's commercially available tools remain fragmented across four discrete families – carrier network management systems, general observability and artificial intelligence for IT operations (AIOps) platforms, distributed fibre-optic sensing platforms, and digital-twin and energy-management tools – which suffer from domain isolation, opaque automation, and vendor coupling. Overcoming these limitations motivates a unified, single-dataset approach, driven by environmental sustainability pressures, the need for operational context awareness, and the decoupling of sensing hardware from specific applications.

Two primary ICON use cases frame the analysis: real-time network sensing and monitoring for proactive infrastructure protection, and long-term network-path availability evaluation through the Availability Assistant. A mapping exercise conducted across project partners (Brno University of Technology, Trinity College Dublin, LightSensAI, Adtran, Tampnet) documents the current presentation and sharing practices and reveals substantial heterogeneity in data formats (.csv, .json, HDF5, TDMS, FLAC, CBT) and in data volumes, ranging from a few megabytes per day for transceiver telemetry to several terabytes per day for phase-sensitive interrogators.

Building on these findings, the report defines seven functional requirements for the ICON cloud platform – covering monitoring, fibre sensing, event correlation, digital twin, AI explainability, reporting, and intent-based control – together with requirements for data formats, interpretation, dual-tier (edge and cloud) storage, retention, time synchronisation, authentication, and Application Programming Interfaces (APIs). A six-tier dataflow and ingestion model, anchored by the Telemetry/Sensing-as-a-Service (T/SaaS) concept and by unified ingestion schemas for telemetry, event, and object data, is proposed to transport reduced, actionable information from the edge to the cloud while keeping raw traces retrievable on demand.

For presentation and sharing, the report adopts a single-dataset principle, in which every application-specific presentation draws from the same observation stream and sets out general presentation principles (actionable-over-raw, correlation-first, role-aware, responsive, and real-time) alongside application-specific rendering, such as the Geographic Information System (GIS)-based Availability Assistant.

Finally, the report analyses the relevant standardisation landscape, focusing on International Telecommunication Union Telecommunication Standardization Sector ITU-T Study Group 15 – in particular Recommendation G.681, the first telecommunication standard for in-service distributed fibre-optic sensing – and on the Institute of Electrical and Electronics Engineers (IEEE) photonics and sensor standards (IEEE 3101 and IEEE 2067), and positions ICON's data-presentation and sharing model within the ITU-T International Optical Networks towards 2030 and Beyond (ION)-2030 framework. These outputs provide direct input to Deliverable D4.3 and to the project's standardisation activities in WP7.

Introduction

Deliverable D2.2: Final report on data presentation and sharing format is the primary outcome of Task 2.3, "Investigation of the standard method of presenting and sharing of the sensing data".

The main objective of this document is to investigate and define optimal methods for collecting, storing, presenting, and sharing the telemetry and sensing data to maximise its usability across diverse applications. It focuses on facilitating the exploitation of this data for real-time network control (e.g., infrastructure protection) and long-term network path evaluation (e.g., Availability Assistants).

The inputs for this deliverable include:

- Data outputs from various sensing and telemetry devices, including Coherent Optical Time Domain Reflectometry (C-OTDR)/Distributed Acoustic Sensing (DAS) and State of Polarisation (SoP) monitors.
- Current data collection and presentation methodologies from project partners
- External non-optical datasets, such as Automatic Identification System (AIS) maritime tracking and weather station data.

The outputs of this deliverable are:

- An overview of the current presentation and sharing methods.
- General requirements for the future telemetry and sensing data collection, storage, sharing, and presentation.
- Technical specifications for the ICON cloud storage implementation, as well as general guidelines of the data presentation and sharing for the ICON project, and
- Standard compliance analyses and improvement recommendations (WP7 standardization efforts), specifically aimed at organizations like ITU-T SG15.

This deliverable is organised into the following sections:

- Section one provides the background, operational motivation, and requirements for integrating sensing into optical networks.
- Section two introduces the primary ICON use cases (infrastructure protection and Availability Assistants) and future sensing grid applications.
- The third section gives an overview of the device-specific telemetry and sensing data handling, highlighting the desired data management workflow for data presentation and current methods of presenting and sharing the sensing data.
- The fourth section presents the seven functional requirements for the ICON cloud infrastructure. Furthermore, general data format, interpretation, and storage requirements are discussed.
- Section five proposes the cloud storage implementation requirements.
- Section six provides the general presentation and sharing guidelines, and
- Section seven provides an overview of the compatibility with standards and recommendations, which outlines the roadmap for presenting ICON's findings to international standards bodies.

1. Integrated communication and sensing

The ICON project [1] builds the research foundation for a new generation of optical networks where sensing is an integral part of the overall system architecture. Historically, fibre-optic sensing has been treated as a specialised add-on; however, the resurgence of interest in IASC aims to turn existing terrestrial and subsea fibre infrastructure into a global-scale sensing manifold.

Through this approach, backbone fibre networks can be transformed into massive sensors capable of detecting geological disturbances such as earthquakes and tsunamis, as well as metro-scale events like road traffic flow. While early experiments have demonstrated these capabilities as proof-of-concept tests, the employed sensing technologies often function as auxiliary, ad-hoc components. ICON aims to move beyond these static implementations by integrating sensing directly with network control and management systems, leveraging emerging virtualised functions and software-defined network (SDN) controls.

1.1. Background

Today, the operational telemetry and sensing for optical networks is addressed by four largely stand-alone families of tools – device-specific Network Management Systems (NMS), general observability and AIOps platforms, distributed and end-to-end fibre-optic sensing devices, and digital twins. Each family is mature in its own domain and individually represents the state of the art for that domain. For example, device-specific Network Management Systems such as Nokia MantaRay [2] and Huawei iMaster NCE [3] provide mature alarm management, fault notification, and large-scale device configuration for vendor-aligned carrier networks. General observability and AIOps platforms such as Grafana [4] and Splunk ITSI [5] offer flexible visualisation, query, and event-correlation capabilities on top of arbitrary time-series data, but with no domain knowledge of optical infrastructure. Distributed fibre-optic sensing platforms such as Silixa iDAS [6] and AP Sensing [7] deliver high-resolution DAS and Distributed Temperature Sensing (DTS) measurements, typically through the C-OTDR with their own front ends and storage. Digital-twin and energy-management tools such as Ansys Twin Builder [8] and Schneider EcoStruxure [9] cover their respective verticals (engineering simulation, sustainability reporting) in isolation from the network.

However, three structural limitations recur across all these families, each of them rooted in the original scope for which the platform was designed. These limitations are:

- Domain isolation. Each platform was designed before fibre infrastructure became a sensing manifold. Network telemetry, distributed sensing, digital-twin state, environmental context, and energy data therefore flow through separate systems with separate data models, and operators must manually correlate them across tools.
- Opaque automation. AIOps and intent-based platforms expose decisions to the operator but rarely the reasoning behind them. Without a transparent rationale, operator trust remains low and adoption in regulated, safety-critical environments is constrained.
- Vendor coupling. Intent and configuration interfaces in NMS suites remain proprietary, locking operators into a single-vendor stack and preventing interoperability between fibre-sensing, optical, and IP layers.

1.2. Motivation

With the advent of the IASC paradigm, the goal of the ICON project is to develop a standard method for presenting and sharing the sensing data as an integral part of the optical network operations and is driven by three primary factors:

- Environmental Sustainability and Efficiency: The telecommunication sector is responsible for 3% to 4% of global CO₂ emissions [10], with data traffic growing at an estimated 60% per year. Traditional efficiency improvements are no longer keeping pace with this growth, forcing a shift toward novel network-level optimisations, such as "low-margin" networks [11]. These

optimisations require high-precision telemetry and sensing data that current transmission systems cannot sufficiently provide.

- **Operational Context Awareness:** Fibre sensing goes beyond today's standard optical monitors by providing data on the internal and external operational context of the fibre. This allows for the real-time detection and management of unexpected perturbations, such as lightning strikes, wind gust loading, wind shear, aeolian vibration, galloping, and ice accretion on aerial cables, cable bridge crossings, or even physical attacks, which were previously managed through estimation rather than precise data.
- **Decoupling Technology from Application:** Current IASC approaches typically use static probes designed for a single application at a fixed location. ICON is motivated by the need to disaggregate the sensing architecture, decoupling sensing hardware from specific applications. This allows for a flexible, dynamic system that can adapt to specific intents, satisfying the requirements of multiple users (both telecommunication and non-telecommunication), while reducing the need for redundant, static probes.

By establishing a standard method of presenting and sharing the telemetry and sensing data, ICON enables the creation of a comprehensive picture capturing the devices, network infrastructure and its environment, facilitating real-time proactive actions to prevent damage and optimise utilisation through tools like the Availability Assistant (AA).

2. Use Cases and Applications

IASC leverages existing optical fibre infrastructure to provide both data transmission and environmental monitoring services. While fibre sensing has historically been treated as a task-specific add-on, modern approaches aim to integrate it directly into network control and management architectures to turn the global fibre plant into a large-scale sensing grid. The usage of sensing data within optical communications can be broadly categorised into internal network optimisation and external environmental awareness.

Internally, sensing data provides crucial insights into the optical fibre's operational context, allowing operators to detect physical stresses, unauthorised intrusions, or unexpected disruptions to the cable locality. This high-value telemetry data is used to improve Quality of Transmission (QoT) estimation, optimise spectral efficiency through "low-margin" network design, and proactively manage service-impacting perturbations such as lightning strikes or wind shear on aerial cables.

From the environmental aspects, fibre networks act as massive, distributed sensors capable of monitoring wide-scale phenomena that traditional sensors cannot easily reach. Key applications include the detection of geological disturbances, such as earthquakes and tsunamis, as well as urban mobility monitoring, including road traffic density and transport system conditions. Sensing data is also vital for the protection of critical infrastructure, where it is used to identify background leakages in water systems, monitor the integrity of bridges, or detect harmful maritime activities like trawling and anchor drags near subsea cables.

The ICON project advances these capabilities by merging the sensing architecture with the network management and coupling the sensing hardware to be used in dynamic applications. This transition to a dynamic and flexible system enables the retrieval of information that benefits multiple users, both telecommunication and non-telecommunication, while significantly reducing the need for redundant, application-specific sensing probes, while following the constraints of the Recommendation G.681 by the ITU-T Study Group 15 for the in-service distributed fibre-optic sensing links.

2.1. Real-time network sensing and monitoring

ICON Deliverable D2.1 [12] specified two unique use cases to be investigated in the project. The first one focuses on real-time network sensing and monitoring and aims to transform optical infrastructure from a passive infrastructure into a proactive sensing channel capable of situational awareness. This approach moves beyond traditional reactive monitoring and targets to identify early indicators of physical risk, such as construction machinery operating near buried routes or unauthorised handling of fibres in node rooms. In general, it focuses on early detection and prevention of fibre issues, such as potential harmful impacts, degradation and impending faults, through continuous, real-time monitoring.

By integrating C-OTDR/DAS and SoP monitoring, operators can detect and localise events like trawling, digging, anchor drags, and environmental loads (e.g., currents, ice, earthquakes) before they escalate into service-impacting incidents. This proactive approach is essential for minimising downtime, reducing repair costs, and safeguarding critical services across both terrestrial and subsea infrastructures.

To support proactive protection, ICON integrates high-sensitivity instrumentation, including C-OTDR/DAS for spatiotemporal awareness, SoP monitoring for end-to-end strain and lightning detection, and transceiver telemetry for end-to-end performance tracking. These technologies provide complementary insights, allowing the system to detect fast-developing threats with response windows ranging from seconds to minutes. For high-risk maritime events like trawling or anchor drags, such systems can notify operators and maritime authorities immediately and prevent vessels from reaching the cables when protective measures are undertaken. However, different applications may have different requirements for notification response times. For example, while immediate response time is required for maritime applications, minute-scale notifications are enough in icing detection or stress response detection. Furthermore, some applications focusing on long-term evaluations may not require the real-time notifications at all, since they use a different rule set for processing the sensing and telemetry data compared to real-time applications,

A core component of this real-time capability is the use of automated algorithms and machine learning (ML) to distinguish between benign environmental vibrations and actual hazardous events. By correlating fibre-sensing data with external feeds like AIS for ship tracking or weather station data, the

system reduces false alarm rates and provides actionable intelligence. This contextual awareness allows for the classification of event severity, quantifying uncertainties in location and intensity to ensure that notifications sent to Network Operation Centres (NOCs) are precise and relevant.

The ultimate operational impact of real-time monitoring is enhanced network resilience and service continuity. When the ICON intelligent platform detects a high-confidence intrusion or physical disturbance, it can instruct the SDN controller to proactively reroute high-priority or mission-critical communication channels to alternative paths before a service-impacting failure occurs. This ability to safeguard infrastructure in real-time significantly reduces downtime and possibly also infrastructure repair costs, while ensuring the stability of the mission-critical services and infrastructure that our society relies on.

2.2. Long-term network paths availability evaluation

The second use case developed in ICON Deliverable D2.1 [12] is the AA use case. This use case investigates the temporal availability of the infrastructure and routes and supports network operators in identifying stable versus risk-prone routes, planning maintenance, and reacting to emerging risks. Unlike real-time protection, this use case emphasises long-term data collection and analysis, combining sensing data, telemetry, and data lake resources with visual interfaces. The insights provided by the AA serve as a critical tool for strategic network planning and maintenance. Through the AA application, operators can access intuitive dashboards that highlight segments more prone to outages due to storms, digging activity, or flooding.

The development of the Availability Assistant relies on reliable and systematic data collection from various synchronised telemetry and sensing data sources, which are then stored in a data lake for multi-tenant access through the Telemetry/Sensing as a Service concept. The primary objective of the AA use case in the ICON project is to derive reliability estimations for the monitored infrastructure segments. This means identifying the infrastructure segments with higher variations from the sensing activity from the ones with steady baselines. By establishing steady baselines for each link, operators can distinguish between routine environmental variability and increasing patterns of instability that may indicate long-term degradation or emerging physical risks.

A core value of the AA is its multi-layered view, achieved by fusing heterogeneous data streams from various sources, including SoP monitors, C-OTDR and DAS interrogators, device telemetry, and reroutable active probes. While SoP tracking provides link-integrated insights into environmental variability, OTDR-based methods offer precise localisation of attenuation and strain evolution over time, which can be complemented with the performance evaluations from device telemetry or active probes. This sensing information can be further enhanced by correlating results with external datasets, such as AIS maritime tracking or weather stations, to identify the specific environmental or human-driven root causes behind detected variations.

The work for the Availability Assistant prototype is already ongoing and includes input from sensing and telemetry devices provided by ICON partners, installed on production fibres from Tele2 Estonia and Elering, passing from Lihula to Keila Substations (link 1) and from TalTech premises to Ülemiste substation (link 2). The primary sensing sources utilised for deriving reliability estimates for Availability Assistant on link 1 are the C-OTDR from Adtran, SoP analysers from Tampnet and Brno University of Technology, TeraFlex transceiver from Adtran, and Erbium Doped Fibre Amplifier (EDFA) from Adtran. The link 2 only utilised the C-OTDR and two SoP analysers for the task. The data collection on the links was carried out between the 27th of October 2025, until the 10th of March 2026 for link 1 and from the 6th of May until the 21st of May 2026 for link 2.

2.3. Future use cases

The current academic consensus emphasises that the viability of sensing grids depends on the algorithmic layer. A recent framework proposed by Shao et al. [13] highlights that converting conventional DAS into an "intelligent sensing grid" requires deep integration with back-end AI agents to manage complex, multi-source urban noise and enable proactive decision-making.

Furthermore, recent validation of SoP-based analysis on live, metropolitan networks by Sadighi et al. [14] demonstrates that robust sensing grids can now accurately classify threats and disturbances

directly on standard communication channels (like DP-16QAM) using ML, entirely bypassing the need for dedicated sensing hardware or dark fibre access.

Existing telecom sensing grids can be extended into several cross-disciplinary applications:

- **Seismology and Earth Observation:** Utilising C-OTDR/DAS to transform terrestrial and subsea optical networks into dense seismic arrays for high-resolution micro-earthquake detection and tectonic monitoring.
- **Smart City and Traffic Analytics:** Processing vibration signatures in terrestrial cables to passively classify heavy vehicle movement, monitor traffic density, and assess the structural health of adjacent roads or bridges without discrete sensors.
- **Multi-Utility Infrastructure Protection:** Calibrating acoustic anomaly detection to identify high-pressure leaks in co-located water and gas pipelines or detecting faults in adjacent high-voltage power cables sharing the same conduit.
- **Oceanographic Monitoring:** Leveraging subsea cables for continuous environmental observation, using DTS for deep-ocean profiling and acoustic sensing for marine mammal tracking.
- **Predictive Degradation Modelling:** Applying machine learning to microscopic precursor signatures, such as SoP transients, to identify potential localised fibre fatigue before measurable macroscopic signal attenuation.

3. Device-specific telemetry and sensing data handling

The core objective of Task 2.3 in ICON is to investigate and define optimum methods for presenting, sharing, and storing collected telemetry and sensing data to maximise its usability across a diverse set of applications. This section details the data management workflow used in the ICON project, analyses current partner methodologies, and addresses the fundamental challenges of data format heterogeneity.

3.1. Data management workflow for data presentation

Today, telemetry and sensing data from various devices are stored in local or cloud storage, operated by the data collecting entity, usually, the device owner. Access to the storage is provided as per request, and analyses are run either in real-time (using Grafana [4] or Influx DB [15]) or through post-processing. However, to accommodate multiple sensing and telemetry devices and a multi-tenant use case, the ICON project has defined a data management workflow, proposed in Deliverable D1.1 Data Management Plan [16] and illustrated in Figure 1. The data management workflow describes the functional blocks for the data throughout different phases, from collection and transmission to validation, storage, and sharing.

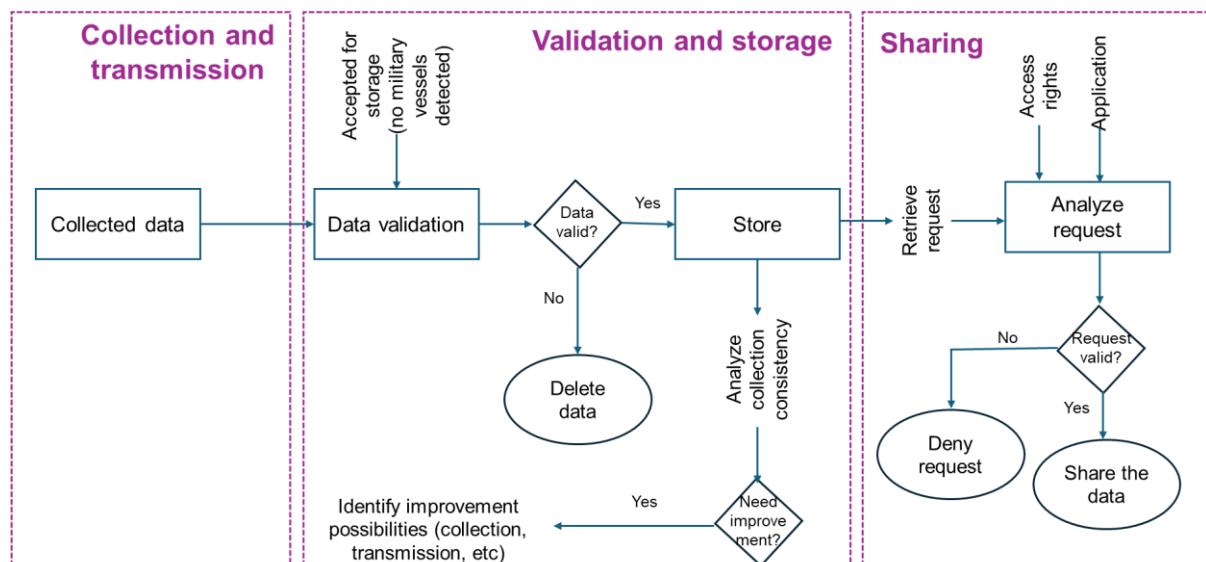


Figure 1: ICON data management workflow

The ICON project shifts away from traditional, isolated storage methods toward a unified Data Management Workflow, as defined in D1.1. This workflow establishes a structured path for data from its initial capture to its final use in diverse applications. The process is divided into three primary functional phases:

- **Collection and Transmission:** Raw telemetry and sensing data are gathered from various physical sources (e.g., C-OTDR, SoP monitors, coherent transceivers) and transmitted for centralised processing.
- **Validation and Storage:** This core phase ensures data integrity and security. Data undergoes a validation process, where it is checked against specific criteria, for example, filtering out sensitive information like military vessel movements. Validated data is archived in the Data Lake, while invalid data is deleted. The exact criteria to classify data in this stage will be developed in the later stages of the ICON project. Furthermore, the system analyses collection consistency to trigger potential improvements in the transmission or collection hardware if anomalies are detected.

- **Sharing:** The final phase manages multi-tenant access through the T/SaaS concept. When an external application or NMS issues a retrieve request, the system analyses it based on defined access rights and the specific needs of the application. Only verified and valid requests result in the sharing of data.

This workflow is designed to be flexible, allowing for future automation and adaptation to changing legal requirements and advancements in sensing technology.

3.2. Current methods of presenting/sharing the sensing data

This section presents the results of a comprehensive mapping exercise conducted among ICON project partners to identify existing methodologies for collecting, storing, and presenting the sensing information. As the primary objective of Task 2.3 is to investigate and define a standard method for presenting and sharing sensing data, it is essential to first establish a baseline of current practices across the consortium.

The information was gathered via a targeted questionnaire (see the questionnaire in Annex A) sent to partners, including Brno University of Technology (BUT), Trinity College Dublin (TCD), LightSensAI (LSAI), Adtran and Tampnet. The primary purpose of this questionnaire was to:

- *Map Heterogeneous Methodologies:* Document the diverse range of sensing and telemetry equipment currently in use, such as C-OTDR, SoP monitors, and coherent transceiver telemetry.
- *Address Format Heterogeneity:* Analyse the variations in raw data formats (e.g., .csv, .json, HDF5, TDMS) and the massive disparity in data scales, ranging from megabytes per day for telemetry to terabytes per day for phase-sensitive interrogators.
- *Define Standardised Interfaces:* Provide a foundational dataset to support the development of standardised API specifications and the transition toward the "single dataset principle" within a T/SaaS platform.
- *Identify Reliability Strategies:* Capture current partner methods for improving data interpretation, such as correlation with external datasets (AIS, weather) or the implementation of Digital Twins.

The insights collected from these responses serve as a direct input for the functional and general requirements for cloud-based storage and sharing interfaces. The following tables and partner-specific summaries provide a detailed overview of the equipment, collection methods, and presentation techniques currently employed by each participating partner.

Table 1: Current presentation method by BUT

Partner name	BUT
Equipment	SoP monitor and proprietary DAS system (OptaSense ODH-F).
Collection method	Data is stored locally at the device level or on a local server.
Primary GUI	SoP: Static .png spectrograms in .pdf reports (offline); DAS: "Waterfall" plots in the system GUI.
Dataset format	TDMS (for SoP) and HDF5 (used for both SoP and DAS).
Compression method	No compression or pre-processing is currently used; data is stored as raw binary values.
Dataset size for 24h	SoP: ~3 GB; DAS: ~18.5 GB (varies based on pulse repetition rate (frequency) and channel count).
Presentation example	Spectrograms and "waterfall" type traces.

Primary extracted findings	SoP: Offline evaluation of discrete events (anomalies, PAPR peaks); DAS: raw trace visualisation.
Secondary findings	Distinguishing environmental effects (wind, icing) from genuine disturbances through correlation with meteorological data (Open-Meteo/ERA5) or Fibre Bragg Grating (FBG) data.
Used ML models	No machine learning has been implemented yet.
Other comments	Analysis is currently exploratory; the focus is on gathering labelled reference data to develop future automatic event classification. Future plans include real-time cloud sync and InfluxDB integration.

Key Insights from the Questionnaire (BUT):

- **Data Volume Challenge:** The DAS system generates a significantly higher volume of data (18.5 GB/day) compared to SoP monitoring (3 GB/day).
- **Reliability Strategy:** BUT emphasises the use of external datasets (weather reanalysis and FBG data) to improve event interpretation and reduce false alarms.
- **Standardisation Gap:** BUT notes they have limited direct experience with the international standardisation workflows (ITU-T/IEEE) at this stage.

Table 2: Current presentation method by TCD

Partner name	TCD
Equipment	Phase-OTDR and C-OTDR.
Collection method	Locally stored (TDC-collected data) or pushed to a centralised cloud with SSH access (TalTech-collected data).
Primary GUI	Power spectral density (PSD) waterfall plots indexed by time and channel
Dataset format	HDF5 and .csv.
Compression method	No preprocessing/compression to reduce raw size; only detected regions are stored after processing.
Dataset size for 24h	DAS interrogator with 20 kHz pulse repetition rate used for TCD data collection: ~11.7 TB; TalTech-collected data varies by pulse repetition rate and channel count.
Presentation example	Frequency domain visualisation of separated events.
Primary extracted findings	Rayleigh backscattering amplitude and phase.
Secondary findings	Discrete event identification through a pipeline: PSD computation, thresholding, and bounding box creation.
Used ML models	Convolutional Neural Network (CNN) for the classification of bounding boxes converted to the frequency domain.
Other comments	Recommends secure shell (SSH) tunnels for remote multi-tenant data access to avoid forcing users to download large datasets.

Key Insights from the Questionnaire (TCD):

- **Massive Data Management:** Working with Phase-OTDR and DAS data, the volume can reach a staggering 11.7 Terabytes per day at a 20 kHz pulse repetition rate.
- **Actionable Intelligence Pipeline:** TCD's methodology includes a sophisticated processing pipeline: computing PSD, de-noising, creating bounding box regions, and utilising CNN for event classification.
- **Sharing Strategy:** To manage these massive datasets, they recommend using centralised cloud storage with SSH access to avoid the need for users to download raw TB-scale files.

Table 3: Current presentation method by LSAI

Partner name	LSAI
Equipment	Non-disruptive probes for SoP and channel monitoring.
Collection method	Private server storage accessed via APIs.
Primary GUI	GUI dashboards and the Grafana visualisation tool.
Dataset format	.json files or TimeScale DB.
Compression method	No compression algorithms are currently employed.
Dataset size for 24h	Dependent on the sampling rate.
Presentation example	Real-time data visualisation via Grafana.
Primary extracted findings	Stokes Parameters (S0, S1).
Secondary findings	Wavelength Dependent Gain and Optical Signal to Noise Ratio (OSNR).
Used ML models	No ML models are currently integrated.
Other comments	No context generation algorithms are currently employed in the existing pipeline.

Key Insights from the Questionnaire (LSAI):

- **Non-Disruptive Probing:** LSAI focuses on non-disruptive probes for SoP and channel monitoring, measuring metrics such as Stokes Parameters (S0, S1), optical signal-to-noise ratio (OSNR), and Wavelength Dependent Gain.
- **Storage and Interface:** Data is stored in .json files and TimeScale DB on a private server and accessed via APIs. Grafana is used for real-time visualisation.
- **Current Maturity:** Their current pipeline is focused on raw data visualisation; they have not yet integrated Machine Learning models or context-generation algorithms.

Table 4: Current presentation method by ADTRAN

Partner name	ADTRAN
Equipment	CC-OTDR, C-OTDR, B-OTDR, and Coherent transceiver telemetry.
Collection method	CC-OTDR: Local storage with event upload to SoFiN cloud platform; C/B-OTDR: Network-attached storage; Telemetry: Kafka pipeline (ETSI standards).
Primary GUI	Web GUI (SoFiN), Grafana, custom dashboards, and an AI chatbot (Adtran Optical Network Scientist)
Dataset format	CC-OTDR: .npz (Python array); C/B-OTDR: Proprietary; Telemetry: .json).
Compression method	None for OTDR systems; Telemetry data is compressed when published to databases.
Dataset size for 24h	C-OTDR: ~2.4 TB per 100 km fibre; Telemetry: ~15 MB (depends on polling interval).
Presentation example	Phase standard deviation and phase "waterfall" plots; Brillouin frequency shift versus fibre length.
Primary extracted findings	Rayleigh probe signal phase changes (strain/temp) and Brillouin frequency shifts (long-term strain/temp).
Secondary findings	Telemetry metrics: pre- and post-FEC BER, SoP changes, SNR, and optical power.
Used ML models	No machine learning has been implemented yet.
Other comments	Telemetry uses standard European Telecommunications Standards Institute (ETSI) interfaces ready for data lake connection. A major hurdle is the lack of common standards for sensing data exchange.

Key Insights from the Questionnaire (ADTRAN):

- **Equipment and Metrics:** Adtran utilises a variety of probes, including CC-OTDR, C-OTDR, B-OTDR, and coherent transceiver telemetry. These measure Rayleigh back-scattering phase changes for short-term strain/temperature and Brillouin frequency shifts for long-term monitoring.
- **Data Volume Challenge:** There is a massive disparity in data scales; while transceiver telemetry generates only ~15 MB per day, C-OTDR systems can produce ~2.4 TB per 100 km of fibre in a 24-hour period.
- **Visualisation and AI:** They use a web-based GUI (SoFiN), Grafana, and an innovative AI chatbot called the "Adtran Optical Network Scientist" for data visualisation.
- **Standardisation Gap:** Adtran identifies the lack of commonly agreed standards for exchanging sensing data formats as a primary hurdle.

Table 5: Current presentation method by Tampnet

Partner name	TN
Equipment	Proprietary SoP monitoring in many fibres and one DAS instrument (Optodas from ASN)
Collection method	SoP collects locally (up to 2 years) and backup in server (indefinitely). DAS collects up to 3 weeks and thus has data rotation. Additionally, selected events of a few GB are saved to a server.
Primary GUI	SoP: On-demand data analysis with time series, spectrograms, and PSDs capabilities. The downsampled streams of data are fed to an InfluxDB and shown live through Grafana dashboards in the form of time series. No live raw data processing and visualisation capabilities. DAS: HDF5 has live data analysis in the form of RMS waterfalls. On-demand data analysis allows for more analysis techniques.
Dataset format	FLAC format (for SoP), HDF5 (for DAS).
Compression method	SoP: Apart from the FLAC format compression, each SoP instrument runs a separate stream of downsampled data with a 10 Hz sampling frequency (100 ms windows) and average, RMS, and maximum data calculations. DAS: Apart from the HDF5 format compression, no further compression or pre-processing is currently used.
Dataset size for 24h	SoP raw data: around 10 GB per day. DAS raw data: around 3 TB per day.
Presentation example	Spectrograms and "waterfall" type traces.
Primary extracted findings	SoP: offline evaluation of discrete events (anomalies, PAPR peaks); DAS: raw trace visualisation.
Secondary findings	Fibre sensing instruments monitoring the subsea infrastructure are correlated with AIS data. Weather data is useful, especially for OPGW monitoring.
Used ML models	SoP has no machine learning and DSP algorithms integrated into the analysis workflow. The DAS runs a vendor-proprietary detection algorithm whose specifications are not known to the end user.
Other comments	Time synchronisation across all devices and AIS data is a challenge. Currently, connecting to synchronised NTP servers and relying on UTC time formatting is the best solution.

Key Insights from the Questionnaire (Tampnet):

- **High-Volume DAS Data Challenge:** Tampnet's DAS instrument generates approximately 3 TB/day, while the proprietary SoP monitoring produces around 10 GB/day. This creates a significant difference in the storage and data-management requirements.

- **Operational Monitoring Maturity:** Tampnet has implemented a mature monitoring workflow for SoP data, including local and server-side storage, downsampled live data streams, InfluxDB-based data access, and Grafana dashboards for live visualisation and alarm triggering, all well-suited and customised for operational needs.
- **Correlation and Synchronisation Gap:** For subsea infrastructure monitoring, Tampnet relies mainly on AIS data, with weather data also supporting interpretation. However, they point out time synchronisation across sensing devices and external AIS data as a key challenge in data correlations.

4. Requirements of the cloud architecture

To bridge the gap between device-specific collection, storage, and presentation and network-wide collection, storage, and presentation, ICON uses an architecture developed in Work Package 5. The functions of this architecture can be divided into three layers: (a) device-specific physical layer, handling the data as described in Section 3; (b) a domain-specific layer with control, processing, and data management capabilities, functions and event detection models; and (c) a cloud layer with TeraFlow SDN Orchestrator for control functions and application layer for various applications using the data. Figure 2 presents the primary blocks of the ICON intent and context aware controller architecture.

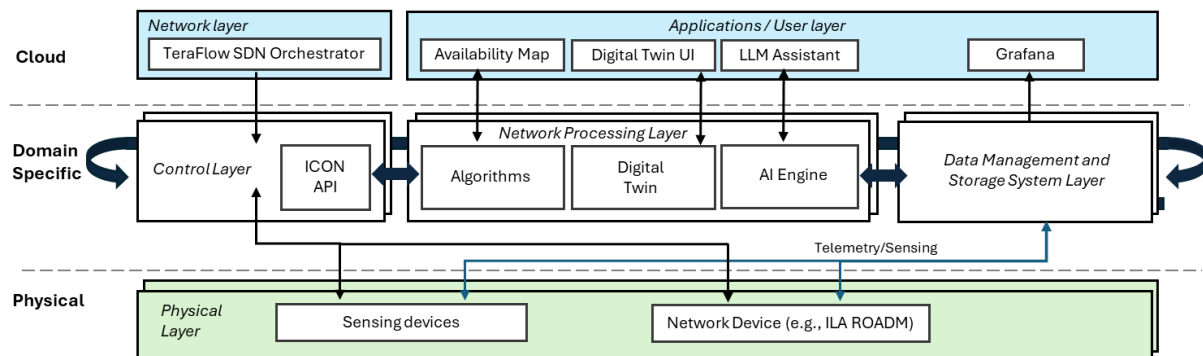


Figure 2: ICON intent-based and context-aware controller reference architecture.

The primary key in extracting value from raw data using actionable intelligence relies on the cloud platform that hosts a suite of software that enables efficient data processing for IASC systems. The suite of software includes a data management and storage layer, machine learning models, and a Digital Twin. The combination of these functionalities enables access to the following value extraction from the raw sensing and telemetry streams:

- **Correlation with External Datasets:** Aligning fibre sensing data with weather stations, maritime AIS, and cameras to validate events and reduce false alarm rates.
- **Processing of network-side sensing data:** Perform the classification, localisation, and correlation of events across the network.
- **Visualisation with smart user interface:** Develop an intuitive user interface to visualise the results without requiring in-depth knowledge of the system.
- **Intent identification and control:** Activation and routing of sensing probes to achieve the intent defined by the sensing application

The following sections discuss the core functionalities that need to be implemented to achieve the core features of the ICON cloud platform. The functional requirements are presented in Section 4, while Section 5 specifies the data flow and ingestion strategies for the cloud infrastructure to support the functional requirements. Section 6 describes the presentation approach. Detailed and in-depth User Interface (UI) specifications, acceptance criteria, non-functional requirements, and validation tests for each functional requirement are covered in the next deliverables of the project.

4.1. Functional requirements for the ICON cloud platform

Following the use cases and user requirements set out in D2.1 [12], the ICON cloud platform shall satisfy the following functional requirements (FR).

4.1.1. Monitoring

Real-Time Telemetry Dashboard (FR-MON-001): A unified operations dashboard shall present live KPIs (availability, faults, latency, fibre-sensing events) across the entire monitored network with data origin-specific refresh rates, replacing the per-tool views that current carrier NMS and observability platforms offer.

Attribute	Description
Unique ID	FR-MON-001
Type	Functional
Priority	Must
Category	Monitoring
Description	Real-Time Telemetry Dashboard
Justification	A unified operations dashboard shall present live KPIs (availability, faults, latency, fibre-sensing events) across the entire monitored network.

Alert Detection and Management (FR-MON-002): The platform shall detect, rank and route alerts from telemetry, fibre sensing, and external feeds through a single alerting pipeline with configurable severities, acknowledgement workflow, and escalation paths.

Attribute	Description
Unique ID	FR-MON-002
Type	Functional
Priority	Should
Category	Monitoring
Description	Alert Detection and Management.
Justification	The platform shall detect, rank and route alerts from telemetry, fibre sensing, and external feeds through a single alerting pipeline with configurable severities, acknowledgement workflow and escalation paths.

4.1.2. Fibre sensing

Fibre Route Visualisation (FR-SEN-001): The platform shall render fibre routes, sensor positions and active events on a GIS canvas at a minimum 1:100 000-segment scale, so that C-OTDR/DAS-class sensing data is operationally usable inside the same tool that runs network operations.

Attribute	Description
Unique ID	FR-SEN-001
Type	Functional
Priority	Should
Category	Sensing
Description	Fibre Route Visualisation
Justification	The platform shall render fibre routes, sensor positions, and active events.

Fibre Event Display (FR-SEN-002): Detected fibre events (vibrations, strain, temperature anomalies) shall be presented with their classification, confidence and supporting raw trace, accessible from the same view as the telemetry that triggered the alert.

Attribute	Description
Unique ID	FR-SEN-002
Type	Functional
Priority	Must
Category	Sensing
Description	Fibre Event Display
Justification	Detected fibre events shall be presented with their classification, confidence, and supporting raw trace, accessible from the same view as the telemetry that triggered the alert.

4.1.3. Event correlation

Event Timeline and Correlation (FR-COR-001): The platform shall correlate events across SoP, C-OTDR, DAS, transceiver telemetry, and external feeds, such as AIS, weather station data, cameras, and Supervisory Control and Data Acquisition (SCADA) on a shared time axis, surfacing the contributing observations rather than only a final verdict. The goal of the event correlation is the ability to present all data streams, regardless of their origin and refresh rate, on a single time axis view.

Attribute	Description
Unique ID	FR-COR-001
Type	Functional
Priority	Must
Category	Monitoring
Description	Event Timeline and Correlation
Justification	The platform shall correlate events across various sensing elements on a shared time axis.

4.1.4. Digital twin

Digital Twin Visualisation (FR-DT-001): The platform must present a twin of the network with different online and offline applications. It shall also provide detailed information regarding the equipment and devices in the optical network and the external sensors. Furthermore, FR-DT-001 holds a synchronised model that the end-user can force into different states, which can then be compared against reality with the Live / Simulated / Diff toggle and the per-state confidence scores included.

Attribute	Description
Unique ID	FR-DT-001
Type	Functional
Priority	Must
Category	Monitoring
Description	Device detail and Telemetry Dashboard
Justification	Situational awareness and proactive management of the optical network depend on a real-time digital twin that reflects both current device state and simulated conditions side by side.

Simulation Execution (FR-DT-002): Operators shall be able to run optimisation scenarios to optimise measurement parameters. The platform shall also provide services such as measurement simulations or predictions of aging and early maintenance recommendations. DT-002 is the engine that runs the what-if scenarios, which are presented in DT-001.

Attribute	Description
Unique ID	FR-DT-002
Type	Functional
Priority	Must
Category	Digital twin
Description	Simulation and Prediction Execution
Justification	Operators shall be able to run optimisations for measurement parameters, simulations of the network link backscattering, and predictions of the device and fibre aging, giving early maintenance recommendations

4.1.5. AI explainability

Explainability Interface (FR-AI-001): Every AI-derived output (classification, correlation, recommendation) shall be accompanied by a human-readable rationale exposing the model's name, the inputs used, and the confidence score.

Attribute	Description
Unique ID	FR-AI-001
Type	Functional
Priority	Should
Category	Monitoring
Description	Explainability Interface
Justification	Every AI-derived output (classification, correlation, recommendation) shall be accompanied by a human-readable rationale exposing the model's name, the inputs used, and the confidence score.

4.1.6. Reporting

Report Generation (FR-REP-001): The platform shall produce structured reports (.pdf, .csv, .json) from the same observation store that drives the live UI, including Environmental, Social, and Governance (ESG)-ready energy reports for regulatory submission.

Attribute	Description
Unique ID	FR-REP-001
Type	Functional
Priority	Must
Category	Reporting
Description	Report Generation
Justification	The platform shall produce structured reports (.pdf, .csv, .json) from the same observation store that drives the live UI, including ESG-ready energy reports for regulatory submission.

4.1.7. Intent-based control

Intent Submission (FR-INT-001): Based on the application's intent, activate and route sensing probes.

Attribute	Description
Unique ID	FR-INT-001
Type	Functional
Priority	Must
Category	Sensing
Description	Sensing probes identification and routing
Justification	The platform shall identify the type of probes and routes across the network to collect the raw measurements required by the sensing application.

4.2. Data format and volume

A key challenge in establishing a sensing-grid architecture is the wide variation in data formats, pulse repetition rates, sampling rates, physical metrics, and data volumes produced by different monitoring and sensing devices. The storage and processing architecture must therefore support a heterogeneous data envelope rather than a single device-specific data model. As discussed in Section 3.2, data volumes and formats vary dramatically across device types. Beyond the systems covered in that survey, sources such as Brillouin-OTDR, Raman-IOFDR, direct-detect OTDR, dedicated SoP analysers, and external reference feeds further expand the diversity in data format and physical interpretation.

The source-side data envelope reported in Deliverable D2.1 [12] includes, for example:

- Phase-sensitive OTDR / DAS: raw data up to the order of terabytes per week per link, with D2.1 §3.3.2 reporting approximately 18.5 GB/day at 10 kHz pulse repetition rate on 15 channels; native formats include HDF5 or CBT.
- Dedicated SoP analysers: approximately 3 to 7 GB/day depending on the repetition rate, parameter mapping granularity, and used storage format. For example, SoP with a sampling rate of 44.1 kHz, 16-bit per S0/S1 parameters, and FLAC format produces 7 GB of data per day, while SoP with a sampling rate of 12.0 kHz, 8-bit per parameter mapping per single channel, and HDF5 formats produces 3 GB per day.
- Transceiver-extracted SoP: approximately 5 GB/month, assuming compression / event-triggered capture at 1 kilosamples per second (kSps) for three Stokes parameters, using HDF5 or .csv formats.
- Direct-detect OTDR systems, such as Adtran ALM and equivalent solutions: on-device processed traces with low-volume .csv export.
- Brillouin-OTDR: minute-scale measurement cadence, custom data format, and a need for external network storage.
- Transceiver and optical line system streaming telemetry: low-rate, long-retention time-series data, commonly using .csv or line-protocol formats.
- External reference feeds, including AIS, weather stations, FBG sensors, cameras, load cells, and SCADA events: mixed .json, .csv, and binary formats.

In addition to data volume and format diversity, the architecture must correctly interpret the different physical metrics presented in the data. For example, C-OTDR systems provide Rayleigh backscattering traces, typically represented as amplitude and/or phase as a function of distance, while SoP analysers provide vector-based changes in polarisation state over time. These data streams must be mapped into a common temporal and spatial framework to enable event localisation, cross-correlation, and network-level interpretation.

4.3. Data interpretation

Traditional sensing workflows often rely on visual inspection of waterfall plots, seismograph-like traces, or device-specific dashboards. In an operational wide-scale sensing grid, however, end-users increasingly require automated outputs, such as event type, severity, confidence level, spatiotemporal location, affected fibre route or network element, and recommended operational response.

ICON uses the Digital Twin as an integration and interpretation layer to support this transition. The Digital Twin enables correlation between fibre sensing data, network telemetry, and external reference datasets such as weather stations, maritime AIS, cameras, and other contextual sources. This correlation is essential for validating events, reducing false alarms, improving classification reliability, and presenting sensing results in a form that is meaningful for network operations.

4.4. Dual-tier storage

The cloud-side storage layer is not intended to absorb the full raw output of every sensing device. Instead, high-volume sensing data should be processed as close to the source as possible, for example, at the interrogator, transceiver, or telemetry agent. Edge processing may include windowed event detection, digital signal processing, machine-learning-based classification, schema reduction, filtering, and noise suppression. This converts high-volume raw traces into reduced feature streams, event clips, and metadata suitable for cloud ingestion, long-term storage, and correlation.

The storage architecture must therefore support a dual-tier data model, where sensing devices should provide:

1. a low-rate, high-retention stream for telemetry, alarms, events, reports, and digital-twin functions to upload sensing events and reduced feature streams to the cloud storage. This stream is recorded in the cloud, where the data can be accessed in real time by the monitoring applications to visualise the fibre route, display events, correlate event streams, etc.
2. local storage at the device level (edge storage), where raw or semi-processed traces should remain retrievable on demand, when requested by explainability, validation, troubleshooting, or offline model improvement functions.

4.5. Data retention policy

Data retention policy for the cloud storage must ensure that the platform provides sufficient visualisation and correlation over various time periods, but does not unnecessarily overload the storage. For this, the data across hot storage that has not been accessed for a period of time will be compressed and stored in cold storage. When a future request for access arrives, the data will be restored and made available in the hot storage. The period that defines when a dataset remains inactive will be evaluated based on the needs of other components, such as the digital twin, and thus is hard to estimate at the current phase. The exact data retention policies are dependent on the stakeholders' requirements, which are not derived in this deliverable and can vary for different applications.

As an example, the following application or device-specific rules can be considered:

- Retention rules derived from the applications may follow the proactive-protection guidance in D2.1 §3.4.2 and the long-term Availability Assistant needs in D2.1 §3.4.3.
- Raw DAS/C-OTDR traces: a minimum of four-week rolling buffer per device at the edge is recommended in D2.1 §3.4.2 [12]; full-fidelity archive restricted to clips associated with flagged events, retained for five years.
- SoP time-series: 24-month hot retention in the time-series tier, indefinite compressed archive in the object store, with down-sampled summaries (per-minute and per-hour aggregates).
- Transceiver and OLS telemetry: 24-month hot retention and a minimum of five-year cold archive aligned with operator fault-analysis windows and with the Failures-in-Time / Mean Time Between Failures (MTBF) / Mean Time to Repair (MTTR) computations in D2.1 §3.4.1.
- Events, alarms, intents, and audit logs: retained for the full operational life of the asset in the relational tier; never deleted, only superseded.
- External reference data (AIS, weather, FBG, cameras, load cells): retained at the granularity needed to correlate with any retained fibre-sensing artefact; camera footage is subject to the privacy rules stated in D1.1 and shall be redacted before long-term archiving.

4.6. Time synchronisation

Because event classification relies on correlation across heterogeneous sources, every stored observation must be anchored to a common, monotonic timeline. It is recommended that sensing and

telemetry devices commissioned in the network would be receiving time from Global Navigation Satellite System (GNSS) sources and fall back to Precise Timing Protocol (PTP) or Network Timing Protocol (NTP) in case of timing failures.

The following rules apply per timestamp:

1. All timestamps are stored in UTC with at least millisecond precision; per-record time zone information is prohibited.
2. An audit log records every ingestion, schema change, export, and query that returns sensitive data.

This creates the maximum potential for the SoP and C-OTDR data to be correlated with weather-station data or AIS vessel positions.

4.7. Authentication and authorisation

For the ICON cloud to run smoothly, all users should be authenticated, and access granted as per their contractual agreements.

It is recommended that authentication is implemented with a token-based mechanism, where tokens will remain valid for multiple sessions. Session length can be specified per application or per user. On expiration of the last session and the token, the system will require a new token to be generated.

To enforce security, the Service-to-service calls use Transport Layer Security (TLS) certificates (either TLS or mutual TLS (mTLS)) for data encryption.

Rate limiting and quotas are applied per tenant and per API key. When a client exceeds its limit, requests are throttled, and the platform returns Hypertext Transfer Protocol (HTTP) 429 (Too Many Requests) with a Retry-After hint, so the surplus load is rejected at the edge rather than reaching the backend. This isolates tenants from one another: a single misbehaving client or traffic burst cannot degrade service for others.

Finally, every time a cloud platform API endpoint is invoked, the system appends a structured audit event to the audit-log. This audit-log is the primary evidence source for the validation framework.

4.8. API requirements

Finally, standardised interfaces are required between the sensing-grid data layer, the cloud platform, the Digital Twin, the NMS, and external applications. These interfaces should expose processed sensing results and relevant metadata in a format-agnostic manner, while preserving traceability to the underlying data sources. The specification must remain flexible enough to accommodate different edge-to-cloud reduction ratios, device capabilities, data formats, and deployment scenarios as ICON progresses towards operational implementation.

To ensure the maintainability of the system, all APIs need to be well defined and documented using the Open API standard. Versioning will be enforced to ensure smooth development, and it will follow the /v{N}/ prefix. All breaking changes require a new major version with a minimum 12-month overlap, and non-breaking additions should be backwards compatible.

5. Proposed data flow and ingestion strategy for cloud

To present and correlate the unique data streams from individual devices and satisfy the functional and general requirements introduced in Section 4, a method for cloud-compatible data ingestion should be developed.

In the ICON architecture, the cloud platform provides centralised data storage and exposes standardised APIs to enable bidirectional communication between the edge layer and higher-level services, including the Digital Twin, the T/SaaS platform, and other application-specific components. This architecture enables data to flow from the edge towards the cloud for storage, correlation, and analysis, while also allowing processed information, control instructions, or contextual feedback to be propagated back towards the edge where required.

The following subsections describe the data flow between architectural tiers using a bottom-up approach. In this structure, the T/SaaS concept introduced in D2.1 acts as a key enabling component for linking edge-based sensing and telemetry sources with cloud-based storage, processing, and service exposure.

5.1. Data flow

To address the requirements listed above, a new updated data flow diagram is proposed. It uses six tiers, each carrying out a specific function:

1. The edge pre-processing tier performs local digital signal processing, feature extraction, event detection, and data reduction close to the sensing source. For high-volume systems such as DAS and C-OTDR, this reduces terabyte-scale raw traces into gigabyte-scale actionable clips, compressed traces, event descriptors, and derived metrics suitable for cloud ingestion, as discussed in D2.1 §1.2.2. Local rolling-buffer policies may be defined by the device owner or infrastructure operator; however, a minimum four-week rolling buffer per device is recommended, in line with D2.1 §3.4.2.
2. The ingestion tier provides the near-real-time hot path from edge sources to the cloud platform. An Apache Kafka message bus, or equivalent streaming technology, carries each observation as a schema-validated record, with partitions keyed by `asset_id` to preserve per-link ordering and support horizontal scaling. This path shall sustain at least 10,000 telemetry events per second end-to-end, according to the FR-MON-001 acceptance criterion defined in the competitor analysis.
3. The time-series tier provides the warm path for structured numerical data. A Mimir database, or equivalent horizontally scalable time-series database, stores telemetry, polarisation parameters, strain-related quantities, and other scalar measurements with millisecond-level temporal resolution where required. It serves as the primary source for real-time dashboards, monitoring views, alerting, and short- to medium-term operational analytics.
4. The object storage tier provides the cold path for large payloads and long-retention artefacts. An object storage holds raw and compressed waveform or trace data, including HDF5, FLAC, and CBT payloads, referenced through the `raw_uri` field of the ICON observation record. This tier also stores external contextual data, including AIS, weather information, FBG sensor outputs, camera artefacts, and other reference feeds.
5. The operational database tier stores relational and transactional information required for platform operation. Implemented using PostgreSQL or an equivalent relational database system, it maintains asset catalogues, source registries, user accounts, intents, incident records, audit trails, report metadata, and other entities requiring consistency and reliable query access, which is achieved through compatibility with the atomicity, consistency, isolation, and durability (ACID) framework [17]. This framework provides the guidelines to ensure that critical data (like asset catalogues, user accounts, and audit trails) remain accurate and uncorrupted, even during hardware failures, power outages, or when multiple users access the platform simultaneously.
6. Finally, data replication, encryption at rest, and backup mechanisms are delegated to the underlying cloud infrastructure and service provider. The ICON platform, therefore, prescribes the logical tiering and inter-tier data contracts, rather than a specific implementation of these

mechanisms, ensuring portability across cloud deployments while preserving interoperability, scalability, and traceability.

Figure 3 presents the proposed data flow diagram, suitable for ingesting data from various physical-level sensing and telemetry devices. It showcases the internal signal processing capabilities of the devices and how the post-processed data is ingested in the cloud, using the standardised APIs. The T/SaaS will act as a key component in linking the edge and cloud.

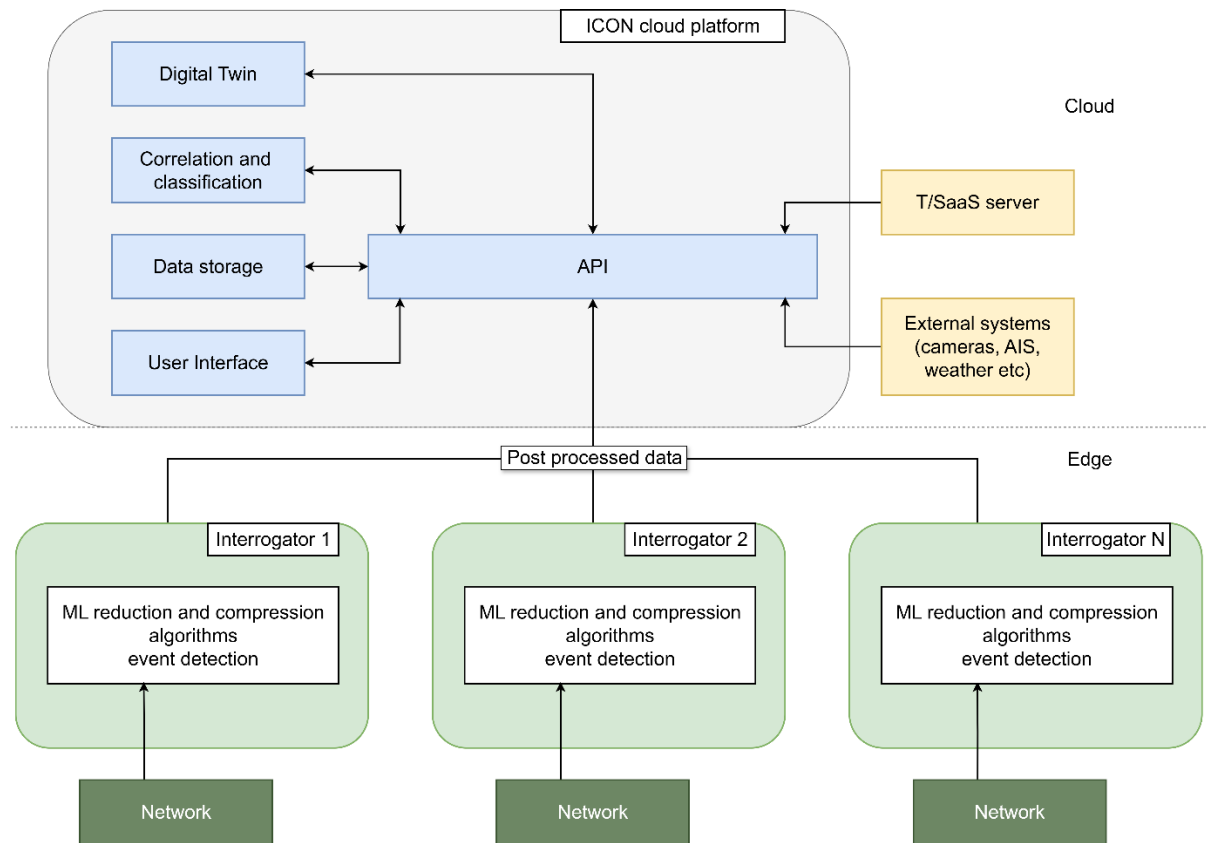


Figure 3: Data flow diagram

5.2. Ingestion strategies

The data modes will be ingested using a collection of services, which are described in greater detail below:

1. **Telemetry service:** Snippets of telemetry data will be processed and correlated to various events. The storage will be held in a hot storage for quick retrieval using a time series-based database. The data will be transmitted using a Kafka bus.
2. **Event service:** Algorithms at the edge will produce events in a pre-agreed format. The service will handle storage and alert triggering. The storage will be held in a hot storage for quick retrieval using a time series-based database. The data will be transmitted using a Kafka bus.
3. **Object store:** Raw measurements snippets are required to further the development of machine learning modules, debugging, and enrichment of the ICON dataset in general. To accommodate such data, an object storage will be implemented utilising cloud blob storage. The storage to the object store is extremely efficient for long-term archiving due to compression algorithms, but it comes at the cost of data access speeds. This solution is also known as cold storage in the context of data lakes.

The following sub-sections provide the detailed instructions for data ingestion. Data ingestion functionality will be extended and adjusted as the system matures.

5.2.1. Unified ingestion data model for telemetry data

Several heterogeneous sensors will be harmonized using the proposed structure. The collected data will maintain the logical connection between assets, which will be formally defined as the system matures. In the current version of the schema, the `asset_id` will be a collection of assets that produce data and can include various types of fibres, sensing and telemetry devices (interrogators), and external systems. All data must be transmitted using the following format:

```
{
  "timestamp": Unix time
  "asset_id": unique ID to asset
  "source_id": network unique ID of source, in case duplicate asset ids exist locally across the network
  "data family": predefined key related to the source of data, i.e. SoP, DAS, etc.
  "data": [
    ]
}
```

Since various interrogators can produce vast kinds of measurement data, the data array will be stored within a flexible field, which will be virtualised based on the data family. This approach enables the definition of various decoders, making the system scalable and extensible.

5.2.2. Unified ingestion data model for event data

Event-related information is logically connected across all collected data. The proposed model aims to transmit only the necessary information from the edge and to delegate data correlation to the cloud, where the asset relationship tables are held together with the start and end reference point IDs or coordinates for the monitored links. All data must be transmitted in the following format and include the following metadata:

```
{
  "timestamp": Unix time
  "asset_id": unique ID to asset
  "source_id": network unique ID of source, in case duplicate asset ids exist locally across the network
  "event_type": type of event triggered
  "data": [
    "event_location_start": distance in meters
    "event_location_end": distance in meters
    ...
  ]
}
```

5.2.3. Unified ingestion data model for object data

In addition to the native data from telemetry and sensing devices, supporting data from the external systems may be uploaded to the cloud platform. This data can be presented in a video format, AIS data format, or other formats. To achieve the uniform processing and presentation capabilities in the cloud, the supporting data must be transmitted in the following format and include the following metadata:

```
{
```

“timestamp”: Unix time
“asset_id”: unique ID to asset
“source_id”: network unique ID of source, in case duplicate asset IDs exist locally across the network
}

The object will be linked to the metadata and catalogued. The retrieval speed of an object will depend on the size or capacity of the transmitted data and its format.

5.3. API interfaces in the cloud infrastructure

To implement a network-wide single dataset principle in the ICON UI, standardised APIs are required. The APIs listed here are designed to be consumed by the ICON UI, by external analytics tools, and by NMS platforms operated by ICON partners. The recommendations are deliberately format-agnostic so that a new sensing technology can be onboarded by adding an entry to the source_type vocabulary, not by defining a new endpoint.

For fluent operation, the following API layering aspects should be considered:

- Ingestion APIs – write path for sensors, interrogators, and telemetry agents. Follows the unified ingestion data model (§5.2)
- Query and analytics APIs – read path for dashboards, reports, and third-party analytics.
- Subscription APIs – real-time push channel for the UI and for downstream alerting.
- Control APIs – intent, simulation, and configuration endpoints for privileged users.
- Administration APIs – catalogue management, Role-Based Access Control (RBAC), audit-log access, and tenant provisioning.

Detailed recommendations for the API families above are being defined under the technical WPs and will be documented via the OpenAPI standard.

6. Data presentation and sharing in the ICON project

This Section provides the general guidelines for the planned presentation and sharing functions in the ICON project and serves as an input to Deliverable D4.3, Cloud infrastructure and integration of signal processing in the ICON system.

In essence, the presentation and sharing layer is the place where the single-dataset principle becomes tangible to the user. Its design is driven by:

- the operational needs documented in D2.1 §2 and §3.4; and
- the absence of a unified interface that combines telemetry, fibre sensing, digital twin, intent control, and reporting

For implementation of presentation and sharing, multiple approaches can be considered, but for future-proof and scalable scenarios required in the ICON project, a multi-tenant scenario is considered. In this scenario, T/SaaS concept can be used for centralised access management. T/SaaS represents a paradigm shift where sensing-enabled telemetry or sensing device owners cooperate, and one single partner is responsible for the data collection, storage, presentation, and sharing. By decoupling sensing hardware from specific, static applications, T/SaaS provides a flexible and dynamic system that satisfies the requirements of multiple users while significantly reducing the need for redundant, application-specific probes.

Under this principle, various synchronised telemetry and sensing data sources ranging from coherent transceiver parameters to dedicated C-OTDR probes are stored in a unified data lake or a T/SaaS platform. This centralised approach enables multi-tenant access, allowing the same underlying physical data to serve a diverse range of applications simultaneously.

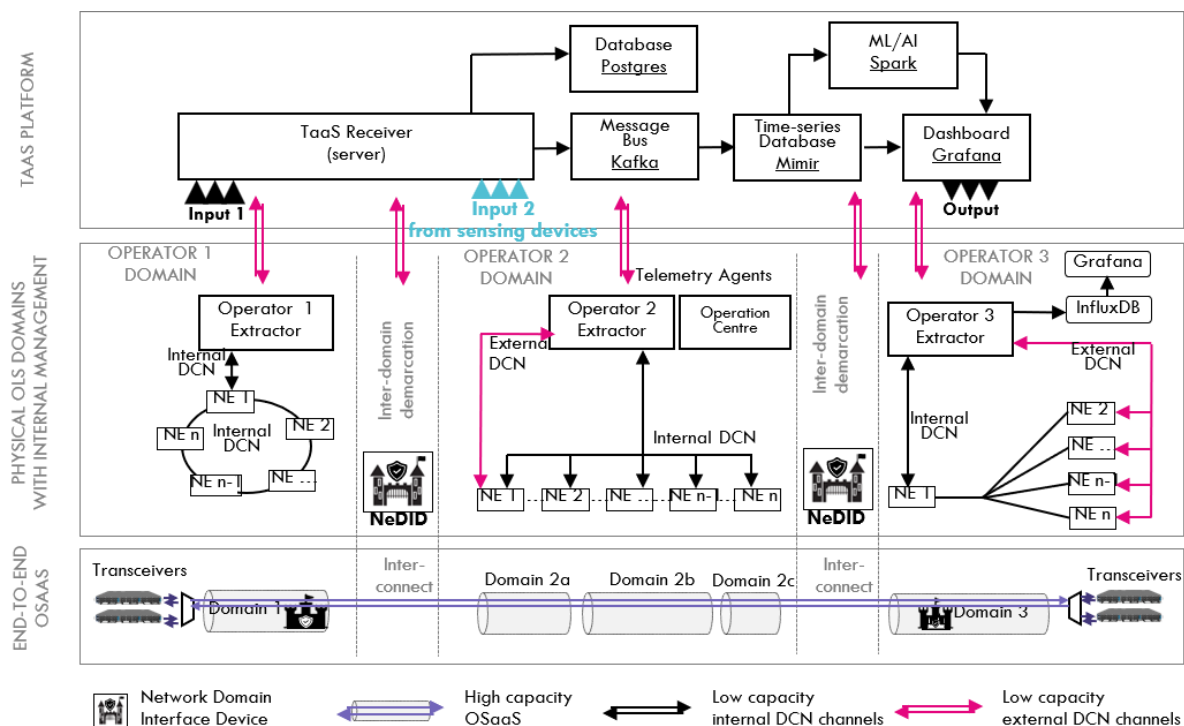


Figure 4: Telemetry/Sensing-as-a-Service (T/SaaS) architecture (reproduced from Deliverable D2.1, Figure 3).

For presentation and sharing, T/SaaS must ensure that collected information meets strict criteria regarding representativeness, timeliness, and availability, which are defined per end-user application [18]. Data representativeness is achieved by fusing heterogeneous information from various sources (DAS, SoP, OTDR, coherent telemetry, and external feeds like weather or AIS) to create a unified, context-aware operational picture of the infrastructure's environment. The implementation of a Digital Twin further enhances representativeness by modelling fibre behaviour and assessing data quality to ensure the virtual representation accurately reflects the physical system's state.

Timeliness is a critical requirement that varies based on the specific use case. Proactive protection demands high-sensitivity, real-time instrumentation with response windows ranging from seconds to minutes to enable intervention before damage occurs, such as notifying operators 10 minutes before a vessel reaches a subsea cable. In contrast, long-term analysis for network planning allows for longer response intervals (hours or days) focused on trend analysis and maintenance scheduling. Finally, data availability is maintained through highly available distributed services that ensure fast and reliable access to the platform's resources. This includes the development of monitoring systems to verify that the T/SaaS platform's collection and sharing services remain continuously operational.

6.1. General presentation principles

A goal of the ICON project is to gather the data across different sensing modalities from the network and unify the data into a single, wide-scale, network-covering large dataset to extract analytics. To achieve this, the presentation must adhere to the following principles:

Actionable over raw: Discrete, classified events and spatially localised anomalies are preferred over the raw data. Waterfall plots and high-resolution traces are available on demand for investigation.

Correlation-first: Maps, timelines, and incident views integrate external context (AIS, weather, cameras, SCADA alarms, etc.) alongside fibre-sensing results so that the user sees the same correlation that the digital twin uses internally.

Single dataset principle: Every visual presentation by the ICON UI uses the same observation stream. There are no simultaneous secondary data sources for the real-time presentation mode.

Role-aware: Data access, dashboard views and enabled actions depend on the Role-Based Access Control. The roles can be Administrator, Operator, or Researcher.

Responsive: The ICON UI must support screen sizes larger than small laptops (1366 × 768) and align with the W3C Web Content Accessibility Guidelines [19].

Real-time without page reload: Telemetry, alerts, and simulation deltas are pushed to the browser without a page reload using a maximum latency defined by the end-user application. This latency, or data refresh interval, is usually defined between 1 second and 1 minute.

To ensure a smooth user experience, the UI organises data in multiple modules. Visualisations of the telemetry and event data will be combined to provide a clear view of incident reporting, both in real-time and historical views. Additionally, users will be able to utilise the digital twin to investigate simulations offered. Specifically, the modules that ICON UI includes are Operations Dashboard, Telemetry Explorer, Incident Management, Digital Twin Interface, and Context & Sensing Map. The following paragraphs provide a short overview of each of these modules. The full overview will be provided in Deliverable D4.3.

6.2. Application-specific presentation

To account for the variations in the requirements by different applications, application-specific presentation rules may be applied for the network-wide dataset collected and provided by the ICON cloud. For example, the Availability Assistant application is the primary consumer of the long-term sensing dataset (D2.1 §3.4.3). It is rendered as a layered GIS view with available infrastructure-specific reliability metrics such as failure rate, MTBF, MTTR, and OSNR/Generalised OSNR (GOSNR) margin per segment. Each segment is shaded according to a reliability score computed from the underlying SoP, C-OTDR, B-OTDR, DAS, and transceiver telemetry, independent of which technology is installed on the segment. This preserves the uniform presentation regardless of the underlying sensing or telemetry device. Clicking a segment opens a detail panel with the contributing sources, recent events, protection-path suggestions, and a link to a segment-filtered telemetry view.

6.3. Reporting and data export

The developed ICON cloud platform will provide a unified reporting hub. Reports are generated by aggregating the various data sources to ensure it provides the user with a collection of events across the network and the correlated data related to the events. Every report is a saved, parameterised query

over the stored data (§5.2) paired with a presentation template. The report catalogue is therefore open-ended rather than a fixed menu: a new report is defined by a new query and template. We foresee the following reports to be supported by the first version of the platform.

Report class	What it aggregates
Network-health/availability	Failure rate, MTBF, MTTR, OSNR/GOSNR margin per segment, reliability score
Event/incident	Correlated event timelines with contributing observations, model name, as well as event_confidence and actions taken if available
Fibre-sensing	Flagged route events, event clips, and distance-resolved anomalies
Historical analytics	Down-sampled per-minute / per-hour summaries over long windows
Audit/compliance	Structured audit events from every API call (who/what/when/scope)

Three export formats are planned to be produced by the cloud platform, each serving a distinct consumer:

- **.pdf** – human consumption; the rendered report (figures, tables, and narrative), suitable for dissemination and regulatory submission (e.g. event reports).
- **.csv** – spreadsheet and statistical analysis; flat tabular rows, one observation or aggregate per row, with columns drawn from the observation schema (timestamp_utc, asset_id, source_type, measurement_family, measurement_name, value, unit).
- **.json** – machine consumption.

7. Related standards and recommendations

To enable interoperability between various devices, the ingestion of a variety of data formats, flexible storage, and real-time presentation and sharing, all components of the ICON sensing system should follow the recommendations and standardisation as in-depth as possible. Distributed-sensing endpoints should inherit the wavelength and interference definitions of ITU-T G.681; DAS-specific payload fields and parameters should follow the terminology and definitions of IEEE 3101-2023. For submarine-cable sensing, the observation-export format should be compatible with ITU-T G.979 and G.9730.1/G.9730.2 so that scientific sensing data can be shared without re-encoding.

The standards overview presented in this section follows a targeted, top-down methodology. Rather than surveying the full landscape of optical-networking standards, the analysis focuses on the standards-developing organisations (SDOs), whose active work programmes directly address the presentation, sharing, interfacing, and terminology of fibre-optic sensing and telemetry data. Two complementary standardisation bodies were selected as the primary focus: ITU-T Study Group 15 (SG15), which governs transport-network and optical-systems Recommendations, and the IEEE (notably the Photonics Standards Committee and the IEEE Standards Association, IEEE-SA), which leads on sensor terminology, data formats, and physical-layer specifications.

Within each subchapter, individual standards and recommendations were selected according to three criteria: (i) direct relevance to distributed fibre-optic sensing (DFOS) and IASC over communication fibre; (ii) coverage of the data-presentation, data-sharing and interface aspects central to this deliverable; and (iii) maturity and trajectory, so that approved Recommendations, technical reports and ongoing draft projects could all be considered as part of an evolving standardisation roadmap. For ITU-T SG15, the review centres on Question 6 (Q6/15, characteristics of optical systems for terrestrial transport networks) and its flagship Recommendation G.681, complemented by related activity in the submarine (Q8/15) and access-network (Q2/15) questions and the strategic ION-2030 framework. For IEEE, the review covers the published interrogator-terminology standards IEEE 3101 and IEEE 2067, the gyro specification standard IEEE 952, and the IEEE-SA white paper on fibre-optic sensors.

The following subchapters examine each of these standards and recommendations, summarising their scope and status and assessing their alignment with the ICON data-presentation and sharing approach.

7.1. ITU-T Study Group 15 (SG15)

ITU-T SG15 is the lead ITU-T study group for transport networks, optical systems, and access infrastructure, and is the principal venue in which fibre-optic sensing is being standardised for telecommunication use. Within SG15, Question 6 (Q6/15) owns the systems-and-interfaces work most directly relevant to ICON, while related sensing activities are distributed across the submarine (Q8/15) and access (Q2/15) questions.

7.1.1. Recommendation G.681 (formerly G.dfos)

The most significant Q6/15 output in this field is Recommendation ITU-T G.681, “Distributed fibre optic sensing system for terrestrial optical transmission systems”, approved on 29 November 2025 [20], [21]. It is the first telecommunication standard for distributed fibre-optic sensing (DFOS), in which the communication fibre itself serves as the sensor. The Recommendation specifies the optical interface requirements for DFOS in in-service terrestrial transmission systems, including wavelength allocation for the sensing signal across the S-, C-, L- and U-bands, allowable crosstalk and interference limits between the sensing and data-communication channels, and both out-of-band (OOB) and in-band sensing scenarios. It defines reference interface points (e.g., DFOS-Sb1 and DFOS-Rb1) and parameters such as peak on-burst power, central frequency and loss budget for ITU-T

G.652/G.654/G.657 fibre. These provisions make G.681 directly relevant to ICON's IASC approach, in which SoP- and OTDR-type measurements must coexist with live DWDM traffic.

7.1.2. Technical Report TR.fsc — Feedforward sensing

A complementary SG15 technical report, TR.fsc ("Feedforward sensing for optical communication networks"), addresses feedforward sensing technology, its use cases and technology requirements, and the path toward future standardisation. It indicates the direction of post-G.681 work and is therefore useful for anticipating future interface requirements in the ICON platform.

7.1.3. ION-2030 framework and the ION-aiDC report

SG15 has established the broader ION-2030 framework ("International Optical Networks towards 2030 and beyond"), which sets a strategic vision for integrated sensing and communication (ISAC) in optical networks. A related technical report on datacentre-centric optical networks and the role of artificial intelligence in their management — referenced in SG15 documentation as TR.ION-aiDC / GSTR.ION-aiDC — extends this vision to AI-optimised data centres and the integration of sensing data. ICON's data-presentation and sharing model is consistent with this trajectory toward AI-assisted, sensing-aware optical layers.

7.1.4. Related activities in other SG15 questions

While Q6/15 handles terrestrial systems and interfaces, sensing topics are also pursued elsewhere in SG15, with which Q6/15 collaborates:

- **Q8/15 (submarine systems):** Recommendation G.979 (characteristics of monitoring systems for optical fibre submarine cable systems) [22] and the SMART-cable Recommendations G.9730.1 (dedicated scientific sensing submarine cable system) [23] and G.9730.2 (scientific monitoring and reliable telecommunication submarine cable systems) [24] specify sensing functions and architectures that enable submarine cables to carry scientific sensing data.
- **Q2/15 (access networks):** the supplement G.suppl.VHSP (published as ITU-T G Suppl. 88, 2025), on PON transmission technologies above 50 Gbit/s per wavelength, includes sensing considerations for very-high-speed PON (VHSP) systems.

7.2. IEEE photonics and sensor standards

To complement the ITU-T network-level Recommendations, several IEEE photonics and optical-sensor standards provide critical frameworks for terminology, data formats, and physical-layer specifications. These are developed primarily by the IEEE Photonics Standards Committee under the IEEE Standards Association (IEEE-SA).

7.2.1. IEEE 3101-2023 — DAS interrogator terminology

IEEE Std 3101-2023 [25] defines terminology and key performance parameters for fibre-optic C-OTDR/DAS interrogators, so that DAS systems from different suppliers can be described fully and compared on a common basis. It is a published, active standard. Such harmonised definitions are a prerequisite for sharing actionable DAS results across the heterogeneous platforms envisaged in ICON.

7.2.2. IEEE 2067-2021 — FBG interrogator terminology

IEEE Std 2067-2021 [26] provides terminology and definitions for Fibre Bragg Grating (FBG) interrogators, with the same objective of eliminating specification ambiguity and enabling cross-supplier comparison across application domains. It is likewise a published, active standard.

7.2.3. IEEE 952 — interferometric fibre optic gyros

IEEE Std 952 — current edition IEEE Std 952-2020, superseding IEEE Std 952-1997 [27] — defines specification formats and test procedures for single-axis interferometric fibre-optic gyros. Although its application domain (inertial sensing) lies outside ICON's scope, it provides a historical baseline and a methodological template for specifying and testing fibre-based sensing components.

7.2.4. IEEE-SA white paper on fibre optic sensors

The IEEE-SA white paper on fibre optic sensors [28] offers strategic guidance for standardising sensing data in a multi-provider environment. Three points are especially relevant to ICON:

- **Alignment with the ITU grid:** standardising Bragg-grating wavelengths on the ITU grid makes grating manufacture and deployment compatible with existing telecommunication applications, directly supporting ICON's IASC approach.
- **Uniform functionality terminology:** an internationally agreed vocabulary for resolution, accuracy, repeatability, and stability is needed so that data from different devices (e.g., C-OTDR versus SoP) can be compared reliably — the basis of ICON's "single-dataset" principle.
- **Gap analysis for industry adoption:** the paper notes that limited standardisation has historically constrained the commercial use of fibre sensors, recommends that IEEE focus on gaps such as data sharing and cost, and refers to the IEC 61757 family (e.g., IEC 61757-1 generic specification and IEC 61757-3 distributed temperature sensing [29]) for the technical characterisation of strain and temperature measurement.

7.2.5. IEEE within the ION-2030 framework

IEEE is identified as a primary collaborator within the ION-2030 framework, which envisions integrated sensing, computing, and AI agents within the optical layer to enable real-time awareness and automation. The synergy between IEEE's component- and terminology-level standards and ITU-T's network-level Recommendations supports the transition of optical networks from passive "connection channels" to active "perception hubs" — the central idea of the ICON architecture.

Conclusions

This deliverable has investigated how the telemetry and sensing data collected within the ICON system should be presented, shared, and stored so that a single, network-wide dataset can serve both telecommunication and non-telecommunication applications. Starting from the use cases and sensing specifications of Deliverable D2.1, it has mapped current partner practices, exposed the heterogeneity and fragmentation of today's tooling, and translated these findings into a coherent set of requirements and a concrete dataflow, ingestion, and presentation architecture for the ICON cloud platform.

The principal outcome of this deliverable is a unified, format-agnostic framework built on the single-dataset principle and the Telemetry/Sensing-as-a-Service concept, supported by seven functional requirements, a dual-tier edge–cloud storage model, a six-tier data-flow with unified ingestion schemas, and a set of presentation principles that prioritise actionable, correlated, and explainable information over raw traces. The accompanying standards analysis – centred on ITU-T G.681 and the relevant IEEE photonics and sensor standards – establishes a clear compliance baseline and confirms that the proposed data model is aligned with the emerging international roadmap for Integrated communication and sensing.

The deliverable has also identified that the functional requirements must be expanded into detailed user-interface specifications, acceptance criteria, and validation tests; the unified ingestion schemas and API contracts need to be finalised and implemented; and stakeholder-driven data-retention policies, role-based access rules, and edge-to-cloud reduction ratios must be quantified against real measurement campaigns, such as the C-OTDR and SoP data collection from two links monitored in Estonia in Elering and Tele2 Estonia's production network. Furthermore, a continued engagement with ITU-T SG15 and IEEE, together with validation under diverse real-world conditions, will be essential to consolidate these recommendations and to support the transition of optical networks from passive connection channels to active perception hubs.

References

- [1] Intent and Context aware Optical Networks, homepage, online, [Last Accessed: June 2026], <https://iconproject.eu/>
- [2] Nokia MantaRay, homepage, online, [Last Accessed: June 16, 2026], <https://www.nokia.com/mobile-networks/ran-operations/network-management/>
- [3] Huawei iMaster NCE, homepage, online, [Last Accessed: June 2026], <https://e.huawei.com/en/products/network-analysis>
- [4] Grafana, homepage, online, [Last Accessed: June 16, 2026], <https://grafana.com>
- [5] Splunk ITSI, homepage, online, [Last Accessed: June 16, 2026], <https://splunkbase.splunk.com/app/1841>
- [6] Silixa iDAS, homepage, online, [Last Accessed: June 16, 2026], <https://silixa.com/technology/idas-intelligent-distributed-acoustic-sensor/>
- [7] AP Sensing, homepage, online, [Last Accessed: June 16, 2026], <https://www.apsensing.com/en/technology-and-products/distributed-acoustic-sensing>
- [8] Ansys Twin Builder, homepage, online, [Last Accessed: June 16, 2026], <https://www.ansys.com/products/digital-twin/ansys-twin-builder>
- [9] Schneider EcoStruxure, homepage, online, [Last Accessed: June 16, 2026], <https://www.se.com/ww/en/work/campaign/innovation/platform/>
- [10] Boston Consulting Group, 2021. Putting sustainability at the top of the telco agenda. Available at: <https://www.bcg.com/publications/2021/building-sustainable-telecommunications-companies> [Accessed: June 16, 2026].
- [11] Yvan Pointurier, "Design of Low-Margin Optical Networks," Journal of Optical Communications and Networking, 9, A9-A17 (2017)
- [12] ICON Deliverable D2.1 Final report on use cases and physical sensing solution specifications, to be published on project homepage [12]
- [13] L.T. Shao, J.Y. Sun, J.M. Zhang, et al „Distributed acoustic sensing system integrable with AI agents“, Opto-Electron Plus 1, 250012 (2025).
- [14] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, "Deep Learning for Detection of Harmful Events in Real-World, Noisy Optical Fibre Deployments," J. Lightwave Technol. 43, 6092-6101 (2025)
- [15] InfluxData, „InfluxDB Open Source Time-Series Database,” 2025. online, [Last Accessed: January 2026], <https://www.influxdata.com>
- [16] ICON Deliverable D1.1 Data Management Plan, to be published on project homepage
- [17] T. Haerder and A. Reuter, 1983, “Principles of transaction-oriented database recovery”, ACM computing surveys (CSUR), 15(4), pp.287-317.
- [18] K. Kaeval, T. Järvelill, J. Müller, and M. Tikas, "Exploring Telemetry Collection Interval and Continuity in a Six-Month Study of a Pan-European Network," Optical Fiber Communications Conference (OFC), 2025, pp. 1-3.
- [19] Web Accessibility Initiative, Standards and Guidelines, online, [Last Accessed: June 26, 2026], <https://www.w3.org/WAI/standards-guidelines/wcag/>

- [20] ITU-T Recommendation G.681 (11/2025), “Distributed fibre optic sensing system for terrestrial optical transmission systems,” International Telecommunication Union, Geneva, approved 29 Nov. 2025.
- [21] Y. Wakisaka, G. Di Rosa, Q. Fan, E. Rochat, A. Sandmann, X. Liu and J. S. Wey, “ITU-T G.681: the first distributed fibre optic sensing standard for in-service terrestrial optical networks,” *Journal of Optical Communications and Networking*, vol. 18, no. 4, pp. B1–B8, 2026.
- [22] ITU-T Recommendation G.979 (11/2016), “Characteristics of monitoring systems for optical fibre submarine cable systems,” ITU, Geneva.
- [23] ITU-T Recommendation G.9730.1 (08/2024), “Dedicated scientific sensing submarine cable system,” ITU, Geneva.
- [24] ITU-T Recommendation G.9730.2 (08/2024), “Scientific monitoring and reliable telecommunication submarine cable systems,” ITU, Geneva.
- [25] IEEE Std 3101-2023, “IEEE Standard for Fibre Optic Distributed Acoustic Sensing (DAS) Interrogator — Terminology and Definitions,” IEEE, 2023.
- [26] IEEE Std 2067-2021, “IEEE Standard for Fibre Optic Sensors — Fibre Bragg Grating Interrogator — Terminology and Definitions,” IEEE, 2021.
- [27] IEEE Std 952-2020, “IEEE Standard for Specifying and Testing Single-Axis Interferometric Fibre Optic Gyros,” IEEE, 2020 (superseding IEEE Std 952-1997).
- [28] IEEE Standards Association, “Fibre Optic Sensors” (white paper, IC15-001), IEEE-SA, 2017.
- [29] IEC 61757-1:2012, “Fibre optic sensors — Part 1: Generic specification,” and IEC 61757-3:2015, “Part 3: Distributed temperature sensing,” International Electrotechnical Commission, Geneva.

Annex A:

Partner Questionnaire for Data Presentation and Sharing

Questionnaire for ICON Partners: Sensing Data Presentation and Sharing

D2.2 partner name:

Objective: To map current partner methodologies for data collection, storage, and presentation to define a standard ICON data format.

Expected Outcome: The responses to this questionnaire will be used to populate Table x: Current presentation method per partner in D2.2 part 3.2 Current methods of presenting/sharing the sensing data, and provide the basis for the proposed standardised API specifications for the ICON intelligent platform.

1. Sensing & Telemetry Equipment

1.1. Which specific sensing or telemetry hardware are you utilising for the ICON project? (e.g., C-OTDR, CA-OTDR, SoP monitors, Coherent Transceiver telemetry, non-disruptive probes).

.....

1.2. What is the primary physical metric being measured? (e.g., Rayleigh backscattering amplitude/phase, Stokes parameters, attenuation traces, OSNR/GOSNR).

.....

2. Data Collection & Storage

2.1. Dataset Format: In what format is your raw data currently stored? (e.g., .csv, .json, HDF5, FLAC, CBT).

.....

2.2. Data Volume: What is the approximate size of the dataset generated by your equipment over a 24-hour period?

.....

2.3. Collection Method: Is data stored locally at the device level, on a private server, or automatically pushed to a centralised cloud/Data Lake?

.....

2.4. Compression: Do you currently use any pre-processing or compression algorithms (e.g., DSP/ML-based volume reduction) to reduce the data footprint?

.....

3. Data Presentation & GUI

3.1. Current Interface: How is the processed data currently visualised for users? (e.g., GUI-based dashboards, "waterfall" plots, seismograph-like traces, or infrastructure-specific Availability Assistants).

.....

3.2. Actionable Insights: Does your current system provide discrete event identification (e.g., localised alarms for digging/trawling) or only raw trace visualisation?

.....

4. Reliability & Interpretation

4.1. Reliability Means: What methods do you use to ensure the accuracy of data classification and reduce false alarms? (e.g., correlation with external AIS/weather data, digital twin modelling).

.....

4.2. ML Models: Which specific Machine Learning models or signal processing algorithms are currently integrated into your analysis workflow?

.....

5. Future Functionality & Standardisation

5.1. Future Needs: What additional functionalities or data sharing capabilities (e.g., specific API specifications) are required for your sensing data to be usable by multi-tenant applications?

.....

5.2. Standardisation Gaps: Based on your experience, what are the biggest hurdles in synchronising your data format with international standards (e.g., ITU-T SG15 or IEEE)?

.....

6. Additional Comments

Please provide any additional insights regarding future use cases (e.g., formulation of sensing grids) or technical constraints not covered above.

.....