ITUPublications

International Telecommunication Union
Telecommunication Standardization Sector

AI Ready – Analysis Towards a Standardized Readiness Framework

Version 2.0 Interim Report

January 2026

Disclaimer

ISBN

978-92-61-41911-0 (electronic version)
978-92-61-41921-9 (EPUB version)

AI Ready – Analysis Towards a Standardized Readiness Framework

Version 2.0 Interim Report

January 2026

Foreword

Artificial Intelligence (AI) is reshaping the way we address complex societal challenges, offering new possibilities in areas such as healthcare, climate resilience, education, and digital inclusion. The ITU AI Readiness project was launched in 2024 to measure the ease/difficulties and the ability to reap the benefits of AI integration.

Last year, to further advance the discussions, ITU launched the ITU AI Readiness pilot Plugfest to collate and study projects on applying AI to solve real world problems. The ITU AI Readiness project also called for engagement of experts to provide strategic feedback and guidance. 88 experts from 38 countries were carefully selected. Mentoring and comments on the Plugfest projects were provided by the experts in addition to valuable regional perspectives to shape the ITU AI Readiness Framework. This project brings together contributions from multiple sectors – industry, academia, government, and civil society – creating a collaborative environment where ideas, knowledge, and experiences are shared to develop the standardized AI Readiness Framework.

Bringing the experience from analysing use cases, in 2025, an analytical approach was followed in combination with a bottom-up approach. This approach derives dimensions and metrices for readiness analysis from the Plugfest project reports. A way forward for integrating regional customizations is provided in the form of Indices. In addition to the analysis, a practical, living toolkit is designed and presented which can be used by countries, enterprises, Non-Governmental Organizations (NGOs), and other 3rd parties.

We acknowledge the support and are very grateful for the encouragement provided by the Kingdom of Saudi Arabia and the Ministry of Industry and Information Technology of China during this project. We acknowledge also the work done by ITU Members in ITU Study Groups and for their contribution to AI Readiness standards.

As we continue developing the AI Readiness Project, we look forward to deepening our collaboration with partners worldwide, developing AI Readiness standards, building AI Readiness capacity, and contributing to multi-level AI Governance.

Table of contents

List of contributors

Name	Affiliation
Ahmed Said	Ministry of Communication and Information Technology, Egypt
Alireza Yari	ICT Research Institute, Iran
Ameny Khachlouf	Tunisie Telecom
Amit Kumar Srivastava	Department of Telecommunications, India
Amjad Maawia Elnayal	Telecommunications Regulatory Authority of Bahrain, Bahrain
Ammar Saleh Ali Muthanna	Saint Petersburg State University of Telecommunications
Anna Abramova	Moscow State Institute of International Relations (MGIMO)
Antonia Moreno	The National Center of Artificial Intelligence in Chile (CENIA)
Álvaro Soto	Pontificia Universidad Católica de Chile, The National Center of Artificial Intelligence in Chile (CENIA)
Asrat Mulatu Beyene	Addis Ababa Science and Technology University
Aysha Ahmed Alkoheji	Telecommunications Regulatory Authority of Bahrain, Bahrain
Chenxi QIU	China Academy of Information Communications Technology, MIIT of China
Fahad Albalawi	Saudi Data & AI Authority, Kingdom of Saudi Arabia
Habib Mohammed Hussien	Addis Ababa Science and Technology University
Halima Mohamed Ismaeel	Ministry of Transportation and Telecommunications, Bahrain
Ian Nyasha Mutamiri	Postal and Telecommunications Regulatory Authority of Zimbabwe, Zimbabwe
Innocent Nzimenyera	GGGI
Katarzyna Wac	University of Geneva
Kiran Raj Pandey	Health AI for All Network (HAINet)
Lilibeth Acosta	GGGI
Marcelo Gabriel Mendoza Rocha	Pontificia Universidad Católica de Chile, The National Center of Artificial Intelligence in Chile (CENIA)
Maxwell Ababio	Shield Tech Hub
Mohammed Alawad	Saudi Data & AI Authority, Kingdom of Saudi Arabia

(continued)

Name	Affiliation
Munezero Mihigo Ribeus	GGGI
Osmar Bambini	umgrauemeio
Prasha Sooful	NT Health, Australia
Rim Belhassine Cherif	Tunisie Telecom
Shan XU	China Academy of Information Communications Technology, MIIT of China
Shweta Khushu	Vector Institute
Tsafak Djoumessi Pauline Gnimpieba	Ministère des Postes et Télécommunications de la République du Cameroun, Cameroon
Xingzhi MA	China Academy of Information Communications Technology, MIIT of China
Yue QIN	China Academy of Information Communications Technology, MIIT of China

Acronyms

AI	Artificial Intelligence
AI-RE Toolkit	AI Readiness Enablement Toolkit
API	Application Programming Interface
CPU	Central Processing Unit
EG	Expert Group
GPU	Graphics Processing Unit
IAP	Incident Action Plan
IoT	Internet of Things
IP	Intellectual Property
KB	Knowledge Base
KPI	Key Performance Indicator
ML	Machine Learning
NGO	Non-Governmental Organization
SDK	Software Development Kit
TAC	Technical Advisory Committee

1. Introduction

Background

This report provides an analysis of the Artificial Intelligence (AI) Readiness study aimed at developing a framework for assessing AI Readiness, which indicates the ability to reap the benefits of AI integration. By studying the actors and characteristics in different domains, a bottom-up approach is followed, which allows us to find common patterns, metrics, and evaluation mechanisms for the integration of AI in these domains.

The ITU AI Readiness framework aims to engage with multiple stakeholders around the world, assess and improve the level of integration of AI in various domains, study use cases to validate the weightage of the key factors in those domains, improve global AI capacity building, and foster opportunities for international collaboration.

In September 2024, ITU published its first version of the AI Readiness report, where 6 key fundamental factors were identified:

- Open Data: Accessibility and quality of datasets for analysis of AI applications.
- Research: Collaboration between domain-specific and AI research communities.
- Deployment: Infrastructure and ecosystem readiness for AI deployment.
- Standards: Ensuring trust, interoperability, and compliance.
- Open source: Enabling rapid adoption through an open developer ecosystem.
- Sandbox: Platforms for AI experimentation and validation.

To further study the role played by these components in the real practice, ITU and the Kingdom of Saudi Arabia called for engagement from the field and launched a pilot AI Readiness Plugfest during the 2024 GAIN Summit in Riyadh. The ITU AI Readiness Plugfest is an initiative to explain the AI Readiness factors to various stakeholders and allow stakeholders to "plug in" their regional AI readiness factors, such as data accessibility, AI models, compliance with standards, toolsets, and training programs. Additionally, the Technical Advisory Committee (TAC) and Expert Group (EG), composed of experts invited through AI for Good initiatives, provide strategic guidance and feedback on AI readiness projects.

Expert Groups are composed of global experts with different backgrounds coming from 38 countries. Experts are mainly from Academia (33%), government ministries/regulatory authorities (32%), telecommunication companies, research institutes/Think Tanks, regional/international organizations, and private companies. There are 88 experts in EGs, among whom 62.5% come from developing countries. 32 experts are women leading figures in the countries and the domain, representing 36% of all experts.

To study the sandbox environments and their influence on AI readiness, cloud credit support is provided to selected projects, further facilitating the development and deployment of AI solutions in real-world applications.

In July 2025, the third ITU AI Readiness workshop at the ITU AI for Good Global Summit was hosted. The workshop invited global stakeholders, industry leaders, and researchers to foster collaboration on ITU AI Readiness. The workshop served as a compilation of projects towards

ITU AI Readiness 2.0, featuring the sharing of plugfest project learnings along with the partner presentations centering on their understandings of AI Readiness. During the workshop, ITU announced its further steps towards ITU AI Readiness 3.0 activities.

One of the main contributions of this report is the further development of the framework for assessing AI Readiness, which indicates the ability to reap the benefits of AI adoption. After the AI for Good Global Summit in July 2025, we continued our analysis, summarized the learnings from the plugfest project reports. By continuing AI use case studies, initiating consultations with experts from industry, research institute, academia and government, we derived 13 generic dimensions from the expert guidance during the plugfest. Metrices quantify and measure detailed domain-specific values under each dimension. Indices serve as filters or weightages, which capture the granular priorities of the user. Indices could be applied to both dimensions and metrics to allow users to adjust the relevant importance when self-evaluating.

The basic framework and the details are complementary to each other, making the framework available for both policymakers with guidance on AI and domain experts with technical and actionable recommendations.

For better stakeholder engagement around the ITU AI Readiness Framework, ITU designed a pilot AI Readiness Enablement Toolkit (AI-RE Toolkit), which is a dynamic model and a living tool that enables **self-evaluation** for the users. The toolkit uses the principle of a foundational model built from the ITU AI Readiness Knowledge Base (KB) in the ITU AI for Good Sandbox and a finetuned model integrating regional customizations for users to self-assess the AI performance in their context. The ITU AI Readiness Knowledge Base functions as the brain of the toolkit. It is built with AI techniques and gathers input mapped to 6 fundamental factors in the framework. Output from the framework contains the evaluation of the status quo, gap analysis, and customized actionable recommendations. Each time users input new materials, such as the latest version of the report, unstructured data, and deployment stories, the knowledge base can iteratively learn from the new input.

To increase adoption from general users, the ITU AI Readiness Challenge, with a specific focus on 6 factors, was launched by the end of October 2025 during the AI for Good Impact Africa event in Johannesburg, South Africa. Participants were requested to build the basic framework of the knowledge base.

To review the framework, dimensions, the pilot toolkit design, and the standards gap on the ground, several rounds of review meetings with experts from EGs were held, with a specific focus on collecting feedback and potential inputs. From the feedback with experts, potential users of the toolkit were identified, pain points of the users on the ground were noted, and contributions from the member states were discussed.

Insights from AI Readiness Study

1. Strengthening ICT-related higher education, leveraging open-source ecosystems, and engaging with international education and training platforms, and enabling leapfrogging opportunities can accelerate AI skills development.
2. A strong positive correlation exists between national income levels and general digital literacy, measured through ICT skill penetration. However, substantial variation exists within income groups. Middle-income countries often exhibit higher optimism and trust toward AI technologies than high-income economies, creating favorable conditions for large-scale AI adoption if skills gaps are addressed.

Digital skills development accelerates most rapidly at the middle-income stage. ICT skill penetration typically remains low in low-income economies. Policy choices and education investment during this phase play a decisive role in widening or narrowing national AI readiness gaps.

3. Data readiness is a critical determinant of effective, trustworthy, and inclusive AI adoption. Beyond data scale and accessibility, the quality, diversity, representativeness, and labeling of datasets directly shape AI system performance, as well as their fairness, transparency, and adaptability.

4. Insufficient data quality and biased datasets risk reinforcing discrimination and limiting real-world impact, particularly in localized deployment contexts. Strengthening public data openness and data service capabilities – including data collection, data cleaning, and data labeling – is therefore essential to enable scalable and localized AI adoption across priority sectors such as education, agriculture, and transportation.

5. AI readiness globally is constrained by limited data scale and uneven Internet penetration. Global Internet usage stands at 55.56%, indicating that nearly half of the world's population remains outside the digital ecosystem required for large-scale AI data generation. While 57% of countries have Internet penetration above 60%, nearly half remain below 50%, and only 18% of countries exceed 90% penetration, highlighting persistent constraints on global data scale for AI development.

6. Data readiness gaps are driven by service capability and governance, not access alone. On average, developed economies have more than three times as many Internet service providers per million inhabitants as developing economies, with median values showing an even larger gap. In addition, lack of data governance frameworks limits effective and trustworthy data use.

7. Basic network coverage supports entry-level AI use, but advanced network readiness remains uneven. While 96% of the global population is covered by mobile broadband, access to advanced networks remains highly uneven. Global 4G coverage reaches 93%, but only 56% in low-income economies. Global 5G coverage stands at 55%, compared to 84% in high-income economies and just 4% in low-income economies, with significant regional and urban–rural disparities.

8. Shortfalls in computing infrastructure, energy supply, and edge devices constrain AI deployment. Availability of data centers, per capita electricity supply in developed economies is more than twice that of developing economies. IoT market size in developed economies is on average four times larger than in developing economies, limiting the availability of edge devices for AI-enabled industrial applications.

9. Open-source technologies lower entry barriers for AI adoption worldwide. Contributions to major open-source AI and LLM-related repositories extend beyond application-layer development to include core model architectures, training pipelines, evaluation benchmarks, and governance mechanisms. Measurable upstream contributions to top-tier open-source LLM initiatives and participation in the opensource technology development, especially development of foundational and large language models (LLMs), is an important metric of AI readiness.

10. Overall, the level of open-source engagement correlates strongly with other readiness dimensions, including R&D, computing capacity, and the overall innovation ecosystem. R&D capacity is an important dimension of AI readiness, leading to metrics such as stronger AI research output, higher publication impact, and greater resilience in talent development. At the enterprise level, company investment in emerging technologies – including AI, data platforms, and advanced computing – plays a critical role in translating research into scalable systems. Corporate AI R&D expenditure brings cumulative advantages along with robust public research institutions and innovation support mechanisms.

11. Investment patterns influence AI readiness levels. Public investment in AI, supported by effective national AI strategies, will help establish research and innovation systems. Dedicated, multi-year public funding mechanisms for AI research, experimentation, and standards engagement with supportive private investment, including venture capital investment in AI startups, are important. These ecosystems benefit from mature financial markets, strong exit pathways, and dense networks linking startups, research institutions, and large technology firms.

    Investment patterns influence startup formation, scale-up potential, domestic commercialization, and enable AI ecosystems to focus on not only deployment and adoption but also endogenous innovation.

12. Regional evaluation of AI Readiness could be linked to strong performance across all dimensions of AI readiness.

    In some cases, tight linkages between academia, industry, government, and active participation in international AI standardization processes play a decisive role in shaping global technical specifications, reference architectures, and evaluation methodologies.

    In contrast, in some other cases, expanding AI adoption, selective research strengths, but limited influence over foundational technologies, with moderate engagement in open-source AI projects (primarily at the application and integration layers), growing public AI funding, but fragmented governance and coordination, lead to limited participation in core open-source LLM development and international AI standardization.

    Lastly, if there are structural constraints across all dimensions, it will lead to minimal upstream engagement in open-source AI and LLM projects with limited access to private capital and global AI investment networks. AI deployment in such cases is frequently driven by imported technologies, increasing dependency risks and limiting national influence over interoperability, security, and long-term system evolution. Participation in international AI standardization processes remains low, further reducing visibility of local needs in global technical frameworks.

13. Policy interventions should complement AI adoption with investments in research and technical capacity. Consistent investment helps move beyond pilot initiatives towards scalable and interoperable systems. Complementary and mutually reinforcing public and private investment play a catalytic role in enabling research, experimentation, and standardization engagement, while private investment is essential for commercialization and scale. Weak coordination between these channels leads to fragmented ecosystems and limited global competitiveness.

14. AI readiness encompasses not only technological capacity but also the ability to participate effectively in the development, adoption, and implementation of international technical standards. Active engagement in standardization on AI, data, and emerging technologies, contributions to technical specifications, reference architectures, and evaluation frameworks, and alignment between national AI strategies and international interoperability requirements.

Report Structure

This report is structured as such: The *Introduction* part serves as the overview of the report; *ITU AI Readiness Basic Framework* section is a summary of the general framework for public use, containing the explanation of the key dimensions and the indices; the *Structural Approach* section complements the basic framework with more actionable metrics for domain experts; *AI Readiness Gap Analysis* section identifies the gaps notified during the studies in standards, policy and implementation area; *AI Readiness Framework engagement* section introduces the ITU AI Readiness Enablement Toolkit design, its requirements, and the ITU AI Readiness Challenge, which serve as the engagement bridge with general users; the *Future Work* section outlines the work that will be accomplished in 2026-2027 including the expansion of the plugfest projects, launch of the ITU AI-RE Toolkit, ITU AI Readiness challenge, development of AI Readiness standards, and expansion of ITU AI for Good Sandbox Network.

2. ITU AI Readiness Basic Framework

Dimensions	Indices	Metrics
Data	Data Accessibility	Open Datasets and models
		Data collection source
	Data service capability	Data Quality Metrics
		Data Representativeness and Diversity Metrics
		Data Labelling Capacity Metrics
	Data Governance	Bias Detection and Mitigation
		Fairness and Accountability Safeguards
	Data Interoperability	Standard data formats
Digital Infra-structure	Connectivity	Fixed-broadband subscriptions
		Fixed Broadband Download Speed
		Mobile-cellular subscriptions
		Mobile Download Speed
		Network Coverage
	Computing Capacity	Compute availability per capita
		Number of Data Centres
		Energy Supply Per Capita
	Device	IoT Market Size
		Access to chipset
		Robotics platform adoption
		Smart sensors deployment
		Usability
	Automation	Levels of automation [Y.3173]
	Access to AI	Location of AI (edge/cloud)
		Contextualization level

(continued)

Dimensions	Indices	Metrics
Digital Skills	Education	Number of STEM Graduates
		Access to AI courses
	Digital literacy	National ICT Skills Level
		AI skills level
	AI application development	AI application development trainings
Innovation Ecosystem	Standards	Data standards
		AI pipeline standards [ITU-T Y.3172]
		Benchmarking standards
		Energy management standards
		Vertical Applications standards
	Open source	Engagement/adoption of open-source projects/models
	R&D	R&D Investment as a Percentage of GDP
		Number of AI Publications
	Investment	Annual public investment in AI
		Private investment/VC in AI startups
	AI Technology Source	Domain-wise AI Technology exported
		Domain-wise AI Technology imported
AI Policy	AI Policy and Regulation	National AI strategies
		National ethics framework
		AI policy tools
	Regulatory Quality	AI regulation implementation in the country
		Flexibility
		Sandbox
	Implementation	Implementation guidelines and priorities
		Supervision guidelines
		AI content guidelines

Data

The Data dimension evaluates foundational data readiness for AI development and deployment. Four core dimensions are selected – Data Accessibility, Data Service Capability, Data Governance, and Data Interoperability- as they collectively address the key aspects of AI data. Accessibility ensures data exists and can be acquired, Capability determines whether data can be transformed into AI-ready inputs, and Governance guarantees data is used ethically, securely, and sustainably, and Interoperability ensures compatibility. This four-in-one mirrors the journey from raw data to trustworthy AI solutions: without accessible data, projects cannot launch; without capability, data remains unusable; and without governance, AI adoption risks ethical failures or public rejection.

To operationalize the findings in section 1 above, and to develop data readiness metrics within the AI Readiness Framework, a set of actionable metrics is proposed, focusing on data quality, data labeling capacity, and bias and fairness risks.

1. Data Service Capability – Quality and Labeling Metrics

(a) Data Quality Metrics

These metrics assess whether datasets meet minimum requirements for AI training and deployment:

- Dataset completeness: proportion of datasets meeting minimum completeness thresholds
- Accuracy and consistency: error rates and internal consistency checks
- Timeliness and update frequency: frequency with which datasets are updated
- Multi-source diversity: diversity of data sources across institutions, regions, and data modalities

(b) Data Representativeness and Diversity Metrics

These metrics evaluate the extent to which datasets reflect real-world conditions:

- Demographic and geographic coverage: representation of population sub-groups and regions
- Sectoral coverage: coverage across priority sectors (e.g., education, agriculture, transportation)
- Local data share: proportion of datasets collected locally versus externally sourced or synthetic data

(c) Data Labeling Capacity Metrics

These metrics assess the ability to transform raw data into AI-ready training data:

- Availability of labeled datasets: proportion of datasets with usable labels
- Labeling quality: consistency or inter-annotator agreement in labeling processes
- Localization of labeling: extent to which datasets are labeled using local languages, contexts, and domain expertise
- Scalability of labeling processes: cost, time, and workforce required per labeling task

2. Data Governance – Bias and Fairness Metrics

(a) Bias Detection and Mitigation

These metrics assess safeguards against bias and discrimination at the data level:

- Existence of bias audits: whether datasets undergo bias or fairness assessments

- Bias documentation: availability of documentation describing known dataset limitations or biases
- Corrective mechanisms: processes to rebalance or refine datasets when bias is identified

(b) Fairness and Accountability Safeguards

These metrics connect data readiness with trustworthy and accountable AI deployment:

- Alignment with data governance frameworks: consistency with national or sectoral data governance policies
- Transparency mechanisms: disclosure of data provenance and labeling practices
- Monitoring and review: periodic reassessment of datasets used in deployed AI systems

Digital Infrastructure

The Digital Infrastructure dimension is a foundational element for the development and adoption of artificial intelligence, as it provides the essential physical and technological conditions for AI systems to be trained, deployed, and accessed. It is divided into five key dimensions: connectivity, which ensures fast and reliable data transmission across devices and platforms; Computing Capacity, which supplies the processing power required to run complex AI models; Device, which determines how widely AI applications can reach end users through smartphones, sensors, and IoT devices; Level of Automation; and Access to AI. Together, these capture the full stack of AI enablement – from core infrastructure to edge deployment – making this dimension critical for scaling and democratizing AI across sectors and populations.

Digital Skills

One of the major challenges for AI adoption in developing countries is the low level of general digital literacy and a shortage of specialized technical skills among the population. Universal digital literacy provides a foundation for the inclusive use of frontier technologies and AI systems. Widespread application of AI, simultaneously cultivating AI talent and vertical domain talent, while actively fostering their exchange and collaboration, holds strategic and critical importance. Under this dimension, there will be three indicators: education, digital literacy and AI application development.

Innovation Ecosystem

The Innovation Ecosystem assesses the broader environment that nurtures and accelerates the advancement and adoption of Artificial Intelligence. It focuses on the critical inputs and collaborative dynamics that transform research into tangible progress and practical applications. We examine six key dimensions: standards, Engagement in Open source, R&D, investment, GenAI Content, and AI Technology market size. These elements collectively capture a nation's capacity for pioneering research, its engagement in global knowledge sharing, and the financial commitment required to translate innovation into impactful AI solutions. A vibrant ecosystem requires not only cutting-edge research foundations but also active participation in open communities and sustained investment to bridge the gap between discovery and deployment. Weakness in any of these areas can significantly impede the pace and scale of AI-driven progress.

AI Policy

The AI Policy dimension evaluates the maturity of institutional frameworks essential for trustworthy and responsible AI ecosystems. We selected AI Policy and Regulation, Regulatory Quality, and Implementation as they collectively represent the indispensable triad for accountable AI systems.

3. Structural Approach

The ITU AI Readiness framework identified 6 fundamental factors, which set the foundation for the AI Readiness study. Under factors, 13 dimensions are derived from the Plugfest project reports, each of which is mapped to at least one factor.

The dimensions are chosen based on whether they fit some axis (e.g., X, Y) and whether it is possible to plot the progress of different levels of entities across these axes.

Indices are extracted and summarized based on the domain-specific metrics called out from the plugfest projects. Two types of indices (1) 0/1 filter, (2) weightage, which goes from 0-1, are designed. Weightages reflect the relative importance given to different metrics and dimensions in different countries or companies.

Metrics are domain-specific Key Performance Indicators (KPIs) under each dimension, which will be designed and reviewed by domain experts. Metrics are used to measure the output of the toolkit. The method or process of measuring a metric depends on the domain.

Users of the framework include countries, enterprises, Non-Governmental Organizations (NGOs), and other 3rd parties. The user can choose the level of self-assessment by applying the relevant set of indices.

Factors

The preliminary report of AI Readiness, published in May 2024, sets the foundation for the ITU AI Readiness study. The preliminary report identified six key readiness technical factors:

1. Data: Accessibility and quality of datasets for analysis of AI applications.

The availability of data is crucial in training, modelling, and applications of AI, irrespective of the domain. Data availability for analysis may be private or public. Metadata for private data may be published (e.g., data types and structures). However, public data, open for analysis by anyone, requires cleaning and anonymization to remove confidential or personal information.

2. Research: Collaboration between domain-specific and AI research communities.

Balancing the two main aspects of research, namely advancements in domain-specific research and advancements in AI research, requires collaboration between domain experts and AI researchers. Providing a platform for collaboration with experts from different realms of knowledge, facilitating cooperation, and the exchange of information among them is key to creating a sustainable ecosystem for AI-based innovation.

3. Deployment Support: Infrastructure and ecosystem readiness for AI deployment.

Two major categories of infrastructure are studied – physical infrastructure and communication infrastructure. Physical infrastructure elements play an important role in the integration and application of AI in data collection, aggregation – at the edge or core, training – federated or centralized, and in the application of AI and Machine Learning (AI/ML) inference using actuators. In addition, there is backend infrastructure, such as compute availability, storage availability, fiber/wireless availability for the last mile, and high-speed wide area network capabilities, which would democratize AI/ML solutions and create scalability for innovations.

4. Standards: Ensuring trust, interoperability, and compliance.

Interoperability and compliance with standards build trust. Secure standards lead to AI Readiness, as global participation and consensus decide whether pre-standard research could be adopted into the real world. Vendor ecosystems, including open source, are diverse in different domains of use cases. Adoption of AI-based solutions that involve humans, such as mobility inclusion, requires their trust and perception of using AI-based solutions.

5. Open source and Code: Enabling rapid adoption through an open developer ecosystem.

An energized third-party developer ecosystem not only fast-tracks adoption but also enables revenue generation. Developer ecosystem bootstraps reference implementations of algorithms, with baseline and open-source toolsets. Third-party applications, Application Programming Interfaces (API), and Software Development Kits (SDKs), along with crowd-sourced solutions, increase the generalizability of AI/ML solutions across regions and domains via transfer learning. Hardware implementations, especially open-source IoT boards, are evolving to host edge data processing.

6. Sandbox Environments: Platforms for AI experimentation and validation.

Many use cases require an experimental sandbox, creating experimental solutions, and validating them using experimental setups. While real-world data would imply a more reliable source of data and a realistic testing environment, not all scenarios could be encountered in the real world, especially when catastrophic events and related data are rare.

Dimensions

In this session, we will introduce the dimensions that are derived from the plugfest projects with examples of AI integration scenarios, so that a straightforward understanding can be provided. The evaluation could be done at different levels of the entity, including country, industry, and enterprise levels, based on needs. The evaluation of the status quo, gap analysis, and recommendations for users of each value could be provided accordingly.

Dimension 1: Data/model Marketplace

This dimension is derived from projects where the importance of exchanging data among partners and creating value out of it was realized. In the scenario where open data and models are available on the table, ontologies and connections within the data can be identified in the system; new ideas, business values, or concept notes could be developed based on the exchange.

This dimension aims to measure the creation of an ecosystem/environment for startups, business-to-business, or other types of value-providers to create services such as (un)structured data, expert knowledge base, and general platform, and monetize them.

This dimension helps in measuring the readiness of integrating AI to provide business value, especially for deriving value from existing unstructured data, models, and domain knowledge and business workflows. The more value that can be generated from adopting AI, the easier it is to use the AI techniques on a larger scale. Metrics such as "properties of the data and models", "properties of the marketplace ecosystem" could be used, so that the preparedness for the data and model marketplace could be evaluated.

This dimension is mapped to the factor of open data and open source due to the reliance on open-source data and models.

Dimension	Metrics
Data/model marketplace	• Properties of datasets (metadata: types structured/unstructured, number of datasets, volume, velocity, variety, quality). • Properties of models (metadata: types, size: number of parameters, performance: accuracy, ML/DL, i/o: data and inference, training dataset parameters). • Data collection sources (location: home/enterprise/public, heterogeneity: image/audio/video, number of sources, privacy/trust, synthetic/real world, type of source: streamed, e.g., mounted cameras, satellites, drones, field IoT reports) • Diligence metrics for the marketplace (license and participation agreement templates – trust). • Data marketplace: the number of data producers, data consumers, and agreements integrated in the digital data marketplace. The number of transactions in the data marketplace. • The number of open datasets and downloads from the data marketplace for such datasets. The number of global citations for the datasets. • Marketplace metrics, including the number of active participants and transaction frequencies. • Metrics for fairness and bias: safeguards to manage data bias or data quality risks. • Privacy and security – metrics to measure the levels of assurance, such as privacy-preserving mechanisms for using datasets, number of sources that can contribute operational data in a privacy-preserving manner. Authentication, Authorization mechanisms, and just-in-time deletion of private data. • Standards on personal data protection. • Data governance metrics, such as: the percentage of datasets under partner agreements (communities/NGOs/public agencies/private), license metrics (types of licenses). • Standards-compliant data formats: the amount of data that is available in a pre-specified standard, amount of data that is available in an interoperable manner. • Metrics and Properties of open-source models, such as openly published weights, or open-source models. E.g., the number of open-source models in different domains like coding and mathematics.

Dimension 2: Generated Content Marketplace

GenAI has been a heated discussion recently, and usually, the focus is on using AI to generate new content. Yet when studying the plugfest projects, one unique perspective came to attention: generating new datasets/models, so that they can be integrated into the new AI services and then be used/traded. New content could be generated for the purpose of AI services. Are we prepared to provide an ecosystem where new ideas can be generated by plugging in existing materials, connecting with other innovations, and being turned into new services?

In this dimension, new (innovations) Intellectual Property (IP) may be created as part of creative sectors. New datasets and models may also be created, which may be used to create new services that use these datasets and models. We aim to measure the ease of creation of new services using AI.

This generated content marketplace should allow users to generate new services based on the resources plugged into the ecosystem, which are IP databases, multi-modal content, arXiv papers, open source models, and codes. Metrics such as the properties of this ecosystem, parameters of the datasets or models in the ecosystem, the interoperability among resources when trying to generate new content, and the ability to detect hallucination could be considered.

This dimension is mapped to the open data and open-source factor, as the marketplace relies on open data and open models to generate new content.

Dimension	Metrics
Generated content marketplace	• Parameters of the ecosystem, which include datasets, models for content generation (including open-source models), pluggability of new services for content generation, and trading/monetization. • Availability of guardrails for hallucinations and ethical content. • Support for multi-modal content. • Evaluation techniques for fake detection. • Customization mechanisms for regional content.

Dimension 3: Cross-domain correlation analysis

AI could be adopted in various domains. In scenarios where the co-benefits of the integration of AI, such as economic, social inclusion, and environmental benefits, are to be studied, cross-domain correlation analysis is needed. If AI is integrated into one workflow in some steps, it could be adopted in similar ways or modified manner in other workflows. This dimension aims to measure cross-domain correlation in integrating AI. The metrics here would find similarities and patterns in different domain workflows and opportunities for integrating AI.

The availability and quality of the published domain report, domain-specific models, and KPIs can serve as metrics for this dimension, and the benefits due to correlation analysis will also be evaluated.

This dimension is mapped to open data and standards. Cross-domain analysis requires a large amount of data, reports, use cases, and information about domains so that correlations among domains can be established. To compare the workflow of the domain, it is required to have a standardized representation of the workflow; thus, ITU standards have to be introduced.

Dimension	Metrics
Cross-domain correlation analysis	• The existence of an integrated workflow including prevention (e.g., risk analysis), detection (e.g., monitoring), response (e.g., resource management), and impact analysis (e.g., post facto analysis). Based on: o Status of the application domains using available data (e.g., from published reports) and regional Readiness parameters (if available). o Domain-specific workflows/models, e.g., fire propagation and detection models. o Domain-specific KPIs (e.g., reduction in the burnt area). • Availability of representation schemes for infrastructure for deployment (e.g., geographic distribution, geographic information system, ArcGIS-based representation). Including city building plan + weather info as input (e.g., CityGML). • Benefits due to correlation, such as: o Cycle time reduction via integrated workflows (e.g., time delay between detection to response). o Coverage (in terms of area covered) and scale (in terms of deployments).

Dimension 4: Contextualization and Regional Impact

When adopting AI solutions that are originally coming from other regions/ other domains, it is observed that contextualization and adaptation are needed. This includes the choice of datasets, models, research, guidelines, toolsets, and standards developed with regional inputs and developed regionally.

Enlarging the regional impact of the AI solution to a larger scale should also be captured. What are the differences between local solutions with those in other parts of the world? What might be the gaps to bridge and to improve?

This dimension handles indigenous solutions, contextualization of the overall solutions with regional inputs for maximizing the impact on communities and the region, and the adoption level of regional solutions on a larger scale.

Metrics to evaluate this dimension are numbers and quality of locally collected data, innovation and patents, including models, toolsets, AI solutions, research, guidelines, number of users for local services, and the adoption level of local services in other markets.

This dimension is mapped to open data, research, and deployment. Local innovation involves large amounts of data and research efforts. The contextualization will facilitate the deployment of AI integration.

Dimension	Metrics

Contextualization & Regional Impact	• Number/quality of regionally developed patents/technology components/solutions.
	  o Datasets, models, research, guidelines, toolsets, and standards developed with regional inputs and developed regionally.
	  o Availability of structured and accessible local datasets for training AI models.
	• Number of users of indigenous services.
	• Customizations for regional applications.
	  o Generalized vs. contextualized solutions (with local inputs).
	  o Gaps for local industry and researchers to develop and contribute with respect to global components and technologies.
	  o Analysis of patterns in customizations to derive potential points for customization, e.g., model training based on regional skin patterns in dermatology.
	• Adoption and scaling of local technologies in other markets.
	  o Mapping the technology adoption in different domains and entities to regionally developed components.
	  o Level of locally developed technologies in domain-wise end-to-end solutions currently deployed.
	  o Level of contribution to global standards.
	• Knowledge products: number of localized standard operating procedures, after-action reviews, and "hybrid AI + traditional practice" playbooks published.
	• Cultural diversity brought by regional inputs.
	• Adoption of best practices across regions: the number of new regions adopting an AI-based workflow and best practices from other regions.
	• Representation of sub-groups in the dataset for fine-tuning, prompt-tuning or evaluations.

Dimension 5: Level of Integration of AI in Workflows

AI is now widely used in different industries, such as manufacturing, education, agriculture, international trade, and so on. It can be used to detect wildfires and provide alarms for the local population. It can also be used to smooth the logistical processes when international trade is made among business partners. But how well is the AI integration in the workflow, and how many benefits does the AI provide? In this dimension, AI is seen as a tool used to optimize different domain workflows.

This dimension can be measured by efficiency, redundancy, and other metrics of AI integration. This dimension helps in inferring recommendations for improved integration of AI in workflows may be produced.

Some gaps in interoperability are noticed. The use cases, represented as workflows, would integrate AI at various points, where 3rd party APIs as tools will be called out. A standardized interface to host the APIs will be needed. The optimization and design of APIs for tool usage in the workflow by the models will be studied based on the analysis.

This dimension is mapped to the Standards factor. Integrating different AI techniques into various domain workflows needs standards to guarantee their interoperability.

Dimension	Metrics
Level of integration of AI in workflows	• Level of automation achieved by integrating AI. • Benefits achieved by integrating AI (which will reflect the usefulness of AI). • Time/energy saved. • Efficiency. o Containment time: average days with active direct combat per operation (e.g., fire management in 2024 in Pantanal (Brazil) uses an average of 31 days). • Redundancy. • Scalability of AI techniques. • The level of AI integration in workflows. o Design level. o Coding level/implementation level. o Evaluation/testing level. • Discovery and the level of suitability of the model and dataset for the workflow. • Number of AI services that can be provided for the same set of resources, such as data and funding. • Improvement in the quality of services by integrating AI. • The extension of scope and range of services (without compromising the quality and requirements). • Cost (demonstrable reduction in cost for similar services).

Dimension 6: Human Interface

AI solutions exist not only in the backend, but also in the front end. In scenarios where AI solutions need to interact with human beings, sometimes with special needs, the evaluation of human interfaces is needed. How accessible are the interfaces for users, and do they contain multi-modal content? Is the coverage of AI everywhere? What local languages are available for users? Is the last-mile coverage of the AI pain point being solved?

In this dimension, we aim to measure the use of AI for ease of human interaction with systems/workflows. The measurement considers various devices such as chatbots, robots, and other channels of interaction.

The availability and penetration of the AI-supported services in devices and local languages, the ease of interaction for people with special needs, are examples of metrics to be evaluated.

This dimension is mapped to Standards and deployment factors. Integrating AI techniques into various human interfaces requires standardized rules. The availability of local language in AI solutions and the existence of AI solutions in near-human devices will facilitate the AI adoption and usage scaling.

Dimension	Metrics
Human interface	• Availability/Penetration of applications with AI AI-ready human interface. • Note – "AI-ready human interface" is one that makes interaction easier (using AI) between humans and the applications. E.g., sentence completion using AI in messaging applications. • Availability of AI models in the local language used in the human interface. • Interaction channels and devices (e.g., mobiles, kiosks, smart panels, sensors, wearables, smart home devices, etc.) for humans (and robots) to interact with AI-integrated applications. • Usability level of AI-integrated applications and services. • The level of cultural sensitivity applied to human interfaces, e.g., the ability to detect an appropriate joke in the local language context. • The level of safety integrated into the human interface, which, while satisfying the application requirements, also considers responsible interaction with humans. For example, the level of safeguards on destructive practices and steps while applying the inference or different nodes in the AI pipeline.[1]

Dimension 7: Strategy Alignment

Adopting AI on a large scale is never an easy task; this is where "alignment" plays a role. Some AI integration scenarios have country-level or even bigger plans and require clear instructions from the higher level to the executive level. Not limited to the technical requirements, coordination with other departments is needed. In the case of AI diabetic retinopathy detection, the high-level strategy from the central authority is conveyed to medical experts, side-by-side with the help of bioengineering experts, AI scientists, logistics... Miscommunication in any step would slow down the process.

This dimension describes the overall gains of coordination for AI integration strategies across distributed entities, including stakeholders such as industry, academia, and government entities, and domains such as logistics, transport, healthcare, etc.

The overall steps of intent distribution involve:

- Step 0: The actors are the intent service provider (who will host the intent management service) and the intent service user (who will utilize the service provider's service).
- Step 1: A hosting entity registers a sub-intent with a national or international intent service provider.
- Step 2: A high-level intent is issued by a potential intent service user.
- Step 3: The intent service provider would decompose the intent into sub-intents and assign them to corresponding hosting entities.
- Step 4: the assigned hosting entities would complete the tasks and return the results to the intent service provider.
- Step 5: The intent service user would get a collated result from the intent service provider.

The interoperability gap here appears as the intent description from the intent service user to the service providers. A standard template for tasks should be developed.

Measuring effective coordination, relevance of results, identification, and achievement of specific requirements is a challenge for multi-level intent orchestration. Metrics for the evaluation are

1 ITU-T Y.3172 : Architectural framework for machine learning in future networks including IMT-2020

the coordination level between service users and providers, availability and numbers of actors mapping to intents, and time delay in each step.

This dimension is mapped to the Standards factor due to the need for a standardized format for intent coordination.

Dimension	Metrics
Strategy align-ment	NOTE 1) The coordination starts at top top-level intent/vision. (e.g., nutrition level among school children). 2) The top-level intent/vision is broken down into sub-tasks (e.g., service providers are schools and education departments, health departments, and meal vendors). 3) These subtasks are given to different service providers. 4) Finally, when all the sub-tasks are completed, then the top-level intent is satisfied. (e.g., after tracking for 4 quarters, the nutrition level among school-going children rises above xxx). • Level of coordination between service providers o communication channel availability o vertical coordination o coverage of services • Time delay in solution delivery, considering the coordination o Time-to-go-live: from design to deployment. For example, the average days from onboarding a new territory for fire management to rolling out the first operational alerts. • Top-level intent/vision representations (formats) and break down into task/domain-specific representations. • Standards on the interoperability of AI systems in public administrations and the private sector. • Availability of service providers and their capabilities mapped to the top-level intent/vision, and the number and profile of the service providers. • The domains impacted by the top-level intent/vision. • The correlation (and cobenefits) by analysing the relation between the higher-level intents and the sub-tasks (e.g., a top-level intent on nutrition may trigger sub-intents on schools and meal providers, etc.) This includes the level of alignment (1) between the institutions and industries at the country level and (2) international (e.g., across the European Union or the African Union). • International cooperation and outreach: metrics related to international liaisons and agreements on AI solutions.

Dimension 8: Collaboration with AI

Humans are not only using the output of AI solutions, but also dynamically interacting and shaping the output of AI solutions synchronously. In the scenario of disaster management, traditional and indigenous wisdoms are consulted to build AI models, showing the valuable contribution of local communities. In this dimension, we measure the level of collaboration between humans and AI in various workflows.

Metrics for this dimension could include the prompting effort to the AI model and the quality of the AI-generated component in the output, as well as the comparison of the AI co-created output with baseline solutions on correctness (pure AI-generated solution and pure human work).

This dimension is mapped to the open-source and Standards factor. The collaboration between humans and AI via different interfaces requires standards to bridge the gap. The collaboration may require both open-source and closed-source models.

Dimension	Metrics
Collaboration with AI	• Ease of collaboration and co-creation (human collaboration is used in the process of inference itself). • Ability to integrate traditional knowledge into AI-based solutions (to create hybrid solutions, e.g., both AI models and traditional knowledge used for inference). • The value added by the exchange between humans and AI. The final inference is decided collaboratively by humans and AI. E.g., In a text generation system, AI is used to produce drafts of emails or text content, and the suggestions from a human user are incorporated to make the final output. • The prompting effort (overhead) vs. the quality of the AI-generated component in the output. o Effectiveness of user feedback integration loops. o trust metrics in AI-generated components (e.g., availability and number of models with expert-validated results). • The comparison of AI co-created output with baseline (baseline may be only AI with no human-in-the-loop, or only human with no AI assistance, and comparison may be with respect to parameters such as correctness of output. • Incident Action Plans (IAPs) supported, e.g., number of IAPs produced with AI techniques inputs per fire season. • Finetuning and optimization efforts vs. results from the model in case of open-source models compared with closed models.

Dimension 9: Impacts of Humans in AI Integration

It is humans who build the scenario, train the model, and use AI. When studying the plugfest reports, the importance of human impact on AI integration is noticed.

There are scenarios when AI decisions are referred back to domain experts to guarantee the accountability of the decision. There are also cases where domain experts and AI practitioners are cooperating and training AI to identify diseases, so that doctors in the hospital can save their time in identifying healthy cases and concentrate on those that are pruned to be diagnosed.

Humans, including AI experts, domain experts, and general users, are shaping AI by modeling, training, fine-tuning, and adopting AI. Without human impact, AI integration cannot maximize its potential.

This dimension has 3 parts. (1) AI experts: This part measures the impact of human experts on AI in terms of model training, labelling, evaluating model output, etc.; (2) Domain experts: the benefits and tradeoffs on humans from integrating AI; And lastly, (3) General users: the awareness level in AI integration. Training and awareness to use AI-integrated services and solutions, along with region-specific analysis, are measured.

To measure the impact of humans in AI integration, the current skilling programme of AI and domain experts, the intersection between the two groups, the awareness and the adoption level of general users will be considered, and thus the corresponding metrics, such as the level of expertise (credentials and time of experiences) in each domain, overlap of the knowledge, ease of onboarding to AI solutions could be used to measure the level of benefits that can be reaped from human impact. Based on the evaluation results, a relevant capacity-building plan could be carried out for AI experts, domain experts, and general users to improve the understanding of domain knowledge and AI techniques and raise awareness of using AI techniques in workflows.

This dimension could be mapped with the Sandbox factor, as scenarios generated, training materials, and research papers for higher intersection among the three types of users could be used in the sandbox.

Dimension	Metrics
Impact of Humans in AI Integration	• Skill distribution and skill levels in technologies (AI, Cloud, Internet of Things (IoTs), 5G). o number of academic programmes or certified training courses in AI. • Intersection of domain experts who are aware of and using AI techniques on one side and AI practitioners aware of domain-specific techniques/needs on the other side. o number of experts trained per year on AI-supported workflows (including women/indigenous participation share). • Ecosystem readiness/ability to develop AI skills and AI talents. o skills gap analysis: an important output of the framework would be to identify what skills are currently lacking.
	• AI awareness and adoption levels among end-users. o inference-to-action latency: average time from first inference to human action. E.g., Detection-to-dispatch latency for fire brigades: average time from first detection to brigade mobilization. o decisions on human actions based on AI alerts: number and percentage of incidents where humans decided to act based on AI-based alerts. e.g., Fire brigades' decision to act following AI-based alerts. o number of citizens trained with AI-related skills – effectiveness of AI training programs.

Dimension 10: AI & Policies

To adopt AI and deeply integrate AI in our lives, policies play a key role.

This dimension is two-fold. It measures the ability of decision makers to experiment, extrapolate, and review the policy impacts on the ground by using AI, and have policies ready to enable smooth AI integration in various verticals. This dimension is useful for policymakers and regulators as it handles (1) the horizontal generic policies that apply to multiple verticals, for example, the privacy rules of usage of data in model training in healthcare. And (2) vertical-specific policies that apply to only a single domain.

"AI for Policies" aims to create a virtual sandbox for policy and regulatory interventions with a simulated timeline and predict outputs from intended interventions and desired results. The gaps between the target and achievement milestones are depicted, and actionable recommendations for bridging the gaps are given.

In terms of policies for smooth AI integration, new policies may be recommended, related to new sources of data: datasets, provenance of data, models, and risk assessment via Sandbox/Policy/SINK as per ITU-T Y.3172.

For "AI for Policies", metrics to evaluate are the availability of the sandbox with simulated scenarios for policy extrapolation and domain-specific policy evaluation mechanisms. For "Policies for AI", the availability of policies about the AI lifecycle, which may include design/creation, training, deployment, upgrade, and demission of the AI models, datasets, and solutions. Policies for nodes in AI pipelines, identified in ITU-T Y.3172, and policies that are granular to the level of country/enterprise will be considered.

This dimension is mapped to Sandbox and deployment factors. Domain-specific policies will be tested in the sandbox with a simulated timeline to see their potential impact on the ground. Policies regulate AI and pave the way for better AI integration on the ground.

Dimension	Metrics
AI & Policy	• Sandbox for evaluating domain-specific policies (related to the vertical applications). o Availability of vertical sandboxes, which evaluate the performance of models in domain-specific applications. E.g., 5G sandbox which evaluates the AI integration in 5G. o Domain-specific policies and evaluation mechanisms for AI-integrated solutions. E.g., metrics for measuring the effectiveness and accuracy of diabetic retinopathy detection using AI. o Extrapolation mechanisms for policy and regulatory interventions with a simulated timeline and predict outputs from intended interventions and desired results. The measured impacts of policies on AI solutions may include the usefulness/effectiveness/completeness of the policies on the domain workflows and the AI artifacts (datasets, models, AI integrated services). • Horizontal generic policies that apply to multiple verticals. o Availability of policies on. – sources of data. – AI lifecycle (design, training, deploy, upgrade/manage, decommission). – explainability, trustability, and reliability. – E.g., similar to ITU-T Y.3172, which applies to ML pipelines across verticals, including AI lifecycle management somewhere (from model training, deployment to decommissioning). o Number of instances where AI enables community Benefit – the inference includes the local citizens' inputs and key insights into decision-making. o Availability of policies for safe, responsible, and reliable operation of AI while satisfying the application-specific requirements. o Metrics to measure explainability and transparency so that policies for accountability of outcomes could be framed. This includes accountability of final decisions, which remains with humans.[2] o Number of Policy inputs from case briefs: number of published briefs/case studies turning operational lessons into governance lessons (e.g., triggers, thresholds, data flows). o Metrics from different types of audits, e.g., models, datasets, and toolsets, may be audited for bias, safety, explainability, etc., e.g., Presence of a national AI ethics committee or mechanism for standards on AI ethics. o Data sovereignty metrics: policies related to ownership and movement of data.

Dimension 11: AI for Inclusion

Simulation in virtual world, GenAI, Edge AI, and techniques on wearable techniques appear to be the key focuses from the plugfest projects. Researchers and scientists are using simulators to create sandboxes where people with mobility difficulties may use AI techniques to bridge the gap; GenAI can be used to generate scenarios for validation or human-like avatars alongside sign languages for easier communication; Edge AI and techniques on wearable devices can be used for general users to have AI embedded in every corner of their lives.

[2] Artificial Intelligence Assurance Framework | Digital Territory

This dimension aims to study bridging the divides in society using AI. This includes accessibility, gender divide, and other forms of inclusion. Measure the use of AI techniques for people with special needs. This dimension measures the availability of 1) simulations, 2) distributed, large-scale edge AI, and 3) the level of usage of GenAI for bridging the inclusion and gaps.

The first sub-dimension is the use of digital twins in a virtual world; metrics for this sub-dimension include types and numbers of sensors, the availability of sandbox environments, and the capability of simulators with respect to the real application. Researchers might use a sandbox environment to test mobility-related AI solutions. The interoperability gap is noticed as some domain-specific AI solutions in real practice need strict standards of evaluation, such as safety level standards for AI in mobility scenarios. The adoption of AI in such domains needs careful evaluation along with experts.

The second sub-dimension is the use of edge AI, which aims to evaluate the closeness to end users. In scenarios such as AI solutions for accessibility, applications are built with 2 steps in mind: upstream tasks that are more generic in nature (e.g., speech to text, text to speech, image to text) and the downstream tasks that are specific to domains (e.g., distance estimation and guidance to objects, and navigation). For the sake of latency requirements, the deployment of the downstream tasks is usually done at the edge, while the generic task is modeled and validated at the cloud.

The metrics to evaluate the usage of edge AI are availability and capability of hardware at the edge, connectivity, and privacy of the data.

The last sub-dimension is GenAI, aiming to measure accessibility for general users. GenAI can be used not only to generate training scenarios in the sandbox but also to mitigate the communication gap between users with special needs via sign languages and generated human avatars or other formats.

Metrics for GenAI that are used to bridge the divides include the number of sign languages supported, the use of language repository from the global baseline model, availability of customization in presentation, multi-modal and data support, the coverage of beneficiaries, GenAI capabilities (customization of the avatar, privacy protection).

This dimension is mapped with the Sandbox factor due to its large involvement of sandbox in the training phase.

Dimension	Metrics
AI for inclusion	• Metrics related to digital twins: o The number of sandbox environments. o Capability of simulators with respect to the application. o Standard integration methods towards the real world. o The types of sensors (laser, cameras, gyroscope, obstacle avoidance). • Metrics related to edge AI (closeness to end users): o Use of agents for the user interface, e.g., question answers on the domain, such as navigation. o Hardware availability (affordability) at the edge. Description of Hardware at the edge (as against the cloud). Capability at the edge (Central Processing Unit, Graphics Processing Unit, memories, network). o Connectivity: 5G, to connect the edge to the cloud, helps in data transfer, model transfer, inference, etc. – Coverage (geographical, vs. coverage holes). o Privacy: Standalone Edge (with no data leakage to the cloud) vs. Cloud (with data transferred to the cloud). o Latency vs. compute tradeoff and energy tradeoff. • Metrics related to GenAI: accessibility for general users. o Use of language repository created from baseline global vocabulary but including local grammar, dialects, and context for interaction (e.g., banking, school, hospital, etc.). o Regional languages may be collected into the repository in the form of “branches” (using GitHub terminology), using the same template as the main branch. – Offline/local-language access: the number of communities with offline alert kits (radio/voice prompts/regional language user experience) for natural disasters. – Support for regional languages in AI use cases, cultural diversity standards on linguistic and cultural inclusion (taking local languages into account in AI systems). o Customizations in presentations (e.g., avatar appearances) can be implemented. o Different types of data can be processed (multi-modal, e.g., audio, video, images) using specific standard representations (e.g., audio streams, MPEG files embedded). o Number of users/beneficiaries. – e.g., Availability of public audio-video material with sign language interpretation. – e.g., Number of sign languages supported. o Characteristics of the avatar (e.g., skin colour). o Multi-modal support (e.g., written -> avatar, audio -> avatar. sign language -> sign language). o Gaps in the form of new vocabulary, skeleton key points, and multi-modal materials. o Mapping to the common standard intermediate representation. – Skeleton and Key points for avatar generation and data collection. – Language templates. (e.g., noun – verb – subject).

(continued)

Dimension	Metrics
	o Generative AI capabilities (for avatar generation). – Customizations (e.g., skins). – Minimization of data transfer (e.g., only the key points are transferred). – Privacy protection (e.g., only the key points are transferred to the cloud, not the users' images). o Women's leadership and indigenous leadership in AI, e.g., metrics such as the number of women-led institutions in the country and international level. o Considering social inclusion scenarios and corresponding metrics. E.g., the impact of AI algorithms and datasets on social inclusion decisions, such as social security. o Generative AI capabilities (for avatar generation). – Customizations (e.g., skins). – Minimization of data transfer (e.g., only the key points are transferred). – Privacy protection (e.g., only the key points are transferred to the cloud, not the users' images). o Women's leadership and indigenous leadership in AI, e.g., metrics such as the number of women-led institutions in the country and international level. o Considering social inclusion scenarios and corresponding metrics. E.g., the impact of AI algorithms and datasets on social inclusion decisions, such as social security.

Dimension 12: Granular Priorities

For whoever designs and aims to integrate AI techniques, the granular priorities of the user need to be considered. Some regions have a clear focus on agriculture, trying to improve the quality and quantity of maize and crops by designing the best use of pesticides, while others might focus on diabetes and relevant complications, which is the second biggest health problem for the local communities. Each scenario has its own focus; thus, when integrating AI or adopting general AI solutions, it is important to accommodate the local needs and customize the model to the context so that the most appropriate adoption is used.

In this dimension, we measure the availability of granular priorities in adapting AI-based solutions to a subset of users. Examples of metrics for this dimension are the following: the availability of granular priorities that could be clearly mapped to broader solutions, the existence of an organizational structure for the evolution of granular priorities, and a fine-tuned model that could reflect the priorities.

This dimension is mapped to the deployment factor, as understanding the local needs and contextualizing and fine-tuning AI solutions contribute to deploying new models.

Dimension	Metrics
Granular priorities	NOTE – regional priorities are applied to global workflows and models to enable customized workflows reflecting regional, granular priorities. • Regional/domain-specific priorities/focus. • Finetuned workflow and models. • The existence of an organizational structure for the evolution of granular priorities. • The evolution of downstream customized (finetuned) workflows derived from global workflows/models/solutions.

Dimension 13: Digital Infrastructure

The application of AI cannot be realized without the help of digital infrastructure on the ground. The best distribution of the devices (sensors, cameras...) in the region/country might influence the next step of AI deployment in society.

This dimension measures the availability of digital infrastructure, including devices, computing capability, connectivity, and energy.

To measure the digital infrastructure readiness on the ground, metrics could include the number, quality, and distribution of the AI-enabled scenario-specific sensors and nodes identified in ITU-T Y.3172; the number of digital services integrated in the overall solution; the number of digital infrastructures at the edge; and the energy consumption level.

This dimension is mapped to deployment, as it facilitates the adoption and implementation of AI in real practice.

Dimension	Metrics
Digital Infra-structure	• Number/quality/distribution of AI-enabled devices/sensors (ITU-T Y.3172 nodes: SRC, trained Models, SINKs, Sandboxes for validation of the models). o e.g., Number of AI-enabled imaging cameras. • Number of data centers and APIs integrated in the overall/national solution. • Number of digital services (integrated in the overall cloud cross-domain), including the number of entities linked together to provide an integrated solution. • The availability and utilization of edge clouds, compute infrastructures at the edge, connectivity to the edge (e.g., length of fiber deployment, connecting the edge servers). • Energy consumption at the edge, devices, and infrastructure (for AI and domain-specific). • Energy reliability and sustainable energy usage considerations (e.g., power usage effectiveness related to computing for AI pipelines). • Energy consumption for (a) compute, including algorithms, (b) storage and retrieval from memory (c) processing packets in the network. • Uptime: percentage uptime of towers/cameras & control room feeds. • Area covered under AI services: including the area covered by sensors for data collection, and actuators for inference application. e.g., total coverage by AI-based fire management solutions each year: hectares of indigenous territories under monitoring. • Quality of Network connectivity (e.g., fiber, wireless, 5G coverage) [ITU-R M.2410].[3]

[3] Minimum requirements related to technical performance for IMT-2020 radio interface(s)

4. AI Readiness Gap Analysis

The ITU AI Readiness Project adopts a unique bottom-up method of analysis. The ITU AI Readiness Project started the research with use case analysis, with an aim to find out the common patterns of AI solutions and the key common factors that contribute to the successful deployment of the AI techniques. A preliminary report was published in the first ITU AI Readiness workshop to call out these factors. For further studies, ITU AI Readiness Project launched the Plugfest in the second workshop in Riyadh. The Plugfest initiative collected regional projects that could influence the AI Readiness and could bring in local nuances. During the third workshop in Geneva, reports on the plugfest projects were given. We continued to follow the bottom-up approach and derived the 13 dimensions and metrics for the AI Readiness self-evaluation framework out of the plugfest project reports.

The design of the dimensions and the metrics went through several rounds of consultation with experts from academia, industry, and member states of different countries. Different types of inputs were received, including country-specific AI integration strategies, suggestions on dimensions and metrics, and indices that reflect the regional/country-level priorities. Suggestions and input to the AI Readiness Toolkit design were also received, including interest in running the pilot in the field and the performance and non-performance requirements for the toolkit. Standards gaps, consultation, and engagement strategies for AI Readiness were also identified and provided by member states.

After collecting and collating the analysis and inputs, three types of gaps were identified.

(1) Gaps in international standards

Standards gaps, such as data harmonization, generative AI, and energy readiness for AI, remain critical to be addressed.

Data Readiness: What are the quality levels of datasets? A need for a framework to standardize the steps for collecting and preprocessing multi-modal, heterogeneous data? Standard mechanisms for discoverability, reusability, and customization of datasets would further enhance the AI Readiness. Data Readiness gaps are also identified in our study of domain-specific data that is collected from services and application scenarios in different use cases. Data Readiness is assessed considering the operations done with the data and to the data. Addressing these gaps is important to enable AI integration in real-world scenarios.

GenAI Readiness: Generative AI is increasingly used in daily lives, yet fine-tuned generative AI tools (using local data or domain-specific data) are still not prevalent. How do we train generative AI, integrate it in the domain, and customize it for local context? How do we ensure a trustworthy generative AI model? GenAI Readiness standards would address these questions and contribute to overall AI Readiness framework.

Energy AI Readiness: AI models, especially large models, not only provide significant benefits but also consume increasing amounts of energy. What is "green" from a technical perspective? In which format should we count the environmental footprint for AI models? To align with long-term sustainable development, to rationally define the energy consumption of AI models for fair comparison, energy-related standards for AI are required for consideration. A systematic study of AI use cases in networks, while mapping the pipeline nodes for each use case, studies the questions below: what are the energy characteristics of the nodes? E.g., how often is data collected from the source? How much data is needed? Is the data real-time, or does it need storage? What are the uniform ways of classification of use cases based on the energy characteristics of the nodes? Standard metrics are needed to achieve Energy AI Readiness, which includes the energy characteristics of the AI pipeline, the energy cost of data collection, such as frequency, volume, and location of the data.

(2) Gaps in implementations

As the AI techniques evolve rapidly, it is important to keep the implementation capabilities ready to deploy AI solutions across regions. Our analysis of the plugfest projects and expert engagements shows that some countries achieve high efficiency of AI technology, yet due to the lack of supporting infrastructure, suggestions and predictions provided by AI could not be transformed into effective applications. Some countries have rich resources of local data and traditional practices. Yet the AI techniques are unable to be deployed due to the lack of digitalized datasets and models. Different levels of AI implementation exist, and an AI readiness toolkit based on a standard framework would be needed for inclusive progress of AI integrations in different domains.

(3) Gaps in policies

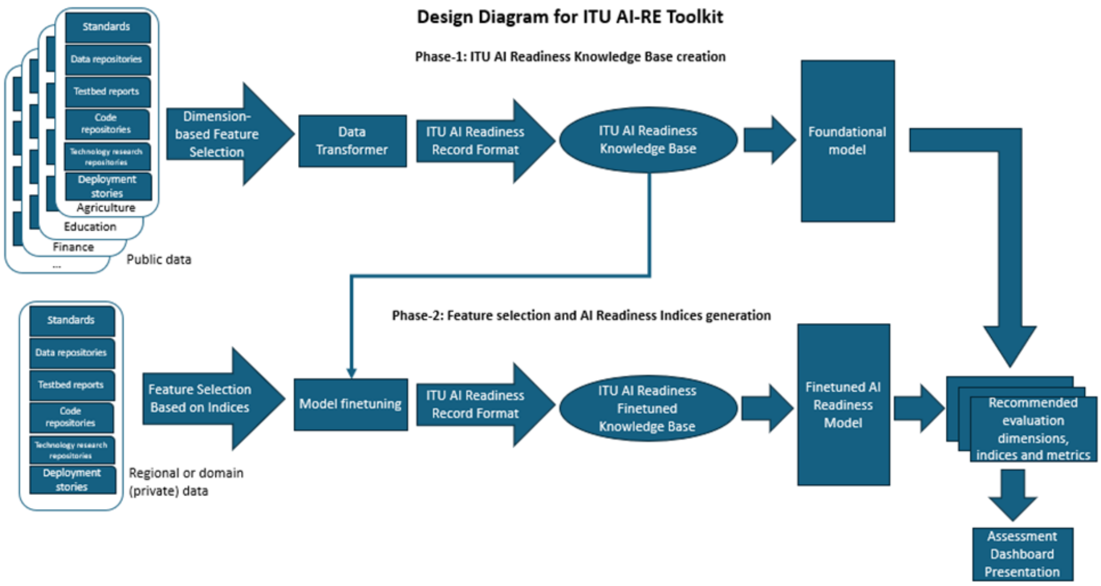
Approaches to framing AI policies, especially for cross-domain AI governance and policies for domain-specific integration of AI, differ between countries. In some countries, dedicated strategies have evolved for AI integration and related fields for AI governance. Yet in some resource-constrained countries, national strategies and policies might be derived from different ministries and might need policy alignment. A standardized framework would help to identify the gaps and provide customized and actionable recommendations for references.

5. AI Readiness Framework Engagement

This section will explain the method for further stakeholder engagement with AI Readiness Enablement toolkit and AI Readiness Challenge.

AI Readiness Toolkit

A dynamic ITU AI Readiness Enablement toolkit (ITU AI-RE Toolkit) is designed to allow users to (1) self-evaluate the status quo readiness level of AI integration and (2) receive customized recommendations, gap analysis, and actionable plans for the future. Users of the toolkit (regions, countries, enterprises, NGOs, and other 3rd parties.



Users provide unstructured data, such as domain-specific PDF reports, deployment stories, best practices, and use cases, into the toolkit as input. The materials will be transformed into knowledge records using the standardized ITU AI Readiness record format. The knowledge records will be stored in the ITU AI Readiness Knowledge Base, which serves as the brain and the foundation of the toolkit. A foundational model will be generated based on the knowledge base materials. Users may use the foundational model to evaluate the performance of AI techniques and recommendations, and gap analysis. Users may also input domain-specific data and user priorities to fine-tune the knowledge base and receive a fine-tuned model and customized analysis results.

Users may attach indices to address granular priorities. There could be two types of indices: (1) a 0/1 filter, (2) a weightage which goes from 0-1. Weightages reflect the relative importance given to different metrics and dimensions in different countries or companies. The indices could be attached to metrics or dimensions in general.

The weightages allow the model to focus on improvement areas with respect to the current status quo. To guarantee a transparent and fair comparison, ranking of self-assessment based on weightages would need (1) publishing the weightages in a policy paper, (2) results from the baseline foundational model.

An example of the index could be "job market conservation". By applying the index, the status quo analysis of the assessment under "Impact of Humans in AI Integration" might specifically reflect metrics such as hours of work saved by adopting AI. Similarly, the recommendations would be aligned with the index applied.

Indices are provided based on country-level needs. Policies and filters are applied on the inferences from the model to reflect country-level priorities and weightages. However, the input materials and private data come from enterprise-level users; thus, the recommendation or status quo analysis from the framework reflects the characteristics and the level of the input material.

Examples of indices are training and awareness-raising activities to create a pool of local AI skills; the development of appropriate legal and ethical frameworks; the exploration of international partnerships for technology transfer; and roadmaps for 2025-2035.

Different levels of AI skills could serve as another example of indices. AI literacy level, which ensures that all members of society, including children, the elderly, and everyday technology users have a basic understanding of AI concepts, could serve as the basic level of AI skills. Specialized AI skills, which indicate the capability of contributing to research, innovation, and AI solution development, could be at the medium level. Workforce upskilling, upskilling employees across all sectors to adopt and apply AI in daily tasks to enhance productivity, could be the highest level of AI skills. [4, 5]

Usage scenarios

There are several scenarios where the toolkit can be used.

User story-1: an entry-level country wants to self-assess the AI integration levels.

Considering the developing stature, the country would like to concentrate on basic domains such as agriculture and education. The assessment team from the country provides materials corresponding to the domains and the reports. The materials are handled in a trusted sandbox and analyzed by the ITU AI RE Toolkit. Two levels of output are produced. Foundation output (which uses the same set of metrics for all) and finetuned output (incorporates the country preferences).

User story-2: regulators want to use the ITU AI4G Sandbox to evaluate the impact of policies in the domain.

The ITU AI-RE Toolkit is hosted in the ITU AI4G Sandbox, where simulators are equipped. Following the recommendations of the ITU AI-RE Toolkit, the user wants to use a sandbox to evaluate and extrapolate the impact of a given policy on the ground. The user could pose "what if" questions and receive corresponding results by using the input data or those stored in the sandbox and adjusting the parameters of the simulators.

User story-3: iterative learning.

4 General Policy for the Use of AI – Final 30 Jul 2025.pdf
5 Sixth National Telecommunications Plan | Ministry Of Transport

An NGO wants to use the ITU AI-RE Toolkit to continuously self-evaluate the performance of AI techniques in a given field. The user could input the yearly report into the toolkit. The material will be captured and stored in the ITU AI Readiness Knowledge Base in the form of ITU Readiness Records. By iteratively inputting the material in the same domain, the user could receive updated results from the toolkit accordingly. The capability of learning iteratively ensures the evolution of the ITU AI-RE Toolkit.

Requirements

ITU-AI-RE-Toolkit-REQ-001: it is required that the toolkit output assessment results be displayed in a dashboard, based on the inputs from the user.

NOTE: this provides a status quo assessment.

ITU-AI-RE-Toolkit-REQ-002: it is required that the toolkit output recommendations, gap analysis.

NOTE: this provides improvement recommendations.

ITU-AI-RE-Toolkit-REQ-003: it is required that the toolkit recommend relevant dimensions, indices, and metrics based on the inputs from the user.

NOTE: this provides the feedback loop for better selection of dimensions, indices, and metrics for the user.

ITU-AI-RE-Toolkit-REQ-004: it is required that the maturity levels per dimension be evaluated.

ITU-AI-RE-Toolkit-REQ-005: it is required that the toolkit generate a finetuned AI model based on the foundational model and considering the regional inputs in the form of indices and corresponding weightages. The toolkit outputs the assessment with respect to the finetuned model and any suggested (new) metrics in this regard.

NOTE: finetuning is not a must. It is a user's choice to fine-tune the model. Finetuning depends on the resources and expertise of the user. For example, a country at entry level may not fine-tune the model initially, but rather use the foundational model, self-assess the output, and then decide to input specific indices for fine-tuning.

ITU-AI-RE-Toolkit-REQ-006: it is required that customizable dashboards for different stakeholders be implemented to show customized dimensions and aggregate views.

ITU-AI-RE-Toolkit-REQ-007: it is required that both private and public data can be used as input in the toolkit.

NOTE: multiple instances of the toolkit can exist at the same time, being connected to the ITU AI4G Sandbox in parallel (for KB records transfer only). Since these instances do not transfer data between each other, there are no data privacy concerns between the users.

ITU-AI-RE-Toolkit-REQ-008: it is required that the toolkit be modular and map to the ITU Readiness Framework dimensions, allowing selection, weighting, and customization of metrics aligned with regional priorities and sector-specific needs (e.g., telecom-specific metrics).

ITU-AI-RE-Toolkit-REQ-009: it is required that the data collection framework within the toolkit support multi-source metric inputs, such as quantitative metrics, qualitative survey-based metrics.

ITU-AI-RE-Toolkit-REQ-010: it is required that the toolkit verify 1) dataset metadata and provenance verification, 2) model training, tuning, versioning, and deployment lifecycle assessment, and 3) data privacy, security, and ethical compliance tracking consistent with international standards.

ITU-AI-RE-Toolkit-REQ-011: it is required that the toolkit use ontologies guided by experts to validate the data.

ITU-AI-RE-Toolkit-REQ-012: it is required that cross-dimensional correlation analytics are used to highlight interdependencies and compound gaps.

NOTE: compound gap refers to a combined or interconnected gap that arises when multiple dimensions interact and collectively result in a more overall deficiency. Rather than looking at gaps in isolation or within single dimensions (such as human impact, data marketplace, or AI and policies individually), a compound gap shows how the weaknesses in one area can amplify challenges in others.

ITU-AI-RE-Toolkit-REQ-013: it is required that the toolkit embed support for 1) language and cultural adaptation of assessment instruments, 2) inclusion metrics like accessibility, gender divide, and regional language support for AI interfaces, and 3) modular extension points for local-specific metrics and emerging AI capabilities.

NOTE: examples could be generative AI usage statistics.

ITU-AI-RE-Toolkit-REQ-014: it is required that the toolkit evaluate the maturity of collaboration with AI based on 1) AI-assisted co-creation, inference, and user-AI interaction quality across devices (chatbots, robotics), 2) multi-modal interaction support, including audio, text, sign language, and avatar representations, and 3) tracking of prompt quality, user feedback integration, and trust measures.

ITU-AI-RE-Toolkit-REQ-015: it is required that the knowledge base be hosted in the ITU AI4G Sandbox.

ITU-AI-RE-Toolkit-REQ-016: it is recommended that the toolkit interface with the ITU AI4G Sandbox with input data or simulated data in the sandbox.

NOTE: the sandbox could be used to 1) simulate the regulatory interventions and their impact on the ground; 2) evaluate domain-specific AI application performance in controlled settings; and 3) provide actionable gap analysis and scenario-based recommendations.

ITU-AI-RE-Toolkit-REQ-017: it is required that potential evolution of new dimensions be recommended by the toolkit as the Knowledge Base is scaled with respect to the diversity of the AI Readiness Records.

NOTE: this output provides the feedback loop intended to evolve the toolkit.

ITU-AI-RE-Toolkit-REQ-018: It is required that the toolkit input unstructured data formats in the form of documents and reports compiled from different domains, such as agriculture, education, health care, and natural disasters.

ITU-AI-RE-Toolkit-REQ-019: it is required that the unstructured data formats include case studies of the application of AI in the domain, specifically mapped to the different factors (open data, open source, research, standards, deployment, and sandbox)

NOTE: examples of unstructured data formats are ITU reports on standards, number/characteristics/metadata/statistics of open-source repositories, deployment studies for AI in specific domains, etc.

ITU-AI-RE-Toolkit-REQ-020: it is required that the toolkit learn iteratively and provide continuous user feedback.

ITU-AI-RE-Toolkit-REQ-021: it is required that composite scoring is calculated by aggregating weighted metrics within each dimension.

ITU-AI-RE-Toolkit-REQ-022: it is recommended that the unstructured data be captured in local languages.

ITU-AI-RE-Toolkit-REQ-023: it is required that the structured data formats include machine-readable properties of the application of AI in the domain. x

NOTE: examples of structured data formats are tabular representations of statistics on open-source repositories, deployment studies for AI in specific domains, etc.

ITU-AI-RE-Toolkit-REQ-024: it is required that the knowledge base uses a uniform/standard record format for the storage and processing of knowledge records.

NOTE: knowledge records are a processed form of data inputs. An example of a knowledge record format is vectorized data records.

ITU-AI-RE-Toolkit-REQ-025: it is required that the knowledge base be distributed, and the consumers of the knowledge base select, process the knowledge records from the distributed knowledge base based on user preferences/priorities.

NOTE: examples of user preferences/priorities are the choice of crops and pesticides in certain countries.

ITU-AI-RE-Toolkit-REQ-026: it is required that the toolkit be used locally, and data not be moved from the country of use.

NOTE: knowledge base is updated in the ITU AI4G Sandbox with AI Readiness Record formats.

ITU-AI-RE-Toolkit-REQ-027: it is required that local rules be respected for data privacy and governance.

NOTE: regular compliance reports would be produced for this by agents.

ITU-AI-RE-Toolkit-REQ-028: it is required that the only persistent records in the toolkit be knowledge base Records, and input materials be stored only temporarily during training and fine-tuning and be deleted immediately afterwards.

Mapping between Toolkit Requirements and Framework

Analysis of the metrics and the toolkit requirements is done to derive the mapping between requirements and metrics. Analysis of toolkit requirements shows 2 types of requirements:

1) Requirements that are related to the usability and functionality of the toolkit (e.g., customized recommendation, gap analysis, and status quo evaluation presented via dashboard)
2) Requirements related to the assessment of dimensions and metrics in the framework.

The Table below provides a mapping between toolkit requirements and the framework.

Toolkit Requirement	Mapping to the Framework
ITU-AI-RE-Tookit-REQ-001	Not applicable
ITU-AI-RE-Tookit-REQ-002	Not applicable
ITU-AI-RE-Tookit-REQ-003	Not applicable
ITU-AI-RE-Tookit-REQ-004	Not applicable

(continued)

Toolkit Requirement	Mapping to the Framework
ITU-AI-RE-Tookit-REQ-005	Granular priorities Data/model marketplace Contextualization & regional impact
ITU-AI-RE-Tookit-REQ-006	Not applicable
ITU-AI-RE-Tookit-REQ-007	granular/data/contextualization
ITU-AI-RE-Tookit-REQ-008	granular
ITU-AI-RE-Tookit-REQ-009	Data/model marketplace
ITU-AI-RE-Tookit-REQ-010	Data/model marketplace
ITU-AI-RE-Tookit-REQ-011	AI for Inclusion
ITU-AI-RE-Tookit-REQ-012	Cross-domain correlation analysis
ITU-AI-RE-Tookit-REQ-013	AI for Inclusion
ITU-AI-RE-Tookit-REQ-014	Collaboration with AI
ITU-AI-RE-Tookit-REQ-015	Not applicable
ITU-AI-RE-Tookit-REQ-016	Not applicable
ITU-AI-RE-Tookit-REQ-017	Granular priorities Contextualization & regional impact
ITU-AI-RE-Tookit-REQ-018	Data/model marketplace
ITU-AI-RE-Tookit-REQ-019	Data/model marketplace
ITU-AI-RE-Tookit-REQ-020	Not applicable
ITU-AI-RE-Tookit-REQ-021	Not applicable
ITU-AI-RE-Tookit-REQ-022	AI for Inclusion Data/model marketplace
ITU-AI-RE-Tookit-REQ-023	Data/model marketplace
ITU-AI-RE-Tookit-REQ-024	Not applicable
ITU-AI-RE-Tookit-REQ-025	Contextualization & regional impact
ITU-AI-RE-Tookit-REQ-026	Not applicable
ITU-AI-RE-Tookit-REQ-027	Not applicable
ITU-AI-RE-Tookit-REQ-028	Not applicable

NOTE: Some of the requirements are internal facing, e.g., usability requirements for the Toolkit. These are not mapped to any dimension and marked as "Not applicable" in the table above.

6. Future work

Expansion of Plugfest Initiative: To further the work, the ITU AI Readiness project will expand the plugfest initiative to collaborate with more regional projects to set up the sandbox environment and demonstrate AI integration addressing real-world problems, supported by ITU AI for Good cloud credits. This would involve on-boarding experts from diverse countries and domains. These compute resources empower teams to collect and process local data, experiment and train models with scalable compute and storage resources, enabling rapid prototyping and validation of AI models. Ultimately, the initiative bridges research and implementation, fostering globally connected, standards-based AI ecosystems.

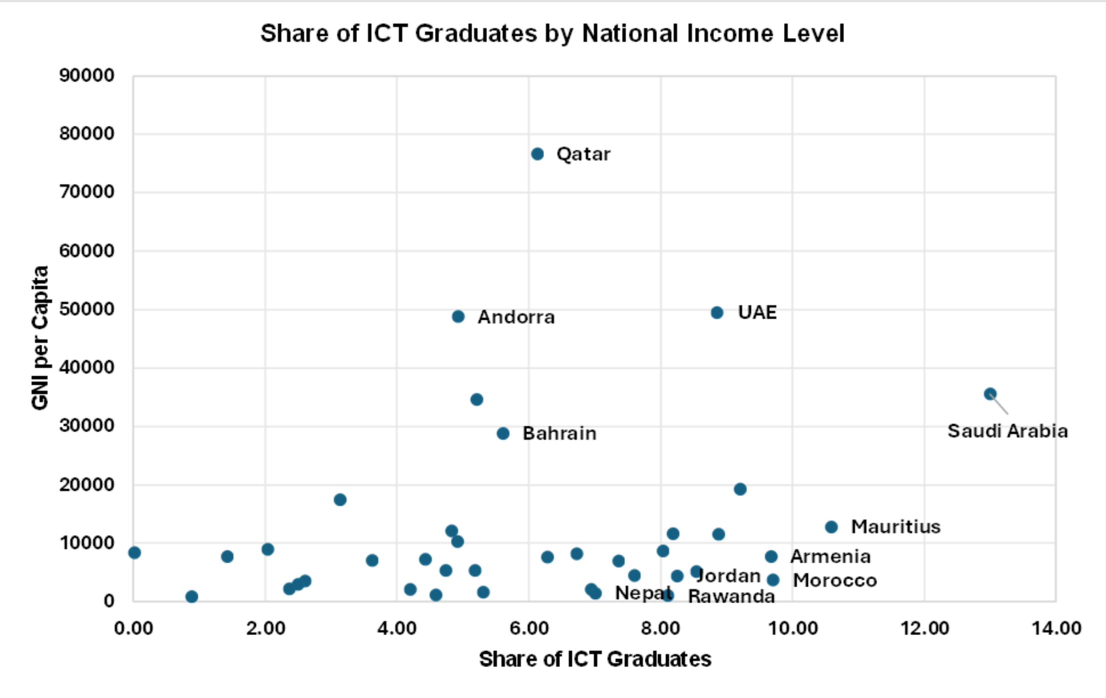
Collaborative Standards Development: Standards gaps identified as part of the plugfest project analysis for the AI Readiness framework would be proposed by members of ITU Study Groups. This standards development work will facilitate interoperable AI solutions, thereby providing more choices for early adopters of AI solutions. Open standards development processes, such as ITU, increase the trust among stakeholders in AI solutions, leading to increased adoption and successful deployments. Standardized AI Readiness evaluation and engaging with more members lead us to collectively govern AI effectively in the future.

ITU AI-RE Toolkit Pilot Launch: The ITU AI-RE Toolkit demo pilot will be launched in July 2026. Pilot users would be chosen and encouraged to run the toolkit locally. Regional data and indices would be applied, a self-evaluation of the AI Readiness level in the country/enterprise/organization would be piloted, and results would be analysed together with the relevant stakeholders. Continuous feedback from the users would help to further optimize and improve the toolkit.

ITU AI Readiness Challenge Launch: The ITU AI Readiness Challenge, with a specific focus on the 6 key factors, will be hosted in different regions to crowdsource solutions for building an open knowledge base. This not only helps in creating awareness of ITU AI Readiness efforts but also channelizes the research in multiple domains to solve relevant problems using AI. The selection of the problem statement and data used for the challenge would both be based on regional priorities and preferences. Prizes will be awarded for winning solutions. Mentoring sessions by local and international experts will be provided, with an aim for local capacity building on AI Readiness. Winning solutions may also lead to contributions to AI standards and opensource initiatives in ITU.

Expansion of ITU AI for Good Sandbox Network: The ITU AI for Good Sandbox Network is envisioned as a distributed, standards-based platform that lowers barriers to AI research, experimentation, and innovation. It aims to empower users – ranging from researchers and startups to policymakers and students – to develop, validate, and share AI solutions aligned with ITU-T standards. By integrating datasets, models, and simulators into a unified, trusted environment, the Sandbox enables collaboration across countries and domains, bridging the gap between open science, research, and standardization. Expansion of this Sandbox into new regions and countries would help to bring cutting edge AI technology and standards closer home, providing local population access to skilling and trainings. Toolkit and models would be hosted locally using ITU standards-based Sandbox, providing a trusted and secure environment for self-assessment leading to concrete actions and optimizations.

Appendix: Additional Information



Country	ICT Skills (%)	GNI per capita (current us)	Income level
Malawi	1.99	540	low
Jordan	26.32	4 430	lower middle
Belarus	56.51	8 240	upper middle
Brazil	44.02	9 950	upper middle
Dominican Rep.	26.03	10 280	upper middle
Malaysia	75.28	11 670	upper middle
Türkiye	41.81	13 150	upper middle
Bulgaria	34.97	15 320	high
Russian Federation	45.90	15 320	high
Uruguay	66.92	21 580	high
Slovakia	50.68	23 900	high
Slovenia	54.60	31 640	high
Spain	69.99	33 410	high
Malta	67.44	34 660	high
Korea (Rep. of)	69.24	35 490	high
Saudi Arabia	84.22	35 570	high

(continued)

Country	ICT Skills (%)	GNI per capita (current us)	Income level
Italy	53.66	38 290	high
France	55.83	45 180	high
Canada	66.96	53 340	high
Austria	66.87	54 160	high
Belgium	68.70	54 840	high
Sweden	81.96	58 820	high
Singapore	77.64	74 750	high
Switzerland	72.56	95 900	high

Appendix: FAQ

1. How do we map the ITU framework with other frameworks on AI Readiness?

    a. ITU AI Readiness is a bottom-up approach. Initially, the focus was on plugfest projects and corresponding usage scenarios. Based on this domain-specific analysis, a list of dimensions is derived along with domain-specific metrics. Country-specific inputs are captured in the indices. A practical toolkit is designed based on this analysis as well. This forms the first level of output from the ITU AI Readiness study.

    b. However, this is an evolving study. Just like an AI model, the more use cases are studied using this approach, the more accurate the framework would be. Hence, further stakeholder engagement, expansion of plugfest projects, and crowd-sourced input gathering in the form of ITU AI Readiness Challenges are planned.

    c. We plan to do the mapping of the ITU AI Readiness Framework to other frameworks once we reach a level of stability and maturity with the first round of self-assessment and toolkit feedback in 2026.

2. What is the difference between the data/model marketplace dimension and the generated content marketplace dimension?

    a. Data/model marketplace focuses on the ecosystem that might provide value with the help of AI, while the generated content marketplace is more concerned with the ecosystem that allows stakeholders to generate new content with the help of AI. In the first dimension, open data and open source are already available to be "exchanged", while in the latter, new content/new data/new models are on their way to being created.

3. What is the difference between granular priorities and contextualization and regional impact?

    a. The granular priorities dimension focuses on local policies and priorities, while the contextualization and regional impact dimension focuses on customization of the solutions.

4. How to get evidence on metrics and relevance?

    a. Metrics will be provided by domain experts based on the characteristics of the domain and selected based on the relevance, mapping to the dimension, and importance to the domain.

    b. Regular stakeholder reviews are held to study relevance, importance, and mapping.

    c. The feedback from toolkit pilots is another mechanism to evaluate the relevance and importance of metrics.

5. How would the correlation between domains be derived? Will that be accurate?

    a. The correlation between domains will be built when constructing the toolkit Knowledge Base based on the domain knowledge input. Statistical distance between the relevant features in the domain will be studied. Distance metric and the closeness of parameters will be analyzed so that the correlation between domains can be understood.

6. What is a simulated timeline in the AI & Policy dimension?

    a. Sandboxes will be used to evaluate and predict the potential impact of the policy in the field. A simulated timeline will be generated in the sandbox to observe the policy impact.

   b. A simulated timeline allows the user to ask "What-if" questions regarding the new policy. The user can configure parameters and simulators (number of parameters and simulators) in the sandbox to generate a predicted timeline. What-if questions could be used to evaluate conservative vs. radical changes in policies or resource allocations, or weightages.

7. Are we using AI to define policy on AI? The recursive use of AI should be prevented.
   a. In the AI & Policy dimension, AI is used in the sandbox to provide implications of policy in the verticals. Feedback and advice on the policy, per se, are provided by domain experts based on the extrapolation from the simulated timeline in the sandbox. Hence, the use of AI in policies is limited to validating the impact.

8. What can be achieved with respect to vertical synergy by using the ITU AI Readiness toolkit?
   a. In the "Strategy alignment" dimension, intent-based task distribution was discussed, and the performance would be evaluated. Gap analysis or actionable plans would be provided based on the evaluation results.
   b. The results from the toolkit can point to improvement areas for the country. This could be derived from individual domains, even in countries where the resources for coordinating AI governance are limited. Using the ITU AI Readiness toolkit would bridge the gap between assessment and improvement through intelligent recommendations.

9. How do we address the problem of a lack of local data, data in digital format, in the local language, and the lack of local models?
   a. Data/model marketplace Dimension measures the properties of datasets and models, and the regional data and workflows will be measured in the dimension for contextualization and regional impact. Together, these dimensions track and address the problem of a lack of local data and data formats.
   b. In addition, in the toolkit, we are evaluating the use of local data usage and local languages in plugfests, such as the projects on Kiswahili (Tanzania), Shona and Ndebele (Zimbabwe), and Amharic (Ethiopia) languages.

10. Where in the framework do we handle human capital?
   a. Human capital is captured in the "Impact of human in AI integration" dimension, where the impact of AI practitioners, domain experts, and general users is evaluated. The expertise of experts in terms of their length of domain practices, the credentials held, and the intersection with other types of experts serves as an important metric.

11. How do we address economic, ethical, social, sustainable, and legal problems with this framework?
   a. This framework will look at the benefits of integrating AI into the workflows. Trustworthiness, privacy, security, and explainability issues will be supported by standards along with the framework and the toolkit. Social divide and gaps will be addressed in the "AI for Inclusion" dimension, measuring the impact of AI in this regard. Sustainability and energy concerns will be discussed with energy readiness standards. Legal considerations are not within the scope of this study.

12. How are ethics handled in the framework?
   a. To practically address "ethics" and to evaluate it by metrics, "ethics" has been addressed in the way of "trustworthiness," "safety," and "compliance with local rules," in the dimensions.

13. Is ITU building its own platform or using the existing platform, such as Hugging Face?
    a. The ITU AI-RE toolkit is a self-developed platform that enables the evaluation of the readiness level of adopting AI techniques in a given entity/country. The AI Readiness project itself does not provide or use platforms such as Hugging Face for data/model exchange.
    b. However, interoperability between existing platforms and the ITU AI-RE Toolkit can be realized using ITU standards. Scalability and adoption of the toolkit can be enabled by using ITU standards.
14. Is ITU counting the transaction volume in data/model marketplaces?
    a. Metrics such as transaction volume in the data/model marketplace are important for the evaluation of AI Readiness.
15. What is the model architecture?
    a. The model is created based on LLM and RAG.
16. Is there any information/data needed from the users?
    a. No additional information or data is needed from the users. Even if private data is used to train the model, it will only be stored in the knowledge base temporarily and will be deleted when the training is done.
17. Is it possible to organize dimensions into a few groups, for example, based on their objectives, level of complexity, or phase of implementation?
    a. Dimensions are grouped based on the corresponding factors in the framework.
18. Are there possible interlinkages among these dimensions that can be shown on a flow diagram? It would help provide a framework to visualize relationships.
    a. Interlinkages among dimensions could be studied by the knowledge base with enough domain input. Compound gaps or co-benefits of dimensions will be studied and provided to the users.
19. What are the differences between the framework and the toolkit?
    a. The AI Readiness framework is comprehensive, providing all factors, dimensions, indices, and metrics. At the evaluation time, the user of the toolkit may input a set of materials, and the toolkit will use dimension-based feature selection to derive the subset of dimensions and metrics applied in the specific case and produce visualizations and recommendations as output. Thus, the evaluation results, the recommendations, and the gap analysis are based not only on the input materials but also on the relevant knowledge records stored in the knowledge base.
    b. The toolkit is an implementation instance of the framework, and hence it is bound to have certain variations (such as the selection of factors, dimensions...) at run-time due to various granular priorities applied by the users. However, the overall framework is static and can function as a reference.
20. How do we measure "divide" or "bias" in the "AI for Inclusion" dimension?
    a. The divide is measured using a standard mechanism as defined in UN reports.[6]
    b. Bias is not part of the AI for Inclusion; it is limited to the data and policy dimensions.

6 unstats.un.org/sdgs/report/2025/The-Sustainable-Development-Goals-Report-2025.pdf

21. Which level is the framework focusing on? The evaluation will be done on which level?
    a. The users of the framework could be regions, countries, enterprises, NGOs, and other 3rd parties. These users would mainly apply for self-assessment based on the level of input.
    b. Evaluation and self-assessment may be published or private, and the results would be relevant to the level of input.
22. How do we address the different priorities between country-level users and enterprise-level users?
    a. Addressed by priorities/indices. The foundational model can be applied to everyone, yet users might choose to finetune the model by providing domain-specific materials and priorities.
    b. Based on the indices input by the different users, the metrics applied for self-assessment are different.
23. What is the definition of AI and AI Readiness? How did ITU come up with the last definition?
    a. The ITU AI Readiness project adopts the definition of AI from ETSI[7].
    b. Artificial Intelligence (AI): a computerized system that uses cognition to understand information and solve problems
    c. NOTE 1: ISO/IEC 2382-28 [i.7] defines AI as "an interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning".
    d. NOTE 2: In computer science, AI research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions to achieve its goals.
    e. NOTE 3: This includes pattern recognition and the application of machine learning and related techniques.
    f. NOTE 4: Artificial Intelligence is the whole idea and concept of machines being able to carry out tasks in a way that mimics human intelligence and would be considered "smart".
    g. AI Readiness can be defined in multiple parts. For example, (1) GenAI Readiness is defined as a method for evaluating training and finetuning of generative AI in multiple domains, (2) Energy Readiness for AI is a standardized framework to evaluate and compare the energy efficiency of AI/ML models, (3) Data Readiness is a framework comprising evaluating the quality assessment of heterogeneous data.
24. With indices, some dimensions will be flexibly chosen when assessing the performance of AI. How do we guarantee a baseline comparison?
    a. The basic level of assessment, as defined in the levels, is the first level of adoption, which can provide the baseline for comparison across multiple users.
25. "Data marketplace" is a bit narrow for a dimension, implying data being bought and sold as a product. "Data marketplace" may be changed to "data services", or other general words.
    a. Data is not a product but an asset.

[7] GR ENI 004 – V3.1.1 – Experiential Networked Intelligence (ENI); Terminology

b. ITU-T Y.3176 definition of model marketplace. "3.2.2 Machine learning marketplace: a component which provides capabilities facilitating the exchange and delivery of machine learning models among multiple parties." In this context, the data marketplace and model marketplace are for the exchange of data and models.[8]

26. Why are the indices mentioned only in the toolkit design part? Why are fewer examples available for indices?

    a. We, as of now, focus on the basic framework, which consists of factors, dimensions, and metrices.

    b. Indices, which function as filters or weightages that reflect country-level priorities or local preferences, would be expanded in the next version with more inputs from users.

    c. The engagement with users would be achieved via the application of the toolkit.

[8] ITU-T Y.3176 : Machine learning marketplace integration in future networks including IMT-2020

References

[1] The Agency Fund, Rethinking AI Evaluation for the Social Sector, https://theagencyfund.substack.com/p/rethinking-ai-evaluation-for-the

[2] ITU-T Y.3172, Architectural Framework for Machines Learning in Future Networks including IMT-2020, https://www.itu.int/rec/T-REC-Y.3172/en

[3] The Northern Territory Government, Artificial Intelligence Assurance Framework, https://digitalterritory.nt.gov.au/digital-government/strategies-and-guidance/policies-standards-and-guidance/artificial-intelligence-assurance-framework#:~:text=The%20Northern%20Territory%20Government%20(NT,while%20carefully%20managing%20potential%20risks

[4] ITU-R M.2410-0, Minimum Requirements Related to Technical Performance for IMT-2020 Radio Interface(s), https://www.itu.int/pub/R-REP-M.2410

[5] United Nations, Sustainable Development Goals Report 2025, unstats.un.org/sdgs/report/2025/The-Sustainable-Development-Goals-Report-2025.pdf

[6] ETSI, Experiential Networked Intelligence (ENI); Terminology, https://www.etsi.org/deliver/etsi_gr/ENI/001_099/004/03.01.01_60/gr_ENI004v030101p.pdf

[7] ITU-T Y.3176, Machine Learning Marketplace Integration in Future Networks including IMT-2020, https://www.itu.int/rec/T-REC-Y.3176

[8] Information and eGoverment Authority, General Policy for the Use of Artificial Intelligence, General Policy for the Use of AI – Final 30 Jul 2025.pdf

[9] Ministry of Transportation and Telecommunications, 20231114152812215_5dylpg40_in5.pdf

[10] ITU AI for Good AI Readiness workshop, AI readiness – Towards a standardized readiness framework – AI for Good

[11] Hugging Face Model cards, Model Cards

[12] Google Model cards, Google Model Cards

[13] Timnit Gebru et al., Datasheets for Datasets, Cornell University, Datasheets for Datasets

[14] Shaina Raza et al., HumaniBench: A Human-Centric Benchmark for Large Multimodal Models Evaluation, HumaniBench: A Human-Centric Benchmark for Large Multimodal Models Evaluation

[15] FairSense AI, FairSense AI

[16] AI Fairness 360, AI Fairness 360

[17] Soto, Á., Durán, R., Moreno, A., Adasme, S., Rovira, S., Jordán, V. y Poveda, L. (Coords.) (2025). Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025. Documentos de Proyectos (LC/TS.2025/68). Comisión Económica para América Latina y el Caribe y Centro Nacional de Inteligencia Artificial, Docuemnto-ILIA_WEB.pdf

[18] Índice Latinoamericano de Inteligencia Artificial 2025, Home 2025 – Índice Latinoamericano de Inteligencia Artificial

[19] Kiran Raj Pandey et al., Leapfrogging Health Services Delivery In Resource-Limited Settings With Artificial Intelligence, Leapfrogging Health Services Delivery In Resource-Limited Settings With Artificial Intelligence by Kiran Raj Pandey, Bishesh Khanal, Udunna Anazodo, Rosalind Parkes-Ratanshi, Binod Bhattarai, Chetan Arora, Gilberto Ochoa-Ruiz, Rose Nakasi :: SSRN

[20] WTO, Trading with intelligence How AI shapes and is shaped by international trade, Trading with intelligence | WTO iLibrary

[21] OECD, OECD Framework for the Classification of AI System, OECD Framework for the Classification of AI systems | OECD

[22] Cazzaniga and others. Gen-AI: Artificial Intelligence and the Future of work. IMF Staff Discussion Note SDN 2024/001.IMF.2024.

[23] United Nations Conference on Trade and Development (UNCTAD). Technology and Innovation Report: Catching technological waves innovation with equity. UNCTAD.2021. https://unctad.org/publication/technology-and-innovation-report-2021

[24] UNCTAD. Technology and Innovation Report: Opening green windows technological opportunities for a low-carbon world. UNCTAD.2023. https://unctad.org/publication/technology-and-innovation-report-2023

[25] UNCTAD. Technology and Innovation Report: Inclusive Artificial Intelligence for Development.UNCTAD.2025. https://unctad.org/publication/technology-and-innovation-report-2025

[26] United Nations Development Programme (UNDP). AI Landscape Assessment. UNDP. https://www.undp.org/digital/aila

[27] UNDP. UNDP-UNESCO Joint Collaboration on Artificial Intelligence Assessments.2024. https://www.undp.org/sites/g/files/zskgke326/files/2024-12/undp-unesco-offer-web-7-aug-2024.pdf

[28] United Nations Educational, Scientific and Cultural Organization (UNESCO). Readiness Assessment Methodology – A Tool of the Recommendation on the Ethics of Artificial Intelligence. UNESCO.2023. https://www.unesco.org/en/articles/readiness-assessment-methodology-tool-recommendation-ethics-artificial-intelligence

[29] United Nations General Assembly. Enhancing International Cooperation on Capacity-Building of Artificial Intelligence. A/RES/78/311. adopted July 1, 2024. https://digitallibrary.un.org/record/4054005

[30] United Nations General Assembly. Pact for the Future (including Global Digital Compact as an annex). A/RES/79/1.adopted September 22, 2024. https://www.un.org/en/summit-of-the-future/global-digital-compact

[31] AI Talent Readiness Index for Africa 2025. https://talentindex.ai/overall-performance

International Telecommunication Union
Place des Nations
CH-1211 Geneva 20
Switzerland

ISBN 978-92-61-41911-0

Published in Switzerland
Geneva, 2026

Photo credits: Adobe Stock