6G-MIRAI-HARMONY - Definition of initial common scenarios, data management and benchmarking methodology

Project information

Project name	Machine Intelligence based Radio Access Infrastructure
Project acronym	6G-MIRAI-HARMONY
Grant agreement	101192369
Call	HORIZON-JU-SNS-2024
Topic	HORIZON-JU-SNS-2024-STREAM-B-01-05
Type of action	HORIZON JU Research and Innovation Actions
Start date	1 April 2025
Duration	36 months

Document information

Associated WP	WP3
Associated Task	T3.1, T3.2
Associated Deliverable	D3.1
Main authors	Efstathios Katranaras (Editor) (**SEQ**), Robbert Beerten (**KUL**), Md Munjure Mowla (**ISRD**), Rouaa Diab (**Fraunhofer**), Jochen Fink (**Fraunhofer**), Renato L. G. Cavalcante (**Fraunhofer**), Utku Uçak (**Fraunhofer**), Paul Almasan (**Telefonica**), Henning Sanneck (**Apple**), Irched Chafaa (**UNIPI**), Alberto A. D'Amico (**CNIT**)
Contributors	Rodney Martinez Alonso (**KUL**), Cel Thys (**KUL**), Noor Ul Ain (**Fraunhofer**), Giacomo Bacci (**UNIPI**), Luca Sanguinetti (**CNIT**)
Reviewers	Paul Almasan (**Telefonica**), Tobias Ley (**Ericsson**), Henning Sanneck (**Apple**), Robbert Beerten (**KUL**)
Type	R
Dissemination level	PU
Comment	This deliverable provides the first consolidated framework for the definition of benchmarking and validation methodologies, data management principles, and common scenarios within the 6G-MIRAI-HARMONY project.

Document revision history

Version	Date	Changes	Author
V0.1	2025-10-27	1st version of deliverable	Efstathios Katranaras
V0.2	2026-01-09	Version with content, under editing	Efstathios Katranaras
V0.3	2026-01-19	Version for review	Efstathios Katranaras
V0.4	2026-01-29	Revised clean version	Efstathios Katranaras
V1.0	2026-01-30	Final version of deliverable for submission to EC	Efstathios Katranaras

Disclaimer

Co-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 or the SNS JU (granting authority). Neither the European Union nor the granting authority can be held responsible for them. 6G-MIRAI project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation program under Grant Agreement No 101192369.

Contents

Acronyms and definitions

3GPP	3rd Generation Partnership Project
5G	Fifth generation (of telecommunication networks)
5G-PPP	5G Infrastructure Public Private Partnership
6G	Sixth Generation (of telecommunication networks)
6G-IA	6G Infrastructure Association
AI	Artificial Intelligence
AI-AI	AI-enabled Air Interface
AP	Access Point
CAPEX	Capital Expenditures
BS	Base Station
CFO	Carrier Frequency Offset
CSI	Channel State Information
DL	Downlink
EU	European Union
FDD	Frequency Division Duplex
FLOPS	Floating Point Operations per Second
gNB	Next generation NodeB (Base Station)
HW	Hardware
IEEE	Institute of Electrical and Electronics Engineers
ITU	International Telecommunication Union
JP	Japan
KPI	Key Performance Indicator
KVI	Key Value Indicator

(N) LoS	(Non) Line of Sight
LTE	Long Term Evolution (4G)
MAC	Medium Access Control
(Ma/D) MIMO	(Massive/Distributed) Multi-Input Multiple-Output
ML	Machine Learning
(N/R) MSE	(Normalized/Root) Mean Squared Error
NR	New Radio
OFDM	Orthogonal Frequency Division Multiplexing
OPEX	Operational Expenditures
PHY	Physical Layer
PoC	Proof of Concept
QoS	Quality of Service
RAN	Radio Access Network
RF	Radio Frequency
RRM	Radio Resource Management
RX	Receiver
SDG	Sustainable Development Goal
SDO	Standards Developing Organization
SFO	Sampling Frequency Offset
SGCS	Squared Generalized Cosine Similarity
SME	Small and Medium Enterprise
SNS JU	Smart Networks and Services Joint Undertaking
SW	Software
TDD	Time Division Duplex
THz	Tera Hertz

TLD	Top-Level Domain
TSN	Time-Sensitive Networking (IEEE)
TX	Transmitter
UE	User Equipment (Terminal)
UL	Uplink
URLLC	Ultra Reliable Low Latency Communications
WG	Work Group
WP	Work Package

Table of partners

Short Name	Partner
Ericsson France	ERICSSON FRANCE
Fraunhofer	FRAUNHOFER GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG EV
Telefonica	TELEFONICA INNOVACION DIGITAL SL
ISRD	ISRD SP Z O.O. (IS-WIRELESS)
SEQ	SEQUANS COMMUNICATIONS SA
APPLE	APPLE TECHNOLOGY ENGINEERING BV & CO KG
KUL	KATHOLIEKE UNIVERSITEIT LEUVEN
CNIT	CONSORZIO NAZIONALE INTERUNIVERSITARIO PER LE TELECOMUNICAZIONI
UNIPI	UNIVERSITA DI PISA

Executive summary

This deliverable provides the first consolidated framework for the definition of benchmarking and validation methodologies, data management principles, and common scenarios within the 6G-MIRAI-HARMONY project. As 6G is envisioned to be AI-native, with AI/ML tightly embedded into the radio access network and physical layer, the project addresses the growing need for coherent, transparent, and reusable evaluation practices that go beyond traditional model-based performance analysis.

The report establishes a shared foundation for assessing the technological enablers developed in WP1 and WP2 by clarifying the roles of benchmarking and validation, defining a structured KPI/KVI framework, and aligning these metrics with state-of-the-art activities in standardization bodies and related European research initiatives. Particular emphasis is placed on repeatability, traceability, and comparability of results, acknowledging the challenges introduced by data-driven methodologies, heterogeneous datasets, and diverse training and evaluation pipelines.

To support robust AI/ML evaluation, the deliverable introduces a data lifecycle and management plan that promotes high-quality, well-documented, and reusable datasets, especially to allow reproducibility of evaluations. This approach aims to strengthen the European wireless research ecosystem and enable fair comparison of AI-native solutions across partners, scenarios, and use cases. In addition, the report presents an initial catalogue of 6G-MIRAI-HARMONY technical topics and enablers, with a strong focus on distributed and cell-free MIMO deployments, while making an attempt to organize them into common research areas and scenarios as well as to identify opportunities for potential synergies among partners.

Overall, D3.1 lays the methodological and organizational groundwork for subsequent technical activities in WP1 and WP2 as well as validation activities and system-level demonstrations in WP3, while remaining open to refinement as the project progresses and new insights emerge.

1 Introduction

The 6G-MIRAI (Machine Intelligence-based Radio Access Infrastructure)-HARMONY project addresses the gap between the envisioned potential of AI/ML[1] in 6G and the practical constraints of real wireless systems. By aligning the European academic and industrial partners in 6G-MIRAI with Japan's HARMONY consortium, the project aims to develop resilient and future-ready 6G physical-layer (PHY) and radio access network (RAN) technologies, with a particular focus on user-centric cell-free massive MIMO and the application of robust, reliable AI/ML techniques.

A core objective of 6G-MIRAI-HARMONY project is: "*Unified Benchmarking Platform (O3) – Build a shared EU-JP framework for data management, testing, benchmarking, and validation of AI-enabled techniques and protocols*". Work package 3 (WP3) of the project is tasked with working towards and reaching this objective. In this deliverable report, we establish an initial, coherent framework for benchmarking and validation of AI-native 6G technological enablers in 6G-MIRAI-HARMONY by defining methodologies, data management principles, and common representative scenarios to ensure fair, reproducible, and aligned evaluation of solutions.

In **Chapter 2**, 6G-MIRAI-HARMONY targets first to setup an initial framework for benchmarking (i.e., comparative assessment, using common scenarios, key performance and value indicators, baselines, and methodologies to rank or position solutions relative to reference methods or competing approaches) and validation (i.e., checking correctness and feasibility against predefined requirements or assumptions) of the technological enablers developed in WP1 and WP2.

In contrast to classical model-based signal and data processing, AI/ML-based 6G research faces substantial challenges in coherence and comparability due to variability in datasets, training procedures, implementations, and execution environments. This heterogeneity makes fair benchmarking and reproducibility difficult, underscoring the need for common data

[1] Throughout this deliverable, **"AI"** refers to system-level intelligence concepts (i.e., referring to the system-level capability of a network or entity to exhibit intelligent behaviour such as perception, reasoning, decision-making, and autonomous adaptation), **"ML"** denotes data-driven learning algorithms (i.e., a subset of AI that focuses on data-driven algorithms which learn patterns or representations from data in the form of ML models to perform specific tasks), and **"AI/ML"** is used when both perspectives are jointly addressed (i.e., an umbrella term used when referring jointly to AI-based system concepts and ML-based realisation techniques, especially when their distinction is not essential).

formats, transparent experimental protocols, and open access to code and datasets. Recognizing that progress in AI-native 6G depends on large, realistic, and reusable datasets, **Chapter 3** outlines the objectives, operation, and current assets of a structured data management and lifecycle plan. Setting-up a well-structured data collection and utilization process will help 6G-MIRAI-HARMONY to produce high-quality datasets to strengthen the European wireless ecosystem.

While AI/ML was introduced in 5G-Advanced primarily to enhance network operations, 6G is envisioned as AI-native, with AI/ML serving as a core component rather than an add-on. Beyond the new use-cases and applications anticipated for 6G, particularly those related to distributed techniques, the integration of AI/ML necessitates not only the evaluation of newly developed methods across diverse setups and scenarios, but also the application of AI/ML to test, refine, and enhance existing techniques under multiple network configurations aligned with 3GPP standard models. To this end, **Chapter 4** presents 6G-MIRAI-HARMONY current scenarios of interest for the main topics covered in our collective efforts, with a primary focus on scenarios involving distributed network operation.

2 Benchmarking and Validation Methodology

The 6G-MIRAI-HARMONY project has only recently commenced; therefore, the benchmarking and validation methodology developed in this deliverable represents an initial framework that will be progressively refined throughout the project. The central aim of the methodology is to establish a coherent and transparent process for assessing the performance and value of the technological enablers developed in WP1 and WP2, and to ensure that the evaluation is repeatable, traceable, and aligned with the overall vision for 6G AI-native radio access networks.

2.1 Methodology overview

In 6G-MIRAI-HARMONY, the benchmarking and validation of enablers will rely predominantly on **computer-based evaluations**, **including link-level or system-level simulations**, targeting to follow initial plans described in the project proposal. These simulation-driven studies will allow partners to investigate the performance and behaviour of advanced PHY/MAC-layer techniques, AI-enhanced processing modules, and cross-layer RAN intelligence under controlled conditions and well-specified scenarios. In some cases, where theoretical insights or performance bounds are important, partners may also complement simulation work with **mathematical modelling and analytical evaluations**. As the project evolves, a subset of enablers will additionally progress toward **early PoC-oriented validation** using laboratory setups or prototype toolchains, leveraging the test and infrastructure plans coordinated by WP3.

A consistent and interoperable description of scenarios is essential for enabling coordinated research across 6G-MIRAI-HARMONY partners and ensuring compatibility with wider European and global 6G activities. Therefore, the methodology will start (Section 2.1) by introducing the initial **scenario taxonomy** adopted by 6G-MIRAI-HARMONY. By initiating defining these elements upfront, we target to ensure that scenarios are traceable, comparable, and compatible across the project's end-to-end benchmarking workflow. The taxonomy also provides a stable interface with technical work packages WP1 and WP2. We also outline how our scenario definitions will be **aligned with the broader 6G ecosystem**, including standardization bodies, major European research programs, and relevant industry and regulatory stakeholders.

A cornerstone of our methodology is the **definition of relevant KPIs and KVIs**, which will form the optimization targets for 6G-MIRAI-HARMONY enablers and the basis against which improvements will be assessed. The methodology will therefore continue (Section 2.2) by identifying the metrics that can capture:

i. **technical performance**, such as spectral efficiency, reliability, latency, energy use, robustness, or computational efficiency; and
ii. **value-oriented aspects**, aligned with the broader expectations associated with future 6G systems, such as sustainability, efficiency of resource use, or contributions to user-centric network behaviour.

To help with the process of choosing the appropriate metrics, we will start with a review of ongoing standardization activities (e.g., 3GPP), European 6G initiatives, and state-of-the-art research results that influence how future 6G systems will be evaluated. Once the KPI/KVI sets are established, 6G-MIRAI-HARMONY will be in better position to define **assessment procedures** to quantify these indicators in a rigorous manner. These procedures will include:

- **scenario-driven simulation workflows**, using the common scenario definitions of Chapter 4,
- **evaluation pipelines for ML models**, covering training, validation, and generalisation behaviour, and
- **baseline comparisons**, where classical signal-processing/optimization methods and state-of-the-art AI/ML-driven techniques will serve as reference points.

This will enable comparative analyses that highlight both performance gains and cost or complexity trade-offs. The methodology also places emphasis on robustness and generalization, especially for AI-enabled and AI-native components. 6G-MIRAI-HARMONY will progressively introduce evaluation procedures that test the behaviour of algorithms under data distribution shifts, unseen network conditions, or new scenario variants. This aligns with the project's objective to ensure that 6G-MIRAI-HARMONY enablers perform reliably not only in idealized settings but also in realistic environments represented by diverse scenarios and datasets.

Moreover, the methodology will highlight **principles for training, validating, and evaluating ML models** that will be followed in 6G-MIRAI-HARMONY (Section 2.4). The main purpose here is to describe the considered data-splitting strategies, model selection procedures, and evaluation protocols in order to ensure reliable, unbiased, and reproducible performance assessment. The section also highlights the importance of **generalization analysis**, emphasizing how 6G-MIRAI-HARMONY will evaluate the robustness of ML models under unseen network conditions to support trustworthy deployment in diverse 6G scenarios.

Finally, we provide (Section 2.5) an overview of the potential **tools and infrastructure** available within 6G-MIRAI-HARMONY project that will be considered to demonstrate and validate a selection from the developed AI/ML-based techniques.

2.2 Taxonomy and alignment with ecosystem

A consistent and interoperable description of scenarios is essential for enabling coordinated research across 6G-MIRAI-HARMONY partners and ensuring compatibility with wider European and global 6G activities. This section introduces the initial scenario taxonomy adopted by 6G-MIRAI-HARMONY and outlines how each scenario will be aligned with the broader 6G ecosystem, including standardization bodies, major European research programs, and relevant industry and regulatory stakeholders.

Scenario Taxonomy

To create a uniform way of describing, comparing, and evaluating scenarios, 6G-MIRAI-HARMONY establishes a structured taxonomy encompassing technical, operational, and contextual dimensions. The taxonomy will serve as a reference template for all partners when defining scenarios and enable consistent mapping toward KPIs/KVIs, simulation configurations, and subsequent benchmarking and validation evaluations.

The initial considered taxonomy includes the following core elements:

- **Scenario Identity and Purpose**: An explicit identifier and brief summary of the scenario's technical and research objectives, along with possible enabler applications.
- **Radio Bands and Spectrum Conditions**: Frequency ranges and bandwidth assumptions.
- **Topology and Infrastructure**: High-level architectural properties such as distributed/cell-free deployments, number of APs/gNBs, edge-cloud split, and location of compute resources.
- **Antenna and MIMO Configuration**: Dimensionality of arrays, beamforming strategy (digital, analogue, hybrid), and use of massive (or ultra-massive) arrays where applicable.
- **Mobility and Channel Models**: Mobility profiles (static, pedestrian, vehicular, aerial) and standardised or partner-specific propagation models appropriate to the scenario.
- **Traffic and QoS Profiles**: Application-driven traffic demands, latency thresholds, reliability constraints, slicing requirements.
- **Impairments and Stress Factors**: Key non-idealities and dynamics to be considered (e.g., CFO/SFO, quantization effects, blockage, hardware impairments).
- **Evaluation Level**: The planned evaluation stage for the scenario (e.g., analytical, link-level, system-level, proof-of-concept).

- **KPIs/KVIs**: Primary and secondary metrics relevant for measuring performance and societal value.
- **Dependencies and Data Requirements**: Expected dataset types, simulation tools, models, infrastructure, and inter-partner dependencies.

Alignment with Japan and the 6G Research and Standardization Ecosystem

The scenario definitions developed within 6G-MIRAI-HARMONY will target to remain coherent with the wider 6G ecosystem to ensure relevance, interoperability and potential uptake of project results. Each scenario will therefore try to be aligned with major ongoing initiatives, including:

Japan, HARMONY: The EU 6G-MIRAI project is collaborating with the partner project in Japan called HARMONY. Together, the 6G-MIRAI-HARMONY partnership has been formed with regular information exchange and alignment on technical and standardization questions and proposals. Both partner projects have a separate work package on technology evaluation, testing and benchmarking. Specifically, HARMONY defined a work package "***WP4: Architecture and Component Technology Evaluation via Testbed Demonstrations***," which shall validate the research efforts outcome through use-case-based technology demonstrations. It consists of two sub work packages: WP4-a) the University of Tokyo Campus Testbed (The University of Tokyo), and WP4-b) Proof of Concept at Construction Sites and Built Environment (Shimizu Corporation).

The University of Tokyo will provide the campus testbed, a platform for demonstrating cutting-edge technologies, accelerating implementation through various international and industry-academia collaborations. The objective is technical validation and proof of concept through end-to-end demonstrations. The Shimizu Corporation activities in areas such as on-site construction, smart building, structural maintenance and disaster response, will serve as realistic application test environment for field PoC. HARMONY partners have started describing the communication challenges of the Torch Tower construction environment as it advances over time (poor signal coverage and blockage, changing work areas, simultaneous connections from many workers/devices, high-reliability low latency for safety management and remote monitoring, disasters and emergency communication).

While the technical scope of JP HARMONY is broader than EU 6G-MIRAI, there are two WPs to which a connection will be established, captured in Table 1.

Funding framework / Call	Project	Technical scope relevant for 6G-MIRAI-HARMONY
JP MICT	HARMONY	**WP1:** AI-native user-centric RAN architecture (cross-layer optimization and RAN/Core convergence to harmonize multi-domain AI)
		WP3: AI-native RAN (PHY and MAC layer resource optimization to improve RAN performance)

Table 1 JP HARMONY technical relevance to 6G-MIRAI-HARMONY

As part of the EU-Japan collaboration, it is planned to exchange both research outcomes as well as data and test set-up information to facilitate the international exchange and collaboration. Detail planning of the test, validation and benchmarking will occur during execution of the respective test and validation work packages of both partner projects.

European Research Programs: 6G-MIRAI-HARMONY aims at leveraging and aligning with benchmarking methodology and data management work from earlier and ongoing EU research projects related to the 6G-MIRAI-HARMONY technical scope of AI-enabled air-interface and AI/ML in RAN. At high level, the scenario taxonomy integrates use-case clusters and value perspectives defined in Hexa-X/Hexa-X-II flagship projects [HexaX][HexaXII]. Furthermore, the notion of Key Values and KVIs, as introduced by SNS JU and 6G-IA and developed further by Hexa-X-II, provides an overarching model for mapping 6G-MIRAI-HARMONY scenarios to societal impacts such as sustainability, trustworthiness, inclusiveness, and resilience. Moreover, Table 2 lists additional EU research projects and describes the identified touchpoints with respect to the technical scope. The finalized Call 1 CENTRIC project is strongly related in particular from 6G-MIRAI-HARMONY WP1 perspective. PASSIONATE is a Chist-Era project which has started in early 2024, hence comparable to SNS Call 2 projects, where the 6G-MIRAI-HARMONY consortium expects concrete synergies. The Call 3 projects 6GARROW and 6G-LEADER can be deemed “sister projects” to 6G-MIRAI-HARMONY, considering their technical scope, their project lifetime and their cross-regional mission. All aforementioned projects, therefore, have both technical relevance to AI/ML in air interface & RAN datasets, as well as methodological relevance with respect to their approach to benchmarking, data management, and validation. We consider 6G-XCEL (SNS Call 2) as well as 6G-DALI (SNS Call 3) to be relevant only from an evaluation methodology point of view.

Funding framework / Call	Project	Technical scope relevant for 6G-MIRAI-HARMONY

EU Chist-Era	PASSIONATE	WP3 AI-native RAN (PHY, MAC, RRM): performance vs. energy consumption
EU SNS Call 1	CENTRIC	User-centric AI-enabled Air Interface (AI-AI): customized waveforms, transceivers, signalling, radio protocols and RRM (WP3 AI-AI PHY; WP4 AI-AI protocols / RRM); AI-AI-compatible, energy efficient HW platforms (WP4)
EU SNS Call 2	6G-XCEL (EU-US)	N/A (no air interface / RAN-related work; only methodology-related relevance)
EU SNS Call 3	6GARROW (EU-ROK)	WP4: AI-native RAN (WP4), resource allocation & energy efficiency, optimization, seamless connectivity, AI compute / Internet of Things (IoT) / immersive services WP3: integrated device-network approaches
	6G-LEADER	ML PHY algorithms, full duplex TX/RX (sparse antenna arrays, self-interference cancellation), non-orthogonal multiple access (NOMA) / random multiple access for machine-type communications (MTC), goal-oriented semantic comms, open RAN
	6G DALI	N/A (no air interface / RAN-related work; only methodology-related relevance: AI experimentation-as-a-Service (MLOps), data/analytics collection (DataOps), model monitoring (concept drift detection, retraining triggers))

Table 2 EU projects with technical relevance to 6G-MIRAI-HARMONY

Standardization Bodies: Although 6G-MIRAI-HARMONY is exploratory in scope, each scenario will consider mapping to relevant standards study or work items where applicable. This may include for example radio enablers for advanced MIMO, AI-native RAN functions, etc.

Aligning scenario assumptions (frequencies, mobility, traffic models, antenna configurations) with standards definitions ensures that 6G-MIRAI-HARMONY outputs will remain interpretable in a standardized context and enhances the potential for future contributions to, e.g., 3GPP Rel-20 and beyond.

Vertical Industries and Technical Stakeholders: Scenario alignment will also consider the viewpoints and constraints of real-world stakeholders such as operators, industrial automation vendors, vertical sectors (manufacturing, transport, healthcare, media), and regulators. This will allow 6G-MIRAI-HARMONY scenarios to represent realistic deployment conditions and operational constraints, enabling meaningful future validation at testbeds and supporting potential industrial uptake.

To operationalize the alignment described above, 6G-MIRAI-HARMONY will strive to maintain a systematic mapping between each scenario and the corresponding ecosystem elements. This mapping will be included as part of the scenario catalogue and will evolve iteratively as partner inputs mature. By making alignment explicit, 6G-MIRAI-HARMONY ensures that scenario definitions reflect both the project's research ambitions and the strategic direction of the European 6G ecosystem. This mapping may capture:

- **Standardization References**: Relevant 3GPP technical reports, study items, and normative specifications associated with the scenario's technical enablers.
- **Use-Case Clusters**: Categorization, if possible and relevant, into broader 6G use-case classes (e.g., immersive communications, robotics and automation, ubiquitous connectivity), aligned with European frameworks such as Hexa-X-II [HexaXIID14] and strategic documents of the SNS JU and 6G-IA, ensuring that 6G-MIRAI-HARMONY scenarios contribute to broader strategic narratives.
- **Deployment Domain**: Context such as private industrial facilities, public macro networks, campus networks, aerial / non-terrestrial-networks (NTN), or hybrid terrestrial–satellite deployments.
- **Ecosystem Roles and Stakeholders**: Identification of primary stakeholders (network operators, vertical industries, equipment manufacturers, regulators), their expectations, and potential contributions to the scenario.
- **Regulatory Context**: Licensing regime of radio bands considered, spectrum availability, any relevant national/regional regulatory constraints, and potential cross-border harmonisation aspects.

2.3 KPI/KVI definition

6G-MIRAI-HARMONY project focuses on developing advanced AI-enabled and AI-native components for the RAN, covering PHY intelligence, MAC/RRM innovations, distributed sensing/positioning assistance, and multi-layer optimization. To assess the benefits of these enablers, 6G-MIRAI-HARMONY requires a structured set of performance and value indicators that capture both technical gains (KPIs) and broader 6G-oriented value dimensions (KVIs).

This section introduces the initial KPI/KVI framework that 6G-MIRAI-HARMONY will adopt for evaluating the enablers developed in WP1 and WP2. The objective is to establish a consistent, comparable, and transparent measurement foundation that will later evolve into full benchmarking procedures in upcoming deliverables.

Since WP1 and WP2 work is still at an early stage, the project's "KPI/KVI matrix" established here represents an initial version that will be refined iteratively as scenarios mature, simulation chains are more established and harmonized, and WP1/WP2 enablers reach higher readiness levels.

Performance indicators

The evolution towards 6G connectivity has intensified efforts to formalize metrics that capture not only technical performance but also broader societal impact. KPIs remain the cornerstone for evaluating air-interface and system-level behaviour, whereas KVIs provide a complementary lens through which to assess contributions of 6G technologies to sustainability, trust, societal well-being, and inclusiveness. This section reviews major existing frameworks, drawing from European initiatives (5G-PPP, SNS JU), standardization bodies (3GPP), and AI-based air interface research directions reflected in Horizon Europe projects such as Hexa-X/Hexa-X-II, CENTRIC, etc. The objective is to identify relevant metrics that can feed the KPI/KVI definitions used in 6G-MIRAI-HARMONY.

KPIs proposed for 6G systems extend the classical 5G performance dimensions (latency, throughput, reliability, etc.) and additionally incorporate requirements emerging from AI-native network operation, high-density deployments, and multi-modal connectivity (e.g., integration of different radio access technologies, concurrent use of different bands, coexistence of communication, sensing and computing, heterogeneous traffic types, etc.). The introduction of KVIs is a major conceptual extension from 5G to 6G. KVIs aim to capture network contributions to societal values, sustainability goals, trust and inclusion, dimensions often inspired by the United Nations (UN) Sustainable Development Goals (SDGs). Unlike KPIs, KVIs do not focus solely on technical outcomes but on broader societal impact. The following subsections summarize the most influential existing KPI and KVI definitions and frameworks.

KPIs in 3GPP

3GPP Rel-18 established, for the first time, a formal KPI suite for evaluating ML-based algorithms in NR air-interface functions [3GPP38843]. These include:

3GPP "Common AI KPIs":

- Performance KPIs: end-to-end KPIs or intermediate metrics appropriate for the measuring, e.g., precision, recall, etc., of ML-enhanced function (CSI feedback, beam management, positioning, etc.).
- Over-the-air overhead for ML: additional signalling/data collection or model transfer required by the ML-based components.
- Inference complexity: real-valued operations, memory footprint, pre/post-processing cost.
- Training complexity: computational load and data demands for model training.
- Lifecycle management (LCM) storage / compute requirements: model update, monitoring, activation and fallback mechanisms.

Use-case specific KPIs, where for each ML-enhanced RAN procedure, 3GPP lists dedicated KPIs, for example:

- CSI compression accuracy vs overhead,
- Beam selection/beam alignment accuracy,
- Positioning accuracy improvements enabled by AI/ML,
- Robustness metrics under mobility and channel variability.

Since 6G-MIRAI-HARMONY aims to go beyond Rel-18/19 AI-enhanced NR features, the 3GPP baseline KPIs provide the formal "anchor" for comparability and may allow 6G-MIRAI-HARMONY to harmonise evaluations with standardization trends, ensure that 6G-MIRAI-HARMONY solutions benchmark against realistic RAN complexity/performance trade-offs, and support 6G-MIRAI-HARMONY's long-term potential contribution to Rel-20+ standardization.

KPIs in 5G-PPP / SNS JU

The 5G-PPP Test, Measurement and KPI Validation Working Group developed one of the earliest and most systematic KPI taxonomies beyond 5G [5GPPP1][5GPPP2]. Their framework has been frequently used in the past as reference for EU research projects. SNS JU follows up 5G PPP work from 2021 and its Test, Measurement and Validation Working Group (TMV WG) is advancing the discussion on the formal definition and validation of 6G KPIs and KVIs, as well as on the harmonization and reusability of testing, measurement methods, and evaluation procedures.

Key elements of recent SNS JU TMV WG work [SNSKPI25], include defining, collecting, and validating performance KPIs in support of a unified European 6G vision. Its white paper consolidates KPI definitions, targets, and contexts contributed by multiple SNS JU projects, aligning them with the ITU-R IMT-2030 framework while also categorizing KPIs into **KPI families**, spanning *data rate and capacity*, *latency*, *reliability and availability*, *mobility*, *compute*, *sensing*, *AI-related capabilities*, Electromagnetic field (*EMF) aspects*, *positioning and localisation*, *energy efficiency*, *coverage*, *compute*, and *other metrics*. From the redefined/expanded traditional metrics we stand out the following as the basic list of metrics to start considering in our technical works:

- Data rate/ Capacity: maximum achievable data rate (in bit/sec) under ideal operating conditions (full resource allocation, error-free reception).
- User experienced data rate: 5th-percentile user throughput (defined as the number of correctly received bit/sec), representing realistic performance in non-ideal deployments.
- Area traffic capacity: spatially normalized capacity (Mbit/sec/m^2), relating spectral efficiency, cell density and bandwidth.
- Maximum bandwidth: aggregated system bandwidth supported across single or multiple carriers.
- Connection Density: total number of connected and/or accessible devices fulfilling a specific quality of service per unit area (in Device/km^2).
- Reliability: capability of transmitting a specific amount of traffic (X bytes) within a predetermined time duration (Y msec) with high success probability of at least $1\text{-}10^{Z}$, where Z denotes the reliability level (a.k.a., the number of nines).
- Latency: Defined for control plane as the transition time (in ms) from the idle state to the active state. In user plane, it is assumed that mobile station is in active state and latency is defined as the time needed to successfully deliver an application layer message in either the uplink or the downlink.
- Spectral efficiency (SE): user-level (e.g. 5th percentile user DL or UL SE) or cell-level (e.g. peak DL or UL cell SE) bit/sec/Hz performance.
- Energy efficiency: ability to minimize the radio access network energy consumption in relation the traffic capacity provided.

These indicators form the basis for assessing the ability of next-generation systems to support densification, extreme bandwidths, and ubiquitous service provision. These KPIs offer a mature baseline and can be useful reference metrics when evaluating radio/RAN enablers (e.g., ML-optimized PHY) that appear in 6G-MIRAI-HARMONY's WP1 and WP2 tasks.

KPIs in AI/ML-relevant EU projects

The Hexa-X project (and subsequent Hexa-X-II project refined) introduced an extended set of KPIs tailored to 6G enablers, grouped into three major clusters: Conventional communication, AI-related and Environment/Physical-world KPIs. Together, these provide a multi-layer view, reflecting Hexa-X/Hexa-X-II's ambition to integrate sensing, communication and distributed AI. Notably, Hexa-X project was the first to explicitly address ***ML-model-centric*** metrics, such as:

- Convergence behaviour,
- Model flexibility/adaptability,
- Training data quality,
- Complexity reduction relative to non-AI baselines.

CENTRIC project further refined KPI categories to specifically address the needs of an AI-native air-interface. CENTRIC defines four KPI groups: *Conventional*, *EMF-related, Common AI-related*, and *Enabler-specific*. The latter two or particular interest to examine (conventional are in-line with SNS JU analysis and EMF-related are not under 6G-MIRAI-HARMONY focus).

Common AI-related KPIs were introduced to measure the computational and operational implications of embedding AI/ML throughout the PHY/MAC:

- Training and inference complexity (FLOPS/operations),
- Model size / parameter count,
- Over-the-air overhead for model exchange,
- Model generalization capability,
- Simulation-to-real performance consistency,
- Inference speed,
- Training loss,
- Model lifecycle storage/compute requirements.

Enabler-specific KPIs included intermediate metrics aligned with specific AI-based enablers (e.g., compression ratio for ML-based CSI compression, beamforming gain for AI-assisted beam management, collision rate for learning-based MAC protocols).

CENTRIC's KPI taxonomy is particularly valuable for 6G-MIRAI-HARMONY because it: a) incorporates AI-related complexity and generalization metrics, b) aligns with AI-native RAN concepts also central to 6G-MIRAI-HARMONY's WP3 architecture and use-case design, c) provides a vocabulary to integrate "model-centric KPIs" into 6G-MIRAI-HARMONY's methodological framework.

Other EU projects, including ImagineB5G, 6G-SANDBOX, 6G-PATH and ACROSS, have been defining ***additional AI-related KPIs***, including among others [SNSKPI25]:

- Classification accuracy
- Model accuracy (MSE, RMSE, SGCS, etc.)
- Model precision (to identify a negative instance as positive)
- Model training effectiveness (on unseen data)
- Level of automation

KPIs in 6G-MIRAI-HARMONY

Following 6G-MIRAI-HARMONY's objectives, the initial identified relevant KPI set is grouped into three main categories.

(a) Common KPIs

This category captures traditional metrics can be used to evaluate *radio performance KPIs*, such as throughput or spectral efficiency (DL/UL, bps/Hz), latency (packet and end-to-end, mean/P95/P99), reliability (block-error-rate (BLER), packet-error-rate (PER), outage probability), coverage/range (achievable signal-to-noise-ratio (SNR) vs. distance, coverage holes), and mobility robustness (handover failure, interruption time). *Energy and resource efficiency KPIs* can be also captured here, that are relevant for both device and network, such as energy per successfully transmitted bit, baseband computation energy, and pilot/feedback overhead reduction. Additionally, network-level KPIs, to be used primarily in system-level simulations, can be included in this category, such as resource utilisation efficiency,load balancing and fairness, scheduling delay, interference management effectiveness, user experience stability over time.

(b) AI-specific KPIs

This category captures KPIs aligned to 6G-MIRAI-HARMONY's AI-native objectives and to be evolved with considered and developed scenarios. Considerations within the project so far include AI-specific KPIs such as generalization error under unseen channel/traffic conditions, robustness to impairments (phase noise, CFO/SFO, non-linearities), training overhead (samples, epochs, compute time), adaptation speed to scenario changes, complexity vs. performance trade-off (FLOPS, memory), ML model inference cost (e.g., Joules/inference).

(c) Enabler/Implementation-related KPIs

This category can include intermediate metrics aligned with specific AI-based enablers; in a similar way it was considered in CENTRIC project. Once 6G-MIRAI-HARMONY PoC definitions and evaluations mature more, it will be possible to track also enabler implementation related metrics such as prototype execution latency, hardware feasibility (memory, compute bounds), and interfacing overheads between modules.

Value indicators

KVI approach in SNS JU and EU projects

The Hexa-X project pioneered the KVI concept and structured it into several large categories, which then offered a high-level, conceptual mapping between 6G enablers and societal expectations [HexaXD14, HexaXD42]:

- ***Sustainability-oriented KVIs***: addressing two main axes:
    - Sustainable 6G, e.g., reduction of energy/environmental footprint, and
    - 6G for sustainability, e.g., for enabling climate-critical applications (monitoring, precision agriculture, etc.).
- ***Inclusion-related KVIs***: addressing universal access and equity of participation, covering, e.g., digital inclusion, accessibility of services, coverage equity, affordability.
- ***Flexibility KVIs***: tied to system adaptability and resilience, covering, e.g., dynamic reconfiguration capability, cost-efficiency under fluctuating demand, robust performance under disruptions.
- ***Trustworthiness KVIs***: spanning security, privacy, and system dependability, covering, e.g., AI/ML security, AI/ML governance, transparency, reduced attack surface, privacy-preserving data processing.
- ***AI-air-interface-specific KVIs***: assessing societal implications of AI-native networks, covering, e.g., transparency of AI decisions, explainability, reliability under uncertainty, safe AI-driven control loops.

Hexa-X also maps these value categories to probable enabling technologies (e.g., distributed AI and automated orchestration for flexibility, energy-aware scheduling for sustainability) and stresses that KVIs should be evaluated both qualitatively and quantitatively depending on data availability.

The SNS JU further consolidated the notion of “key values” (KV), providing structured value categories and linking them to measurable indicators (KVI examples) and KV enabler examples. SNS JU’s approach emphasizes mapping technical enablers to societal objectives (often framed with reference to the UN SDGs). Some key points of the proposed framework include [6GIASNVC]:

- **Categorization aligned with policy priorities:** values are framed into environmental, economic and societal groups. This helps translate technical metrics into policy-relevant indicators (e.g., greenhouse-gas reduction potential, cost-per-user improvements, digital inclusion measures).
- **Enabler-KVI mapping:** it is highlighted that each technical innovation (e.g., energy-aware RAN control, advanced multi-connectivity) must be explicitly linked to one or more KVIs so its broader benefit can be argued and eventually measured. This mapping is intended to support regulatory, funding and industry decisions.
- **Measurement pragmatism:** mixed assessment methods are suggested that include direct measurements where feasible (e.g., energy consumption), modelling/estimation where field data are lacking, and qualitative/impact assessments for societal effects (e.g., jobs or access improvements). This blended approach helps produce actionable KVI reports even in early R&D phases.

The CENTRIC project, which had a very matching technical scope to 6G-MIRAI-HARMONY, defines KVIs as complementary to air-interface KPIs, focusing on how AI-enhanced RAN technologies influence sustainability and societal value creation. CENTRIC groups KVIs into:

- ***Environmental sustainability KVIs***, covering:
    - Impact of RAN enhancements on energy consumption across layers,
    - Support for networks with reduced material waste,
    - Potential reductions in carbon footprint via more efficient control and scheduling.
- ***Economical sustainability KVIs***, covering:
    - Effects on network OPEX and CAPEX,
    - Efficiency in use of spectrum, infrastructure, and compute resources,
    - Gains from automation (reduced need for manual optimization, reduced site visits, etc.).
- ***Societal sustainability KVIs***, covering:
    - Trust in AI-driven network decisions,
    - Perceived transparency and accountability,

- Reliability of critical communication services,
- Support for digital literacy and access equity.

KVIs in 6G-MIRAI-HARMONY

Considering the state-of-the-art KVI approaches captured above, we can converge on a small set of core value themes (sustainability, inclusion, resilience/flexibility, trust) to provide a consistent blueprint for 6G-MIRAI-HARMONY's KVI structure.

In general, 6G-MIRAI-HARMONY will strive to adopt:

- the three-pillar KVI taxonomy (environmental, economic, societal) used by CENTRIC for operational assessments,
- Hexa-X's broader value categories to target linkage to strategic narratives, and
- SNS JU's measurement pragmatism (mix of direct measures, models and qualitative assessment).

Regarding the relevant KVI categories for 6G-MIRAI-HARMONY, we believe that CENTRIC's KVI structure is closer to our project and more methodologically grounded, as it directly couples AI-native operations with measurable impacts (e.g., energy savings via ML-driven scheduling, improved efficiency via learning-based compression, etc.). Similar to CENTRIC project, 6G-MIRAI-HARMONY is "a use-case agnostic project which focuses mainly on generic technological enablers rather than on specific 6G use-cases", rendering it hard to assess the impact of its developed technologies on societal Key Values.

Therefore, 6G-MIRAI-HARMONY *AI-specific KVI considerations* will also be promoted to assess contributions to a sustainable 6G network in the three main dimensions:

(a) **Environmental sustainability KVIs,** capturing energy efficiency improvements and material efficiency improvements.

(b) **Economical sustainability KVIs,** capturing CAPEX reductions and OPEX reductions.

(c) **Societal sustainability KVIs,** capturing user data protection and privacy.

Qualitative contributions of 6G-MIRAI-HARMONY's technological enablers will be assessed, while whenever possible, relevant proxy KPIs can be used to, at least partially, quantify the contribution towards the target KVIs.

2.4 Protocols for training, evaluation, and generalization

In 6G-MIRAI-HARMONY, we will follow standardized protocols in the AI/ML field for the training and evaluation of the ML models. The simplest method to build ML models is by

splitting the dataset in three parts: training, validation, and testing sets. Each data split corresponds to a subset of the original dataset, without repeating data or overlapping areas between splits. A typical percentage split for these is the 70/20/10, which indicates that the first 70% of the dataset is used during training, the next 20% is used during validation, and the final 10% is used for testing [MU22]. We note that the *training split* is used to optimize the model parameters; the *validation split* supports hyperparameter tuning and model selection, while the *test split* offers an unbiased estimate of the model's final performance. The splits can be changed and include a larger chunk of the data into the validation/test sets. This will always depend on several factors such as volume of data available, data quality, and computational cost among others.

During the training process of an ML model, typically an error metric for a selected KPI is computed for each sample or batch (e.g., mean squared error of the latency), evaluating the model performance for that sample or batch. When the error is obtained, the model parameters are optimized using gradient descent. This process is repeated until a pre-defined number of epochs or a given threshold is reached. During this process, the validation split is typically used to compute additional KPIs and therefore have a better understanding of the model performance on different data samples. For example, one could evaluate the latency and the throughput. Computing one or more KPIs on the validation split helps choosing the best model that will later be used to make inferences on the test split.

While the simple train/validation/test split is easy to implement and widely used, it also presents several limitations. The performance estimates can be highly sensitive to how the split is done. In other words, different random partitions of the same dataset may lead to significantly different results, especially when the dataset is small or heterogeneous. In addition, splitting method uses only a single portion of the data for testing, which may not be fully representative. This could lead to biased or unstable performance estimates.

A more robust protocol is **cross-validation** (CV). This method is simple to understand and implement, typically yielding a lower bias than the simple train/validation/testing split sets. Several ways exist to implement CV, but they all share the same underlying procedure:

1. They divide the dataset into two non-overlapping partitions: training and testing.
2. The model is trained on the training partition.
3. Using relevant KPIs for the problem at hand (see Section 2.3), the model's performance is evaluated on the testing partition.
4. Repeat steps 1-3 several times, depending on the specific CV scheme.

From all the repetitions done during steps 1-4, the CV produces a set of KPIs scores (i.e., one per fold or repetition). These scores are then aggregated —typically by computing the mean and variance— to obtain a robust estimate of the model's expected performance. For example, the mean absolute error of the latency could be computed for each repetition when evaluating the model's performance on the testing partition. The model returning the lowest error is selected as best model. This process ensures that the chosen model is not overfitting a particular data split.

There is a wide set of CV techniques [TU22] such as k-fold cross-validation, leave-one-out, leave-p-out, stratified k-folds, repeated k-folds, etc. Despite their advantages, these methods also have several limitations. On the one hand, creating multiple data splits can be computationally expensive when working with large ML models or datasets. This is because the training–evaluation cycle is repeated multiple times, and there is a computational overhead to be paid to split the data or initialize the model's training setup. On the other hand, if the dataset is small or highly imbalanced, some folds may not be representative, which can lead to unstable or misleading performance estimates. In 6G-MIRAI-HARMONY, we will target to analyse the requirements of our problems to ensure we use the proper training methodology and KPIs each time.

Beyond obtaining accurate KPI estimates, it is essential to understand how well the ML models' performance transfers to conditions different than those seen during the training process or in the training data split. This property is known as **generalization**, and it is particularly important in 6G-MIRAI-HARMONY since AI-driven functionalities must operate across diverse network deployments, environments, and device behaviour. To evaluate the generalization capabilities, an effective approach is to assess the model's performance on *unseen data* [DOKI18]. For example, we can compute the KPIs on new geographical areas, new traffic patterns, alternative propagation conditions, novel device types, or new carrier frequencies, among others. In 6G-MIRAI-HARMONY therefore, we will target to compare the KPIs of selected ML models —using a subset of the KPIs defined in Section 2.3— across new data distributions, in order to determine whether they can reliably support real-world 6G use-cases.

2.5 6G-MIRAI-HARMONY tools & infrastructure for validation

KUL testbed

The KUL testbed infrastructure is mostly dedicated to collecting channel state information and In-phase and Quadrature (I/Q) samples of Radio Frequency (RF) signals indoors. It is tightly coupled with the cartesian robot positioner and the motion capture system.

It provides a unique platform for distributed, phase-coherent experimentation in both FR1 (450 MHz–6 GHz) and FR3 (7–23 GHz) frequency ranges. The system combines National Instruments (NI) N321 Universal Software Radio Peripherals (USRPs), supporting 200 MHz instantaneous bandwidth, with external digital clock distribution to achieve precise time and phase synchronization across spatially distributed access points.

An additional sub-7 GHz setup features three NI X410 USRPs (1 MHz–7.2 GHz, 400 MHz bandwidth), configurable for stand-alone or networked operation. The entire software stack is fully reconfigurable, allowing deployment of open-source RAN software or custom-developed signal processing pipelines. Researchers can seamlessly transition between standardized protocol experiments and low-level RF prototyping without hardware changes. The setup is described in further detail in [BRICKSD53].

The core functionality comprises coherent joint transmission and reception, essential for advanced beamforming and sensing experiments. Furthermore, the additional FR3 frontends, based on Pi-Radio hardware, support communication in the 6 - 24 GHz with an instantaneous bandwidth of up to 1 GHz.

There is also significant compute power available as a HP Enterprise Edge server with dual Xeon Gold 6430 CPUs, 1 TB DDR5 memory, NVIDIA A100 GPU acceleration, and 100 GbE connectivity to the radio units. This enables real-time digital signal processing (DSP), large dataset storage, ML model training, and data-driven adaptive waveform generation. A Rubidium frequency standard maintains system-wide synchronization, while multi-band antenna arrays and precision robotic positioners allow controlled spatial experiments in beamforming, localization, and propagation measurement.

Fraunhofer testbed

The testbed in Fraunhofer HHI offers a flexible framework for prototyping physical layer digital signal processing algorithms and MAC layer algorithms, such as MIMO precoder/combiner design and scheduler. Off-the-shelf software defined radios with a different range of capabilities make it possible to implement experimental base stations and users. The USRP N320/1 platform is used to deploy multi-channel phase synchronous RF chains in Sub-6 GHz carrier frequency with up to 100 MHz bandwidth. Smaller USRP B210 and B205 mini platform are used as single antenna user equipment.

The software controlling the software defined radios (SDRs) comes from a mix of open-source RAN projects and software developed internally at Fraunhofer HHI. New features such as uplink multi-user MIMO and TDD reciprocity calibration are developed as internal branches based on open-source projects. On top of unit tests and end-to-end simulations, these algorithms are further validated using the testbed framework with over-the-air transmission running in real-time as a component of a complete cellular modem software stack. Finally, the new components are presented as a technical demonstrator with a real-time graphical interface that presents relevant intermediate signals and performance indicators across the complete cellular stack.

To validate our systems in wireless channel conditions beyond the channel inside the lab, we utilize a Keysight PROPSIM F64 wireless channel emulator. With this system, various channel models, either defined by 3GPP or simulation tools like Quadriga can be placed between the transmit and receive antennas as a virtual channel.

3 Data Collection and Utilization

In the past decades, the signal processing and wireless communications communities benefited from a culture of shared **analytical models and mathematical formulations**. Progress, especially in the PHY, was accelerated by the widespread use of standardized and reproducible models, including:

- shared channel models (Rayleigh, Rician, 3GPP based models, etc.),
- shared noise assumptions (e.g., additive white Gaussian noise (AWGN)),
- shared analytical problem formulations

This high coherence in the community facilitated the comparison of new algorithms, the evaluation of improvements, and extensions of previous contributions. This enabled rapid progress. However, a recent survey of European 6G research projects [SNSAI25] indicates a deep interest in AI/ML. Out of 60 surveyed projects, 49 projects indicated AI/ML approaches as one of their key investigations.

In the AI/ML era, achieving similar coherence is significantly more challenging. Modern **data-driven techniques** do not rely on a shared mathematical foundation of system models, but instead vary extensively along following dimensions:

- datasets with complex distributions and many hidden parameters,
- preprocessing pipelines,
- hyperparameter configurations,
- training dynamics,
- code implementations,
- hardware and runtime environments.

Two models trained on different datasets may produce different results, making fair comparison far more difficult than in classical analytical research. Consequently, while the classical signal processing community advanced through shared mathematical benchmarks, the AI/ML era introduces significant variability and opacity. Ensuring fair comparison, meaningful progress, and compounding development now requires:

- **Common dataset formats and metadata standards**,
- **Transparent reporting of experimental protocols**,
- **Code and data availability for repeatability and verification**.

Modern AI-driven 6G research depends on large, realistic, and well-annotated datasets that capture real propagation environments, network behaviour, and user dynamics. However, such **datasets are often difficult to collect, aimed at single tasks and challenging to reuse**.

As such, a key objective of 6G-MIRAI-HARMONY initiative is to **strengthen the European wireless ecosystem by making high-quality wireless datasets openly available**. To maximize data reusability, we design a data lifecycle plan. In the remainder of this chapter, we provide an overview of the goal of our data lifecycle plan and how our data lifecycle plan works (Section 3.1), which datasets/generators we already provide and have identified from other public sources in this project (Section 3.2), and identified ongoing research projects with potential areas for collaboration on datasets (Section 3.3).

3.1 Data lifecycle

A comprehensive data lifecycle plan provides a structured framework for managing datasets from initial conception to long-term preservation. The lifecycle typically encompasses the following **stages**: *dataset definition*, *measurement*, *preprocessing*, *documentation*, *dissemination*, and *archival*. Each stage plays a role in ensuring that data is accurate, accessible and reusable throughout its lifespan.

Clear dataset definition establishes the scope, measurement conditions, and intended use-cases, preventing ambiguity in interpretation and experimental design. Comprehensive documentation captures essential metadata, preprocessing steps, and contextual details, which are indispensable for reproducibility. Additionally, effective dissemination ensures that datasets can be accessed and reused consistently across research groups.

The increasing use of data-driven methods in wireless research highlights the need for clear distinction between two aspects of reusability: **repeatability** and **reproducibility**, particularly when publishing datasets, models, and experimental results.

Repeatability can typically be achieved through straightforward sharing of datasets, scripts, and execution instructions. When code and data are released together, third parties can often re-run the original experiments and obtain matching results, provided that the computational environment is sufficiently similar.

However, reproducibility represents a more demanding standard. It requires that researchers not only can re-run the original work, but also can **re-implement or re-evaluate the methods on different but comparable datasets or experimental conditions**. Achieving this in a seamless manner requires more than data and source code availability. We outline several extra requirements for such reproducibility:

- datasets must be **properly documented**, and annotated,
- assumptions, preprocessing steps, and hyperparameters must be **explicitly stated** or findable in the code,

- models and algorithms must be described in **sufficient detail** to enable independent reproduction,
- **measurement conditions and system configurations** must be recorded,
- **metadata must capture context information** such as hardware limitations, environmental constraints, software version information and data acquisition factors.

Without this level of methodological transparency, results may be repeatable on the original system but not meaningfully reproducible elsewhere. This distinction is especially critical in 6G research, where datasets depend heavily on specific hardware platforms, such as software defined radios, antennas, cables, etc., but also radio environments, deployment conditions, equipment conditions that are difficult or impossible to replicate exactly.

As such, 6G-MIRAI-HARMONY aims to set **best practices for dataset publication**, including dataset cards, standardized metadata, reproducible code repositories, and open documentation in accordance with recent emerging initiatives in the SNS JU Technology Board. Through this commitment to open and reproducible data, we can support the community by:

- providing real measurement data that complements simulation-based AI development,
- enabling reproducible experimental research,
- lowering the entry barrier for researchers without access to advanced testbeds,
- accelerating the development and validation of 6G ML algorithms

Dataset cards

In the context of AI/ML-driven 6G research, datasets play a central role in shaping model behaviour, performance, and usability. Unlike classical analytical approaches, where shared mathematical models enable straightforward benchmarking, AI/ML-based methods are highly sensitive to dataset characteristics, preprocessing choices, and experimental setups. As highlighted in Section 3.1, this variability makes fair comparison and reproducibility particularly challenging and necessitates explicit dataset documentation and traceability.

To address this challenge, 6G-MIRAI-HARMONY introduces dataset cards as a core element of its data lifecycle plan. Dataset cards provide **a structured and standardized documentation framework for describing datasets** used, referenced, or generated within the project. They do not define new datasets, nor impose constraints on dataset formats or sources. Instead, they act as a harmonization layer across heterogeneous datasets, including public datasets, partner-owned datasets, and consortium-generated synthetic or emulated datasets. Within this deliverable, dataset cards serve as a **reference mechanism for three key purposes**:

(i) justify the selection and relevance of datasets used in the project; (ii) establish traceability between datasets, AI/ML methods, and evaluation metrics; (iii) provide solid foundation for transparent validation/benchmarking in subsequent reports.

By adopting dataset cards early in the project, 6G-MIRAI-HARMONY establishes a common language around data, enabling meaningful comparison of AI/ML approaches and supporting cumulative progress across partners and research tasks. Hence, the primary objective of the dataset cards is to **ensure that datasets can be understood, compared, and reused** across work packages and research tasks. By explicitly capturing dataset origin, scope, content, and intended use, dataset cards reduce ambiguity in experimental design and help mitigate hidden biases arising from undocumented assumptions. This is particularly important in 6G-MIRAI-HARMONY, where datasets are expected to support diverse research activities ranging from:

- ***WP1 – AI-native practical 6G air interface:*** dataset cards document channel characteristics, hardware impairments, and PHY/MAC-level features required for realistic AI-based air-interface and baseband design.
- ***WP2 – AI-enabled distributed 6G RAN architecture and control:*** dataset cards describe RAN telemetry, traffic patterns, and control-related KPIs needed for AI-based network control and architecture design.
- ***WP3 – Scenarios, data, validation and benchmarking:*** dataset cards provide the foundation for scenario definition, dataset traceability, and reproducible benchmarking across heterogeneous experimental setups.

Each dataset card shall capture at least a **minimum, essential set of information** aligned with the data lifecycle stages introduced in Section 3.1, including dataset definition, generation or measurement conditions, documentation, and intended reuse. The information structure of the dataset cards is derived from the project's internal dataset inventory and mapping activities and includes dataset identification and availability, data origin and generation method, technical scope and context, data content and features, associated ML models, relevant KPIs and KVIs, target WP and task relevance, and known limitations or research gaps. Table 3 below captures our initially considered dataset card template.

Field	Description
Dataset name	Descriptive name reflecting network domain and research focus
Version	Dataset version or release identifier
Provider / Owner	Partner name / Public source / Consortium
Dataset type	Measurement-based / Simulation-based / Emulation-based / Hybrid

Generation / Collection method	Brief description of how the dataset was obtained or generated (e.g., testbed measurements, simulator, digital twin, emulation)
Network domain	PHY / MAC / RAN / Control / Cross-layer
Topology / Architecture	Single-cell / Multi-cell / Cell-free / D-MIMO / Distributed RAN / Core-assisted
Traffic / Service context	eMBB / URLLC / mMTC / Mixed-service scenarios
Key data content	Main signals, features, or telemetry included (e.g., channel states, SINR, KPIs, traffic traces, control actions)
Time / spatial resolution	Sampling rate, slot-level, frame-level, or aggregated statistics
Associated AI/ML methods	Supervised learning / Unsupervised learning / Reinforcement learning / hybrid- or model-driven AI
Target KPIs / KVIs	Performance or value metrics derivable from the dataset (e.g., throughput, latency, reliability, energy efficiency)
WP / Task relevance	Relevant 6G-MIRAI-HARMONY WPs and tasks supported by the dataset
Intended use	Training, validation, testing, benchmarking, or scenario definition
Known limitations	Key constraints, assumptions, or missing dimensions
Reproducibility notes	Availability of configuration parameters, seeds, or metadata to enable repeatability

Table 3 Template for 6G-MIRAI-HARMONY Dataset Card

Metadata Format

Finally, we conclude by stressing the importance of a shared standardized metadata format. To maximize the reusability of the datasets generated in 6G-MIRAI-HARMONY, we label our datasets based on the metadata format that is currently being discussed in the SNS JU working group on Testing, Measurement and Validation (WG TMV). This metadata labelling format currently comprises five fields.

- **Data Collection Context:** Describes the conditions under which the data was gathered, including who collected it, what was collected, when, where, and how, along with the tools and methods used. Especially, the exact description of what data were collected using which tools.
- **Application:** For which purposes the dataset is intended.
- **Intrinsic Validation:** What is the actual quality of the dataset. Does it suffer from missing values, errors, noise or redundancies?
- **Context:** What is the broader context of the data? This includes fields like consent of participating subjects, potential impacts on individuals, sensitivity, confidentiality, and information about creators, funders, and licensing.

- **Representation:** How is the data represented? This field is essential for reusability of a previously recorded dataset.

In Table 4 below we present the detailed metadata collection form, which is based on previous discussions in the SNS Working Group on Testing, Measurement and Validation. We have adapted this table according to our views on the matter and how 6G-MIRAI-HARMONY will address this issue. Due to the heterogeneity of the targeted dataset, this table might be appended with additional fields later into the project.

Data collection context
Who collected it?
What was collected?
When was it collected?
Where was it collected?
How was it collected?
• Which software was used (all components and software versions)
• Hardware system details (depending on the setup this might include antennas, cables, software-defined radios, …)
• What was the process of data collection (setup of transmitters / receivers, movement of transmitters / receivers, …)
Application
What was the original goal for the dataset creation?
Is there an explicit target variable?
Is the dataset accompanied by baselines or example usages?
What are the tasks in which the dataset has already been used and their results?
What are the recommended use-cases or tasks?
Which use-cases are not recommended?
Intrinsic validation
Are there missing values or features?
What are the errors, noise or redundancies in the dataset?
What are the relevant statistics for the dataset (means, std deviations, etc.)?
Context
What was the time frame of data collection?
Does the data involve humans and are there any ethical considerations?
• How were the ethical considerations evaluated (was there a review process?)
• What was the information on individuals' knowledge of data collection?
• Was there consent for data collection?
• What are the potential impacts of the dataset and its use on data subjects?
• Are there any identifiable characteristics of individual or sub-populations?

Is there sensitivity data in the dataset?
Does the dataset contain dangerous data?
Is there any confidential data in the dataset?
Is there any repository that links to papers or system that use the dataset?
What is the license and terms of use?
Who is supporting, hosting, maintaining the dataset?
What is the owner contact information?
How can the dataset be extended, augmented, built on, or contributed to?
Who are the dataset creators?
Who are the dataset funders?
What is the dataset's DOI?
Representation
What does each instance represent?
How many instances are in the dataset?
Are the instance features in a standardized format?
What was the sampling, preprocessing, cleaning, and labelling procedures?
• Are these procedures standard or custom? If custom, provide the implementation
Who was involved in the data sampling, preprocessing, cleaning, and labelling procedures?
What other possible sampling, pre-processing, cleaning, and labelling procedures can be performed?

Table 4 Metadata collection form with 6G-MIRAI-HARMONY updates

3.2 Existing datasets and generators

In the following, we present the existing datasets and generators that are publicly available 1) from 6G-MIRAI-HARMONY partners 2) from other sources outside the project. The goal of this collection is to better understand what datasets are possible to develop given the current test facilities, as well as to provide a starting point of potentially useful existing datasets/generators for the technical research developments in WP1 and WP2.

Datasets in KUL

1. Ultra Dense Massive MIMO Channel Measurement

- Overview: This dataset contains thousands of Channel State Information (CSI) samples collected using the 64-antenna KU Leuven Massive MIMO (MaMIMO) testbed. The measurements focused on four different antenna array topologies; Uniform Rectangular Array (URA) LoS, URA NLoS, Uniform Linear Array (ULA) LoS and, Distributed (DIS) LoS. The user's channel is collected using CNC-tables, resulting in a dataset where all samples are provided with a very accurate spatial label. The user position is swept across a 9 square meter area, halting every 5 millimeter, resulting in a dataset size of 252,004 samples for each measured topology. To the best of our knowledge, this is the biggest open dataset containing measured MaMIMO CSI samples. The dataset can, for example, be used to visualize precoders and validate positioning algorithms.

- Code example: https://codeocean.com/capsule/0881566/tree/v1
- Dataport: https://ieee-dataport.org/open-access/ultra-dense-indoor-mamimo-csi-dataset
- Related publication: A. Colpaert, S. De Bast, R. Beerten, A. P. Guevara, Z. Cui and S. Pollin, "Massive MIMO Channel Measurement Data Set for Localization and Communication," in *IEEE Communications Magazine*, doi: 10.1109/MCOM.004.2200716.

2. **Multimodal Sensing Dataset**
   - Overview: This is a novel millimeter wave communication and radar sensing co-existing dataset. The measurement campaign was performed for blockage prediction with diverse human activities. 26 GHz Orthogonal Frequency Division Multiplexing (OFDM) multi-beam communication testbed and 77 GHz Frequency-Modulated Continuous-Wave (FMCW) multiple input, multiple output (MIMO) radar multi-monostatic set-up were configured. The corresponding bistatic channel state information and multi-monostatic backscattered channels are pre-processed for preliminary domain shift analysis by means of visual pre-processed sample inspection.
   - Data: https://data.4tu.nl/datasets/838cf4b1-c9c5-4488-bbd7-bd794c4894c1/2
   - Related publication: B. van Berlo *et al.*, "26 GHz OFDM and 77 GHz FMCW Radar Dataset for Domain Shift Invariant Blockage Prediction," *2023 IEEE 3rd International Symposium on Joint Communications & Sensing (JC&S)*, Seefeld, Austria, 2023, pp. 1-6, doi: 10.1109/JCS57290.2023.10107463.
3. **Massive MIMO with User Equipment mounted on Drone**
   - Overview: A UAV MaMIMO communication platform for channel acquisition was first designed. Then, a testbed was used to measure uplink Channel State Information (CSI) between a rotary-wing drone and a 64-element MaMIMO base station (BS).
   - Data: https://doi.org/10.48804/0IMQDF
   - Related publication: A. Colpaert, Z. Cui, E. Vinogradov and S. Pollin, "3D Non-Stationary Channel Measurement and Analysis for MaMIMO-UAV Communications," in *IEEE Transactions on Vehicular Technology*, doi: 10.1109/TVT.2023.3340447.
4. **Respiratory Sensing in Distributed MIMO**
   - Overview: In this experiment, we rely on the UL channel estimation at the baseband unit (BBU). The single-antenna UE is also driven by software defined radio (SDR) connected to a host computer for signal processing. Tight

synchronization between the user and the BS is achieved by connecting the local oscillators by coax cable. The UL channel estimates are collected at a sample rate of 200 Hz. The testbed operates at a center frequency of 3.51 GHz, with 100 subcarriers spanning an 18 MHz bandwidth with a subcarrier spacing of 180 KHz. The single-antenna UE employs an omnidirectional antenna for transmission, while the receiver antennas in the BS are patch antennas. The ground truth is measured via Qualisys Miqus M3 Motion Capture cameras. This is a marker-based multi-camera motion capture system. The 3D motion capture data is collected at a sample rate of 150 Hz and is used as ground truth for the respiration estimation

- Data: https://ieee-dataport.org/documents/respiration-sensing-cell-free-massive-mimo-0
- Related publication: H. Xiong, R. Beerten, Z. Cui, Y. Miao and S. Pollin, "BS-Breath: Respiration Sensing with Cell-free Massive MIMO," *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* Hyderabad, India, 2025, pp. 1-5, doi: 10.1109/ICASSP49660.2025.10888046.

Other existing Datasets/Generators

Table 5 summarizes some relevant identified existing datasets and/or generators from sources outside the project.

Dataset / Generator Name	Source	License / Tool	Description	Reference / Link
KUCG	Kyoto University	Public (evaluation edition, MATLAB)	An interactive and intuitive platform to generate channel impulse responses for 60 GHz (millimeter wave), 95 GHz and 105 GHz (sub-THz) bands, which are based on real-world channel measurements	https://www.dco.cce.i.kyoto-u.ac.jp/ja/ku-gen/kucg/
QuaDRiGa	Fraunhofer HHI	Public (Octave)	QUAsi Deterministic RadIo channel GenerAtor for generating realistic radio channel impulse responses for system-level simulations (Compatibility with 3GPP TR 36.873 v12.5.0, 3GPP TR 38.901 v16.1.0, mmMAGIC channel model)	https://quadriga-channel-model.de/
Sionna	Nvidia	Open Source, (Apache license), Python	Open-source library for research on differentiable communication systems, including ray tracer for radio propagation modelling, link-level simulator for wireless communications systems and system-level simulation	https://github.com/NVlabs/sionna

			functionalities based on PHY abstraction.	
Gym Logs	Open RAN	Open Source (GPL-3.0 license)	Dataset collected from a real Open RAN experimental deployment integrating a Near-RT RIC and learning-enabled xApps. Captures end-to-end interactions between the RAN, control applications, and user traffic, enabling the study of closed-loop control in operational settings. The dataset records detailed time-series KPIs and control actions observed during live experiments.	https://openrangym.com/datasets [BOPO22]

Table 5 Identified existing datasets with potential relevance to 6G-MIRAI-HARMONY

Usability in 6G-MIRAI-HARMONY

The **Distributed Massive MIMO datasets** from KUL contain multidimensional tensors depicting the Channel State Information over frequency, space, and/or time, allowing for full characterization of a certain propagation environment. For 6G-MIRAI-HARMONY, these datasets are readily available publicly to be used, and can allow for upscaling and validating the AI-based end-to-end wireless interfaces for a massive MIMO scenario. This, for instance, allows the ML model adaptation of (part of) the PHY to the specific propagation conditions, also considering the antennas' precoding or vector combining.

Furthermore, several additional publicly available datasets and generators may prove useful to the project:

- **QuaDRiGa** (QUAsi Deterministic RadIo channel GenerAtor) is a flexible and widely used channel simulation tool for generating realistic radio channel impulse responses in system-level simulations of mobile and wireless networks [JARA14]. For 6G-MIRAI-HARMONY, this tool (developed by 6G-MIRAI-HARMONY partner Fraunhofer HHI) can enable the modelling of MIMO radio channels for specific network configurations, such as indoor, satellite or heterogeneous configurations. It can support three-dimensional, geometry-based stochastic modelling of MIMO channels across a wide range of scenarios, including indoor, outdoor, satellite, and heterogeneous deployments, and is calibrated against established 3GPP channel models for LTE and 5G New Radio. QuaDRiGa enables spatially and temporally consistent channel generation, multi-frequency and large-array simulations, and realistic user mobility and environment dynamics. It can be used to systematically generate reproducible and configurable channel datasets tailored to specific network layouts, frequencies, and mobility patterns, providing a reliable basis for data-driven analysis, algorithm development, and performance evaluation. Besides being a fully-fledged three-dimensional geometry-based stochastic channel model, QuaDRiGa contains a collection of features created in SCM(e) and WINNER channel models along with novel modelling approaches which provide features to enable quasi-deterministic multi-link tracking of users (receiver) movements in changing environments.

- **KUCG** (Kyoto University Channel Generator) is a generator of channel impulse responses for mmW/sub-THz bands, i.e., produces an output file of such link-level channel impulse responses, and could be exploited by 6G-MIRAI-HARMONY partners that are interested in link-level simulations on those bands.
- **Sionna**, a GPU-accelerated library for wireless communication simulation, is designed so that virtually all components (including channel models, signal processing blocks, and ray tracing) allow for the calculation of gradients with respect to their input parameters. This enables end-to-end learning, where the entire communication system can be optimized using gradient-based techniques to maximize performance metrics. Due to its modularity and end-to-end differentiability, it can be excellently suited for 6G-MIRAI-HARMONY evaluations to examine gradual improve of a communication system with highly specific AI blocks for tasks such as signal decoding, channel estimation, etc. Furthermore, additional considerations can be added as extra layers in the modular system such as hardware impairments, and highly realistic channel models based on previous measurements.
- The **OpenRAN Gym Logs** are well suited for 6G-MIRAI-HARMONY as they provide experimentally collected data from a live O-RAN deployment integrating a Near-RT RIC and learning-enabled xApps. The dataset captures time-aligned network observations, control actions, and resulting KPIs, which are essential for studying AI-driven closed-loop RAN control. This makes it particularly relevant for WP2, where ML-based algorithms for scheduling, power control, and near-real-time optimization are investigated. The dataset enables realistic training and benchmarking of reinforcement learning–based control policies under operational conditions consistent with O-RAN principles.

3.3 Collaboration framework for datasets

6G-MIRAI-HARMONY aims at leveraging work from earlier and ongoing research projects in terms of framework and methodologies for data management and benchmarking, as well leveraging existing and emerging datasets related to the 6G-MIRAI-HARMONY technical scope of AI-enabled air interface and AI/ML in RAN.

Table 6 shows the projects introduced in Section 2.2, with respect to their datasets framework relevance from a 6G-MIRAI-HARMONY WP3 perspective and gives links to the respective datasets (where available) or information relevant to data management. At the time of writing of this deliverable, results are only publicly available from the PASSIONATE and CENTRIC projects. With work in 6G-MIRAI-HARMONY progressing, we will iteratively examine results from other projects once they become available and also pro-actively reach out in particular to the mentioned "sister projects" to leverage synergies of the concurrent project execution regarding the generation of datasets as well as the data management framework and methodologies.

Funding framework / Call	Project	relevant for 6G-MIRAI-HARMONY WP3	Deliverable	Link / availability
JP MICT	HARMONY	WP4 Testbed and demonstration (building AI optimization profiles for industrial use-cases)	*To be clarified*	*To be clarified*
EU Chist-Era	PASSIONATE	WP5 Data acquisition and performance assessment	D5.1 Open-access database of the measurement campaign	https://passionate.webs.tsc.uc3m.es/deliverables/
EU SNS Call 1	CENTRIC	WP5 AI-AI Testing and Validation	D5.5 dataset of measurement results D3.4 CENTRIC's AI-Based MIMO toolkit	https://zenodo.org/records/15979632 https://zenodo.org/records/13868444 https://github.com/CENTRIC-WP3
EU SNS Call 2	*6G-XCEL (EU-US)*	*(only relevant wrt. methodology) WP5*	*D5.2 Curated Datasets* *D5.4 Developed Benchmarks & Datasets*	*https://www.6g-xcel.eu/deliverables/ (not available yet)*
EU SNS Call 3	6GARROW (EU-ROK)	WP5 Develop Experimental Testbeds and Proof of Concept Platforms	D5.1 / D5.3 6GARROW Validation and Experimental Testing	https://6garrow.com/results/ (not available yet)

		Implement experimental testbeds to validate new AI-native concepts in 6G networks.		
	6G-LEADER	*To be clarified*	*To be clarified*	*To be clarified*
	6G DALI	*(only relevant wrt. methodology)*	*PoC 1: Data management and experiment on demand (Not available yet)*	*https://6gdali.eu/proof-of-concepts/ (not available yet)*

Table 6 Projects with datasets and methodology relevant to 6G-MIRAI-HARMONY

4 Initial Common Scenario Sets

This Chapter provides an initial overview of the research topics and respective scenarios being addressed in 6G-MIRAI-HARMONY. One important target of 6G-MIRAI-HARMONY WP3 is that the continuously developed benchmarking and validation methodology maintains strong links to the broader project architecture. To this end, WP1 and WP2 research topics and prospective innovations/enablers will be mapped to the scenarios, KPIs, and evaluation procedures defined so far in WP3. A preliminary attempt is presented in this report.

We first (Section 4.1) present an initial catalogue of scenarios that will be used to study and evaluate the 6G-MIRAI-HARMONY technological enablers. Our purpose here is to describe how the project's technical efforts are grounded in a diverse yet structured set of representative deployment scenarios, with particular emphasis on distributed and AI-native radio access networks. By outlining the considered topic areas and their associated scenarios, the Chapter establishes a common context for subsequent analysis, benchmarking, and validation activities across the project. It can help us identify commonalities among the various technical works or even novel scenarios and challenges for 6G.

We also make a first attempt (Section 4.2) to organize the various 6G-MIRAI-HARMONY research topics and enablers in main common research areas of interest. The purpose of grouping into broader research areas is to support the identification of shared scenario sets. Within each currently identified common research area, we summarize the respective considered scenarios that address the main technical topics pursued across the 6G-MIRAI-HARMONY project.

Finally, we present (Section 4.3) some identified areas for potential synergies, i.e., research topics of common interest to multiple 6G-MIRAI-HARMONY partners that could benefit in the future from common scenario sets, in terms of, e.g., benchmarking assumptions, datasets, simulation or demonstration tools and setups, etc. As the project advances, the revisited enablers, mappings and synergies will ensure consistency across WPs and facilitate integration into WP3's validation and demonstration activities.

4.1 Research topics and mapping to benchmarking and validation methods

6G-MIRAI-HARMONY project focuses on research topics and respective enablers that aim to future-proof 6G PHY technologies and open RAN architectures, with a strong focus on AI/ML-driven multi-antenna solutions that bridge fundamental theory and practical implementation requirements, while remaining aligned with ongoing and emerging 3GPP standardization for 6G RAN. In the following, we give an overview of the preliminary considerations on the research topics, enablers and scenarios in WP1 and WP2, while we attempt to perform an initial categorization and mapping to considered benchmarking and validation methods.

6G-MIRAI-HARMONY research topics and scenarios overview

Overall, as part of its use-case portfolio, 6G-MIRAI-HARMONY project addresses key multi-node connectivity challenges across three representative deployment scenarios: wide-area distributed MIMO networks, multi-carrier distributed MIMO systems, and multi-carrier cell-free networks. The project pursues a unified next-generation framework that builds upon and extends existing multi-transmission/reception-point (multi-TRP), multi-downlink-control-information (multi-DCI) and non-collocated carrier aggregation concepts, enabling seamless connectivity across multiple network controllers while preserving a single protocol stack at the user equipment. In parallel, the project investigates generative AI approaches for producing realistic, context-aware CSI samples, such as models conditioned on user location, to augment training datasets for AI-native radio algorithms. This capability can significantly reduce reliance on costly measurement campaigns and supports advanced RAN optimization and AI-for-PHY evaluation in the complex propagation conditions anticipated for 6G. In addition, 6G-MIRAI-HARMONY develops the underlying theoretical foundations and AI-driven solutions for CSI acquisition, as well as joint power control, beamforming, and scheduling, while explicitly accounting for practical impairments in next-generation distributed MIMO systems, including phase estimation uncertainties.

In Table 7, we capture in a unified matrix of the ongoing 6G-MIRAI-HARMONY research topics and enablers considered, organized by WP (1/2). For each research topic, we provide a brief description of the enablers currently considered by 6G-MIRAI-HARMONY partners, as well as details on respective scenarios for evaluations.

Research Topic	Partner	Enabler(s) Description	Scenario details
WP1			
SRS-based channel estimation with real data	Ericsson	ML-based trained channel estimation based on real measurement.	Urban Macro, outdoor, 1 BS, 1 UE, MIMO, TDD
CSI-RS port overhead reduction	Ericsson	ML-driven reconstruction of channel based on spatially sparse CSI-RS.	Urban Macro, both indoor and outdoor, MIMO, mainly FDD
MIMO end-to-end AI for PHY	KUL	Train end-to-end AI capability for modulation & signal detection under cell-free scenario. Practical scenario considering measured MIMO channel & HW impairments.	Cell-free, indoor, 1 user antenna, up to 32x2 AP-antenna's, OFDM LTE-like signal, pedestrian user mobility
GenAI for CSI prediction	Telefonica	Develop a GenAI-based channel model for data augmentation. The objective is to improve the state-of-	1 BS, Multiple UEs with different positions, Outdoor, Real-world antenna configuration

		the-art models by including information from the environment when making the inference	
CSI pre-processing for ML-based compression	SEQ	Investigate low complexity CSI acquisition and pre-processing techniques that enable better compression and robustness for ML models while providing high reliability against imperfect CSI	MIMO (e.g., 1 UE/BS - 4x32); 5G-like air interface; 3GPP channel model, sub-6 GHz, 100MHz bandwidth, Low UE mobility (e.g., TDL-A)
Resource allocation for distributed MIMO systems (power control + beamforming)	Fraunhofer	Approximate optimal beamformers, transmit power, and clustering algorithms for resource allocation with AI techniques	Low mobility, limited cooperation between access points, algorithms should be independent of specific simulation parameters
Angular power spectrum estimation	Fraunhofer	Block-sparse with unknown block partition angular power spectrum estimation in MIMO communication systems based on latent optimally partitioned (LOP) -l1/l2 penalty	Uniform linear array with directive antennas based on 3GPP with 8,12,14,16,20,24,28,32 antennas and carrier frequency 2.1 GHz
Cell-free massive MIMO under imperfect LoS phase tracking	Fraunhofer	Channel estimation and beamforming framework exploiting imperfect LoS phase information	3GPP channel model based with 100 access points and 40 users and LoS phases are uniformly distributed random variables over [-pi,pi]. Rician factors considered [1,5,20,100]
Scalable synchronization and reciprocity calibration solutions	CNIT	Maximum-Likelihood algorithms for CFO/SFO estimation in multicarrier (OFDM) communication systems. Theoretical Limits on Estimation Accuracy: Cramér–Rao Lower Bounds	3GPP channel models. Oscillator instabilities with magnitudes on the order of a few parts per million (ppm).
Traffic-aware ML-based beamforming in cell-free mMIMO	ISRD	Design of ML-based beamforming strategies that adapt RU-level transmission weights based on traffic flow, load distribution, and spatial channel conditions in a distributed RAN. Learning models exploit cross-RU coordination enabled by the D2 interface to improve spectral efficiency and traffic balancing.	Cell-free massive MIMO, distributed multi-DU / multi-RU, sub-6 GHz, outdoor and mixed environments, TDD, multi-UE
WP2			
Distributed RAN control for multi-connectivity	APP	Performance comparison between legacy and multi-connected case	1 UE, 2 RU/ DU, 1 CU; 5G air interface; 3GPP channel model, different UE positions, no mobility
Time-sharing /scheduling	Fraunhofer	Mathematical framework for identifying scenarios where time-	Framework independent of scenario/network setup but demonstrated for small uplink

		sharing/scheduling users is needed to improve network performance	cell-free network with 4 access points equipped with 2 antennas and 3 single antenna users in 100m x 100m area
Scheduling in overloaded cell-free massive MIMO	Fraunhofer	Joint power control, beamforming and time-sharing/scheduling in cell-free networks with lower spatial degrees of freedom than number of users	User-centric cell-free network, 9 and 3-APs considered, 30-UEs. Based on 3GPP models
Joint user scheduling and fronthaul resource allocation	CNIT	Fronthaul resource allocation with multiple sequential finite-capacity links.	3GPP channel models, cell-free massive MIMO architectures, single-antenna users.
ML-based power allocation in cell free network with AP-UE association	UNIPI	Train Transformers for joint power allocation and AP-UE association for the max-min fairness problem	UL/DL, 16 APs, 1-40 UEs
ML-based uplink power control, scheduling and flexible fronthaul optimization in distributed RAN	ISRD	Development of ML-based closed-loop control mechanisms for uplink power control, scheduling, and flexible fronthaul optimization targeting URLLC requirements. ML models operate in near-real-time using RAN telemetry and D2-enabled coordination among multiple DUs and RUs.	Cell-free massive MIMO, distributed multi-DU / multi-RU with D2 interface, sub-6 GHz, URLLC traffic, low-latency constraints
Neural receiver	KUL	Neural receivers considering data-driven channel and hardware models, considering realistic impairments and deployment limitations such as finite fronthauling capacity.	UL, 32/64 BS antennas, 1 – 10 users, 3GPP Channel Models as baseline

Table 7: 6G-MIRAI-HARMONY currently identified research topics

Initial benchmarking/validation mapping matrix

Next, we attempt to categorise the research topics and enablers into more general research areas and we present the methods and KPIs currently considered relevant for benchmarking and validation. We note again that research in WP1 and WP2 only recently started, hence, the current benchmarking considerations are not fine-tuned yet and a more generic description is currently adopted.

In Table 8, we note the general research areas per categorise 6G-MIRAI-HARMONY research topics and capture the evaluation plans, simulation tools, and expected metrics currently considered relevant for benchmarking and validation in each topic.

Research Topic	Research Area(s)	Evaluation plans and Simulation tools	Expected Metrics
WP1			
SRS-based channel estimation with real data	Channel information	Testbed platform (real data collection) + system or link level simulation. Internal tool.	DL throughput
CSI-RS port overhead reduction	Channel information	System level simulation. Internal tool.	DL throughput and intermediate KPI (SGCS, NMSE)
MIMO end-to-end AI for PHY	Channel information, Hardware impairments	Link level simulation in Sionna. Testbed platform (real data collection).	Reliability, Spectrum & energy efficiency
GenAI for CSI prediction	Channel information	System level simulation. Sionna Ray Tracing (RT).	Error metrics (e.g., MSE), performance of downstream tasks
CSI pre-processing for ML-based compression	Channel information	Link level simulation. Matlab / Python.	Spectral efficiency, Reliability, Generalization, Learning convergence
Resource allocation for distributed MIMO systems (power control + beamforming)	Resource allocation	Link level simulation. Python, Quadriga channel models.	Spectral efficiency, reliability
Angular power spectrum estimation	Channel information and modelling	Link level simulation	Error metrics – Normalized mean square error
Cell-free massive MIMO under imperfect LoS phase tracking	Synchronization, Hardware impairments	System level simulation.	Spectral efficiency
Scalable synchronization and reciprocity calibration solutions	Synchronization	Link level simulation. System level simulation.	Mean-Squared Estimation Error (MSEE). Evaluation of Spectral Efficiency.
Traffic-aware ML-based beamforming in cell-free mMIMO	Channel modelling, D-MIMO higher layers	System-level simulation with distributed multi-RU channel models; traffic-driven beamforming evaluation using internal simulator and open channel datasets; Python	DL/UL throughput, spectral efficiency, beamforming gain, fairness
WP2			
Distributed RAN control for multi-connectivity	D-MIMO protocol design (e.g., ARQ, timers), Mobility	System level simulation. Internal tool.	DL throughput

Time-sharing /scheduling	Resource allocation	Numerical illustration via toy model	No specific metric, but can be used for e.g. higher spectral efficiency
Scheduling in overloaded cell-free massive MIMO	Scheduling	System level simulation. Python	Spectral efficiency
Joint user scheduling and fronthaul resource allocation	Scheduling, Resource allocation	System- and link- level simulation. Matlab	Spectral efficiency, scalability, energy efficiency
ML-based power allocation in cell free network with AP-UE association	Resource allocation	Link level simulation. Python.	Spectral efficiency, inference time, scalability
ML-based uplink power control, scheduling and flexible fronthaul optimization in distributed RAN	Resource allocation, Scheduling, D-MIMO higher layers	O-RAN-compliant system-level simulation integrated with Near-RT RIC; closed-loop ML evaluation using internal and O-RAN based datasets	End-to-end latency, reliability (BLER), packet success rate, resource utilization
Neural receiver	Resource allocation, hardware impairments	End-to-end differentiable system allowing for hardware impairment-aware physical layer and resource allocation.	Bit error rate, UL throughput

Table 8: Initial mapping of 6G-MIRAI-HARMONY research topics to research areas and benchmarking and validation methods.

In the next stages of the project, each technological enabler evaluation will later define its own specific subset scenario(s) and measurement methods and will additionally consider the KVIs dimension. Generally, each 6G-MIRAI-HARMONY enabler category (e.g., ML-assisted channel estimation, semantic compression, intelligent MAC/RRM, distributed coordination, etc.) naturally maps to a subset of research areas as well as to a subset of KPIs and KVIs:

- **Channel knowledge acquisition and modelling** (including acquisition, representation, prediction, compression, and modelling of radio channels and related state information)
  - ➢ Channel estimation / prediction accuracy, CSI overhead, spectral efficiency impact, reliability, learning convergence and generalization capability, energy efficiency, inference complexity and latency
- **Resource allocation and scheduling intelligence** (including AI-enabled decision-making for power control, scheduling, beamforming, user association, and interference management)

- Spectral efficiency / throughput (DL/UL), resource utilisation efficiency, fairness, stability of allocation decisions, adaptation speed to traffic or channel dynamics, scalability (users/APs), operational efficiency

- **Higher-layer aspects of D-MIMO networks** (including distributed RAN control, mobility management, coordination across nodes, fronthaul-aware operation, and protocol-level intelligence)
  - End-to-end latency, reliability and service continuity, mobility robustness (handover interruption time), control signalling overhead, resilience and robustness, QoS consistency
- **Synchronization and hardware impairments** (addressing non-ideal hardware effects and synchronization challenges in distributed and cell-free MIMO deployments)
  - Synchronization accuracy and stability (phase offset, carrier frequency offset), computational and calibration overhead, spectral efficiency degradation due to impairments, reliability under non-ideal hardware, deployment feasibility, energy efficiency, localization accuracy

The example mapping provided above, informed to some extend by the initial reported plans above, will serve as some basis for the revised methodology per enabler that will be developed later throughout the project.

4.2 Common research areas and scenarios catalogue

This section categorizes 6G-MIRAI-HARMONY research topics into more broad research areas to help towards the identification of common scenario sets. To this end, within each research area, we capture our current scenarios of interest for some of the main topics covered in our collective 6G-MIRAI-HARMONY efforts. It becomes evident that the primary focus is on scenarios involving distributed network operation, but we also target to revisit classical single-BS deployments.

Area-1: Channel Knowledge Acquisition and Modelling

Acquiring channel information is arguably one of the most challenging tasks in network design, even with the advent of AI/ML-based solutions. This is particularly difficult for distributed implementations, since information sharing typically needs to take place. In this context, several different scenarios are being explored in 6G-MIRAI-HARMONY. The first scenario we consider is the 3GPP Urban Macro cell, which supports both indoor and outdoor operation. In this scenario, we consider a MIMO communication with a single base station and a mobile user. Furthermore, time-division duplex is the mode of operation **(Ericsson)**.

Another scenario concerned with CSI acquisition overhead reduction is an interference-limited user-centric massive MIMO deployment with minimal distributed AP cooperation. In this scenario, we assume low mobility and the sub-6 GHz frequency range with bandwidth of 100 MHz (**SEQ**). A closely related problem is channel modelling. In this project, data-driven channel modelling is performed using scenarios derived from 3GPP statistical channel models and ray-tracing models with a single transmitter and multiple uniformly distributed users equipped with linear antenna arrays. Initially, no user mobility is assumed (**Telefonica**). ML-based beamforming and control algorithms use the channel knowledge outputs as inputs to higher level learning mechanisms especially for cell-free massive MIMO scenarios.

Area-2: Resource Allocation and Scheduling Intelligence

Another main topic of interest in 6G-MIRAI-HARMONY is resource allocation, especially on PHY aspects such as power allocation and beamforming. The first scenario considered here regards uplink overloaded user-centric cell free networks, i.e., networks with more user antennas than total number of AP antennas. We consider 3 degrees of overload: (i) Slightly overloaded with 30 single-antenna users and 9 APs equipped with 2 antennas each, (ii) Extremely overloaded with 30 users and 3 APs, (iii) a hotspot scenario that follows the setup in (i), but with users concentrated in a single area (**Fraunhofer**). In addition to overloaded regimes, we also consider regimes with varying user loads. In this setup, the APs have a square deployment, and both user-centric and non-user-centric cell-free implementations are studied, operating in both uplink and downlink. The number of users and APs varies between 1-100 and 1-50, respectively (**UNIPI**). For these two scenarios, the 3GPP Urban Microcell scenario is adopted. In addition to square AP deployments, sequential (radio stripe) AP deployment is considered in conjunction with a limited-capacity fronthaul (**CNIT**). Alongside adopting 3GPP models, we also consider measurement-based channel models for distributed MIMO with uniform linear arrays at APs and randomly positioned single-antenna UEs, which include hardware impairment and finite-capacity fronthaul for a more realistic AI-based resource allocation (**KUL**). This area could also include ML-based solutions for uplink power control, scheduling, and flexible fronthaul optimization, with a particular focus on URLLC services. This latter work targets distributed RAN environments with cell-free massive MIMO and multi-DU/multi-RU coordination via the D2 interface, enabling intelligent and adaptive resource allocation under stringent latency and reliability constraints (**ISRD**).

Area-3: Higher Layer Aspects of D-MIMO Networks

In addition to the distributed MIMO scenarios presented above, we consider different types of D-MIMO networks for handling tasks such as cell handover (**Apple, ISRD**). Namely, we consider wide-area distributed MIMO scenarios, where the RAN is represented by virtual cells and the size of the virtual cell is dependent on factors such as the complexity on the network side. Multi-carrier distributed MIMO networks, where the virtual cells can incorporate multiple component carriers are also considered. In addition, we have multi-carrier "cell-free" networks, where the virtual cells are overlapping (**Apple**). Cell-free massive MIMO architectures with multi-DU and multi-RU coordination via the D2 interface, with a particular focus on URLLC use cases, are also considered (**ISRD**). Within this framework, scenarios with tens to hundreds of UEs and a large number of distributed access points, coordinated by multiple cooperating DUs, are adopted. In particular, the number of RUs/APs, DUs, and UEs ranges from 32–128, 4–16, and 50–200, respectively. Furthermore, low-to-moderate user mobility is considered, and channel models follow sub-6 GHz 3GPP specifications.

Area-4: Synchronization and Hardware Impairments

We also consider scenarios that include hardware impairments, such as finite resolution A-D/D-A and phase oscillator noise, raising synchronization challenges. In one scenario, we consider both square and line AP deployments with 8 users and 128/256 total serving antennas (**CNIT**). Furthermore, we address LoS phase estimation uncertainty in a scenario with 100 APs and 40 UEs adopting 3GPP channel models (**Fraunhofer**).

4.3 Opportunities for potential synergies

In the following, we outline some initially identified opportunities for potential synergies among 6G-MIRAI-HARMONY partners, sharing interest on common research areas and/or scenario definitions.

ML-based channel estimation/prediction

The adequate evaluation of ML-based channel estimation/prediction and comparison against state-of-the-art benchmarks requires the utilization of a variety of system configurations depending on, e.g., the availability of real measurement data or the 6G study item (SI) evaluation assumption requirements specified by 3GPP.

For cases where measurement data is available, e.g., based on the uplink sounding reference signal (SRS) transmissions from one user equipment (UE) towards a massive multiple-input multiple-output (MIMO), one of the adopted approaches consists in integrating the channel measurements into a 3GPP-compliant proprietary link-level. Notably, this allows an expansion of the evaluation regimes to include those beyond those measured in the field, e.g., by artificially introducing noise to the measurements. Since the ML-driven task for channel estimation closely resembles the task of generative channel model techniques, a similar model architecture has been implemented by Ericsson for SRS channel estimation and by Telefonica for ML-based channel generation, i.e., a convolutional neural network with residual connections. In spite of not being identical, this similarity could facilitate joint work on ML model optimization.

For cases where the ML-based channel estimation/prediction techniques are proposed as potential candidates for 6G standardization, comprehensive link- and/or system-level simulations compliant with the agreed 3GPP evaluation assumptions are required to be performed. The investigation of common scenarios (e.g. urban areas, cell edge UEs, etc.) among **Ericsson**, **Telefonica**, **KUL** and **SEQ** can allow an understanding of what kind of scenarios are critical for channel estimation and thus what should be tackled for the data generation (e.g. coherent evolution over time to allow time prediction, or cell edge UE where AI/ML can fail). Additionally, close-to-product simulators integrating proprietary algorithmic solutions and accurately capturing state-of-the-art PHY implementations may be utilized to adequately assess the performance potential of candidate solutions and facilitate the comparison with the results from other 3GPP contributors.

Intelligent handling of synchronization errors

Cell-free MIMO systems are based on coherent joint processing, in which multiple distributed APs cooperate to serve the same user. As a result, their performance critically depends on tight synchronization among the APs. In practical deployments, synchronization impairments such as sampling frequency offset (SFO) and carrier frequency offset (CFO) are unavoidable and can lead to a significant performance degradation. These impairments introduce phase and timing distortions that disrupt coherent combining, thereby degrading the quality of the received signals and reducing the overall system efficiency.

In this project, on-going work on handling synchronization errors is being addressed by **CNIT** and **Fraunhofer**. The research activity of CNIT is focused on the design and analysis of synchronization algorithms among APs for multi-carrier (OFDM) cell-free communication systems. The impact of residual synchronization errors on system performance, for instance in terms of spectral efficiency, will also be assessed. Fraunhofer addresses residual synchronization errors at the channel-model level, specifically imperfect tracking of the LoS phase due to mobility, estimation delays, or limited phase-tracking capability. The imperfectly tracked phase is treated as exploitable side information, and channel estimation and beamforming schemes are developed that explicitly capitalize on this residual structure, rather than discarding the LoS component due to its phase uncertainty. Clearly, the synchronization algorithms developed by CNIT could be leveraged as input for the research activities at Fraunhofer.

ML-aided power control

Fraunhofer HHI has developed a framework for optimal joint power control and beamforming design using fixed-point methods. Since these are iterative methods, their implementation can be slow, especially in large-scale deployments. In collaborative work with **CNIT** and **Fraunhofer**, a master thesis has been completed by identifying cases where the framework developed is slow and utilizing neural networks for faster convergence of the proposed iterative method. In particular, using position information, the optimal power control for an arbitrary number of users is predicted and used to initialize the algorithm.

A further extension of the work is to train an ML model to directly predict the power and beamforming vectors jointly instead of just using the ML outcome as an initialization point for the algorithm. By doing so, we can leverage the learning capabilities of ML to generalize to different scenarios while taking advantage of the available iterative algorithm, developed at Fraunhofer HHI, as a method to provide the offline training data.

5 Conclusion

Deliverable D3.1 marks an important first step toward a coherent and project-wide approach to benchmarking, validation, and data management in 6G-MIRAI-HARMONY. By clearly structuring a methodology for benchmarking and validation and by grounding both in a common sets, scenarios, KPIs and KVIs, and methodological principles evaluation protocols and data management, the report addresses a critical prerequisite for meaningful evaluation of project's AI-native technologies.

The state-of-the-art analysis confirms that, while substantial progress has been made in defining 6G performance and value indicators, the increasing reliance on AI/ML introduces new challenges related to reproducibility, comparability, and transparency. The proposed framework responds to these challenges by promoting harmonized evaluation practices, standardized data handling, and explicit documentation of experimental assumptions and procedures. In parallel, we explore directions for alignment in that scope (benchmarking, validation, shared usable datasets) with Japan, global standards, and the European 6G research ecosystem.

The scenario taxonomy, the KPI/KVI definitions and the data collection and utilization principles presented in this deliverable lead to a structured view of the project's research focus, highlighting common areas of interest and potential synergies across partners. These elements will facilitate consistency within and across technical work packages (WP1 and WP2) and enable systematic integration of research outcomes into validation campaigns and demonstrations.

Given the early stage of the project, the methodologies and assets presented here should be regarded as an initial framework rather than a final specification. They will be progressively refined and extended as new enablers are developed, additional datasets become available, and validation needs evolve. Nevertheless, the foundations established in D3.1 are essential to ensure that subsequent results are credible, comparable, and aligned with the broader vision of AI-native 6G radio access networks.

6 References

[3GPP38843] 3GPP, TR 38.843, Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR air interface, Release 18, V19.0.0, September 2025

[5GPPP1] 5G PPP Whitepaper, “Beyond 5G/6G KPIs and Target Values”, Version 1.0, June 2022, https://doi.org/10.5281/zenodo.6577506

[5GPPP2] 5G PPP Whitepaper, “Beyond 5G/6G KPI Measurement”, Version 1.0, June 2023, https://doi.org/10.5281/zenodo.7963247

[6GIASNVC] 6G Infrastructure Association, Vision and Societal Challenges Working Group, Societal Needs and Value Creation Sub-Group, "What societal values will 6G address?," 6G SNS IA, May 2022, https://doi.org/10.5281/zenodo.6557534

[BOPO22] Bonati, L., Polese, M., D’Oro, S., Basagni, S., & Melodia, T. “OpenRAN gym: An open toolbox for data collection and experimentation with AI in O-RAN”, April 2022, In 2022 IEEE Wireless Communications and Networking Conference (WCNC) (pp. 518-523). IEEE. https://arxiv.org/abs/2202.10318

[BRICKSD53] 6G-BRICKS, D5.3: “D5.3 Final deployment and Configuration of enablers at 6G sites” , June 2025, https://6g-bricks.eu/wp-content/uploads/2025/07/D5.3_Final_deployment_and_Configuration_of_enablers-_at_6G_sites_submit.pdf

[DOKI18] Doshi-Velez, F., & Kim, B. (2018). Considerations for evaluation and generalization in interpretable machine learning. In *Explainable and interpretable models in computer vision and machine learning* (pp. 3-17). Cham: Springer International Publishing.

[HexaX] Hexa-X Project, "Main - Hexa-X”. https://hexa-x.eu/

[HexaXD14] Hexa-X, Deliverable D1.4, “Hexa-X architecture for B5G/6G networks - final release," July 2023, https://hexa-x.eu/wp-content/uploads/2023/07/Hexa-X-D1.4-Final.pdf

[HexaXD42] Hexa-X, Deliverable D4.2, “AI-driven communication & computation co-design: initial solutions", June 2022, https://hexa-x.eu/wp-content/uploads/2022/07/Hexa-X_D4.2_v1.0.pdf

[HexaXII] Hexa-X-II Project, "Main - Hexa-X-II". https://hexa-x-ii.eu/

[HexaXIID14] Hexa-X-II, Deliverable D1.4, “Final 6G value, requirements and ecosystem”, April 2025, https://hexa-x-ii.eu/wp-content/uploads/2025/05/D1.4-final.pdf

[JARA14] S. Jaeckel, L. Raschkowski, K. Börner and L. Thiele, "QuaDRiGa: A 3-D Multicell Channel Model with Time Evolution for Enabling Virtual Field Trials", IEEE Transactions on Antennas Propagation, 2014.

[MIRAID52] Deliverable 5.2 “Ethical and data management guidelines and requirements”. https://6g-mirai-harmony.eu

[MU22] Muraina, I., “Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts”, In 7th international Mardin Artuklu scientific research conference (pp. 496-504).

[SNSAI25] 6G SNS JU White Paper: “AI/ML as a Key Enabler in 6G Networks: Methodology, Approach and AI Mechanisms” January 2025, https://smart-networks.europa.eu/wp-content/uploads/2025/02/ai_ml_white-paper-sns_tb_v1.0.pdf

[SNSKPI25] SNS JU TMV WG Whitepaper, “6G KPIs – Definitions and Target Values”, Version 1.0, January 2025, https://doi.org/10.5281/zenodo.14621168

[TU22] Different Types of Cross-Validations in Machine Learning and Their Explanations. 2022. Online: https://www.turing.com/kb/different-types-of-cross-validations-in-machine-learning-and-their-explanations. Accessed on: 13th January 2026.