Digital Innovation Gateway for Kosovo (DIG-4K)

Digital Europe Programme

This project has received funding from the European Commission's Digital Europe Programme

Project number: 101191620

Start date: 2025-01-01 Duration: 48 Months

https://itp-prizren.com/dig-4k/

Report on Establishment and Operational Readiness of AI Tech Access and Testing Labs

Document Title	Report on Establishment and Operational Readiness of AI Tech Access and Testing Labs
Author(s)	Osman Osmani — Head of Technical Operations Albana Zejnullahi — Software Developer (AI products) Blerta Salihu Krasniqi — Business Analyst
Number of pages	14
Document type	Deliverable
Work Package	WP4 — Tech Access and Testing Lab
Document number	D4.1
Issued by	Quantix L.L.C.
Date of completion	30.06.2026
Dissemination level	Public

Summary

This deliverable reports on the establishment and operational readiness of the AI Tech Access and Testing Lab led by Quantix L.L.C. under Work Package 4 of the DIG-4K project. It provides a detailed analysis of the availability and readiness of the lab's AI infrastructure and resources — hardware, software and cloud services — and identifies the key resources available for machine learning, data analytics and AI model training, including computational power and data storage. It evaluates the availability of skilled personnel and training capability, sets out the lab's readiness to begin service delivery to Micro Small and Medium size Enterprises (MSMEs) and Public Sector Organizations (PSOs), and highlights remaining gaps together with recommendations for the next reporting period. Submitted at Month 18, this report establishes the operational baseline from which service delivery and measurable progress against Work Package 4 indicators will proceed.

Contents

1. Introduction

Digital Innovation Gateway for Kosovo (DIG-4K) is a European Digital Innovation Hub (EDIH) co-funded by the European Union through the Digital Europe Programme and by the Government of Kosovo. The project supports the digital transformation of micro, small and medium-sized enterprises (MSMEs) and public sector organisations across Kosovo, and runs for 48 months from January 2025.

Quantix L.L.C. is the Lead Beneficiary of Work Package 4 (WP4) — Tech Access and Testing Lab. WP4 establishes and operates an AI Tech Access and Testing Lab that gives organisations access to advanced computational resources, AI technologies and technical expertise to develop, test and refine AI-driven solutions.

1.1 Purpose and Scope

The purpose of this deliverable is to:

- provide a detailed analysis of the availability and readiness of the lab's AI infrastructure and resources;
- assess the existing technological infrastructure — hardware, software and cloud services — required to support AI development and deployment;
- identify key resources, including data storage, computational power and specialised tools and platforms for machine learning, data analytics and AI model training;
- evaluate the availability of skilled personnel and training capability to support AI initiatives; and
- highlight gaps in infrastructure and resources, with recommendations to ensure effective implementation of AI technologies.

The scope of this report is limited to the responsibilities of Quantix L.L.C. under WP4, namely the Grant Agreement tasks: T4.1 (establish infrastructure), T4.2 (offer access to emerging technologies for testing), T4.3 (technical support and consultancy), T4.4 (create shared ecosystem for research and education), and T4.5 (test and print with 3D laboratories).

1.2 Position in the Project Timeline

D4.1 is due at Month 18 (M18), corresponding to the end of the first reporting period of the 48-month project. At this stage the deliverable reports on operational readiness: the point at which the lab's infrastructure and resources are established and the lab is ready to begin delivering services to MSMEs. The two subsequent WP4 deliverables — D4.2 (Catalog of AI Technologies Available for Testing, M36) and D4.3 (Compilation of Technical Support and Consultancy Services Provided, M48) — will report on technology offerings and delivered services respectively as the project progresses.

1.3 Methodology

The readiness assessment in this report is based on direct verification of the lab's production infrastructure (hardware specifications, installed software and frameworks, storage and security

configuration), a review of the lab's service model against the WP4 tasks defined in the Grant Agreement, and an assessment of demand identified through the DIG-4K MSME screening process. Infrastructure is described at capability level; technical identifiers and client-confidential details are intentionally omitted in line with the project's Data Management Plan (D1.2) and the public dissemination level of this deliverable. The assessment also draws on infrastructure inventories submitted by the consortium partners (UBT and KREN).

2. The AI Tech Access and Testing Lab

2.1 Concept and Operating Model

The AI Tech Access and Testing Lab is an operational environment, established and maintained by Quantix, that makes advanced AI compute, software platforms and technical expertise available to project beneficiaries (MSMEs & PSOs). Rather than a facility being newly constructed, the lab is an existing, configured and production-grade environment from which a defined share of capacity is allocated to DIG-4K beneficiaries for experimentation, testing and validation of AI-driven solutions. This model follows a structured "test-before-invest" approach where entities can validate AI use cases on real infrastructure before committing to investment.

2.2 Service Architecture

The lab's services map directly to the three WP4 tasks: provision of infrastructure (T4.1), access to AI technologies for testing (T4.2), and technical support and consultancy (T4.3). These are delivered across three layers — an access and delivery layer, a platform and AI/ML layer, and an underlying compute, cloud and storage layer — as illustrated in Figure 1.

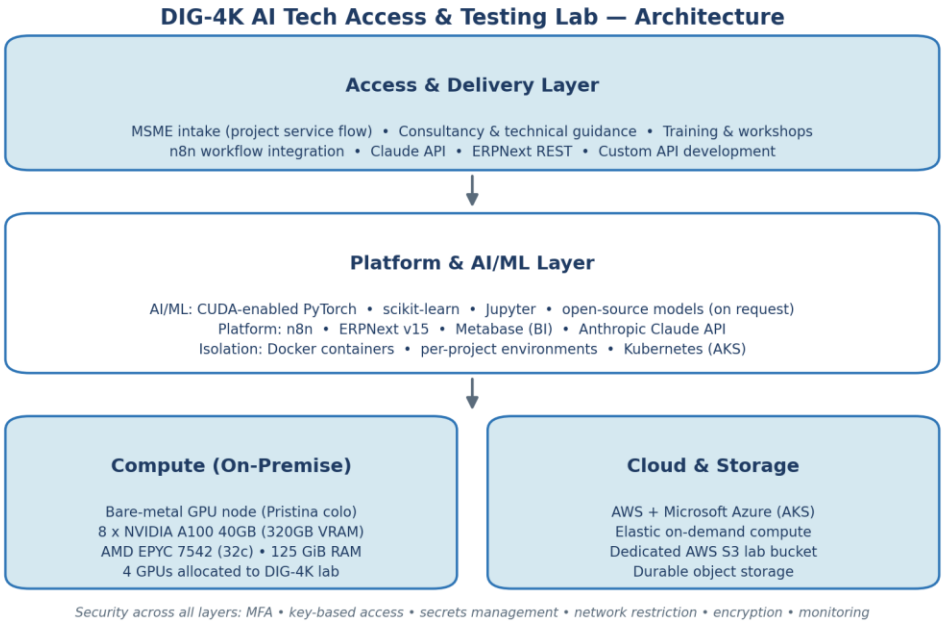

Figure 1: DIG-4K AI Tech Access and Testing Lab — architecture

2.3 Access, Governance and Data Handling

Access to the lab is managed through the DIG-4K service-delivery model and is restricted to authorised lab personnel and approved beneficiaries. Data generated through lab activity — technical specifications of solutions tested, usage logs, performance data and consultancy notes — is handled in accordance with the project's Data Management Plan (DIG-4K Dataset 4) and applicable data-protection law, including the General Data Protection Regulation (GDPR; EU 2016/679) and Kosovo Law No. 06/L-082. Within Quantix, the designated data-protection contact for WP4 activity is Osman Osmani (Head of Technical Operations). Project-level lab records are retained within secure project storage with access limited to lab teams; lab test data and model artefacts are stored in a dedicated object-storage bucket (see Section 3.4).

3. Infrastructure and Operational Readiness (T4.1)

This section assesses the existing technological infrastructure supporting AI development and deployment, addressing the hardware, software and cloud services that constitute the lab, and identifies the key resources available — computational power, data storage and AI tooling.

3.1 Computational Power

The lab's core AI compute is provided by a dedicated bare-metal GPU node hosted in a colocation facility in Pristina, Kosovo. The node is equipped with eight NVIDIA A100 PCIe accelerators (40 GB VRAM each; 320 GB aggregate GPU memory), dual AMD EPYC 7542 processors (64 physical cores / 128 threads in total) and 125 GiB of system memory, running a CUDA 12.0 environment. Of the eight A100 GPUs, four (50% of the node) are allocated to the DIG-4K lab for its beneficiaries experimentation and testing, with the remaining capacity supporting Quantix's own AI research and development. The lab allocation can be extended as demand grows.

3.2 Cloud Services

Beyond the on-premise GPU node, the lab can draw on Quantix's established cloud environments — active Amazon Web Services (AWS) and Microsoft Azure subscriptions, including managed Kubernetes (AKS) — to provide elastic, on-demand compute and isolated, reproducible test environments when a testing workload requires scaling beyond local capacity. This hybrid on-premise/cloud model allows the lab to match resources to the size and sensitivity of each beneficiary engagement.

3.3 Software and AI/ML Frameworks

The lab's software environment is organised in two layers. The AI/ML layer, hosted on the GPU node, provides a CUDA-enabled PyTorch environment together with torchaudio, scikit-learn, and the standard scientific Python stack (pandas, NumPy), with Jupyter / JupyterLab for interactive experimentation. This supports AI model training, testing, inference and refinement on the A100 hardware. The platform and automation layer comprises a self-hosted n8n workflow-automation engine, ERPNext v15 (ERP/CRM), Metabase for business-intelligence dashboards, and programmatic access to large language models via the Anthropic Claude API.

Together these layers enable both model-level experimentation and the delivery of complete AI and automation solutions to MSMEs.

3.4 Data Storage

Lab test data, datasets and model artefacts are stored in a dedicated DIG-4K AWS S3 object-storage bucket, providing scalable and durable capacity independent of local disk limits. The GPU node's local SSD serves as working and scratch space for active jobs. Project-level lab records (usage logs, performance data and consultancy notes) are retained in line with the project's Data Management Plan (Dataset 4).

3.5 Security, Monitoring and Virtual Environments

The lab environment applies layered security controls: multi-factor authentication and key-based access for administrative entry, centralised secrets management, network-level access restrictions, encryption of stored data, and continuous infrastructure monitoring. Cloud resources are governed by role-based access controls (IAM) with account-level multi-factor authentication. These measures align with the DIG-4K Data Management Plan and with GDPR and Law No. 06/L-082 requirements for integrity and confidentiality. The lab provisions isolated, reproducible virtual environments for AI research and development through containerisation (Docker) on the GPU node, per-project Python and Jupyter environments for dependency isolation, and managed Kubernetes (AKS) in the cloud layer for orchestrated, scalable isolation. Virtual-machine-based isolation is available as a capability where a specific engagement requires it.

3.6 Partner and Consortium Infrastructure (UBT and KREN)

Beyond the Quantix-operated environment described above, the DIG-4K consortium can draw on complementary infrastructure provided by two project partners — the University for Business and Technology (UBT) and the Kosovo Research and Education Network (KREN) — which extend the lab's reach for education- and research-oriented engagements.

UBT. The University for Business and Technology contributes higher-education laboratory and immersive-technology resources. The assets made available for DIG-4K comprise a 30-workstation computer laboratory (Lab N 114), five Meta Quest virtual-reality headsets, and access to the EON Reality platform — a VR/AR and AI-enabled content and training environment (ten seats). As a technology-focused higher-education institution with an innovation campus and applied IT laboratories, UBT also provides a route to academic expertise and student talent for the lab's shared research-and-education activities (see Section 6.3).

KREN. The Kosovo Research and Education Network is Kosovo's national research and education network, established under the Kosovo Digital Economy (KODE) project and interconnected with the pan-European GÉANT network. KREN operates a secure, high-speed national network together with a data centre offering hosting, virtual machines, cloud and processing capacity, storage and virtual laboratories, managed and monitored through the KREN Network Operations Centre. For DIG-4K, KREN extends the lab's reach to Kosovo's

higher-education institutions and the wider research-and-education community, providing connectivity, hosting and virtual-laboratory capacity for eligible engagements.

Usage scope of KREN infrastructure. KREN's infrastructure, investments and services are at present dedicated to higher-education institutions and the research-and-education community — that is, education institutions and public-sector organisations (PSOs) — and cannot currently be offered to MSMEs or the private sector. The use of KREN infrastructure to support MSME and private-sector beneficiaries may be considered at a later stage, subject to KREN's governance and the approval of the responsible authorities. Within DIG-4K, KREN infrastructure is therefore positioned to support the lab's PSO- and education-facing engagements, while MSME and private-sector testing continues to be served through the Quantix-operated environment.

3.7 Readiness Status Summary

Table 1 summarises the readiness status of the lab's key infrastructure components.

Component	Resource / Capability	Status
Computational power	Bare-metal GPU node — 8× NVIDIA A100 40GB (320GB VRAM), dual AMD EPYC 7542 (64 cores), 125 GiB RAM	Operational
Cloud services	AWS + Microsoft Azure (managed Kubernetes / AKS) for elastic compute	Available
AI/ML frameworks	CUDA-enabled PyTorch, scikit-learn, Jupyter; open-source models deployable on request	Operational
Platform & automation	n8n, ERPNext v15, Metabase, Anthropic Claude API	Operational
Data storage	Dedicated AWS S3 lab bucket (durable object storage)	Operational
Security & monitoring	MFA, key-based access, secrets management, network restriction, encryption, monitoring	Operational
Virtual environments	Docker containers, per-project environments, Kubernetes (AKS); VM capability available	Operational
Partner infrastructure — UBT	Computer laboratory (30 workstations, Lab N 114), 5× Meta Quest VR headsets, EON Reality VR/AR/AI platform (10 seats)	Available
Partner infrastructure — KREN	National research & education network (GÉANT-connected); KREN data centre — hosting, virtual machines,	Available (education institutions &

	cloud/storage, virtual laboratories (NOC-managed)	PSOs only)

Table 1: Infrastructure readiness status

4. AI Technology Access for Testing (T4.2)

4.1 Available AI Models, Algorithms and Frameworks

The lab provides access to AI technologies across three tiers. First, hosted large language models are available through the Anthropic Claude API, with the ability to integrate additional providers via API. Second, open-source and self-hosted models — including language, speech, vision and embedding models — can be deployed on request on the A100 GPU node using the lab's PyTorch environment, sized to the requirements of each engagement. Third, classical machine-learning methods (regression, classification and clustering) are available through scikit-learn for data-analytics use cases. This tiered model ensures beneficiaries can access both current-generation commercial AI and customisable open-source models appropriate to their needs and budget.

4.2 Tools for Testing, Evaluation and Refinement

The combination of GPU compute, the PyTorch and scikit-learn frameworks, and Jupyter-based interactive environments allows AI solutions to be built, tested, evaluated and refined within the lab. Isolated environments ensure that each engagement can be tested cleanly and reproducibly without interfering with other workloads.

4.3 Integration and API Access for External Platforms

The lab supports integration of AI and automation capabilities into beneficiaries' existing platforms. Workflow integration is delivered primarily through the self-hosted n8n automation engine, which connects external applications, APIs, webhooks and databases without bespoke development. Programmatic LLM access via the Anthropic Claude API can be embedded directly into client systems, and REST API integration (including via ERPNext) or custom API development is available where an engagement requires deeper system-to-system connectivity. This integration tooling is installed and operational; the integration work itself is carried out per engagement.

4.4 Test-Before-Invest Validation

The lab applies a structured validation approach in which AI use cases are mapped to specific business functions and tested on real infrastructure and data before an enterprise commits to investment. This protects MSMEs from investing in AI solutions that do not fit their operations, and is enabled by the lab's ready compute, tooling and isolated test environments.

5. Technical Support and Consultancy Readiness (T4.3)

5.1 Service Offering and Intake

Access to the lab's support and consultancy services is integrated with the DIG-4K service-delivery model, in which beneficiary needs are identified through the consortium's structured assessment process and qualifying AI and automation cases are referred to the WP4 lab. Quantix has established and defined its intake and engagement process for the lab, and is ready to receive and act on referrals as they are channeled through the project's service flow.

5.2 Consultancy and Guidance Capability

Consultancy and technical guidance are delivered by Quantix's WP4 technical team and external consultations with other experts from other EDIHs are practiced depending on the cross-border availability services, and beneficiary specific requests . The service covers guidance on AI model development and deployment, one-to-one consultancy for specialised projects, and the troubleshooting and optimisation of AI implementations. This capability is established and ready to be delivered to beneficiaries as engagements commence.

5.3 Cross-Border Expert Collaboration

As part of the European Digital Innovation Hub (EDIH) network, comprising more than 200 EDIHs across Europe, DIG-4K can facilitate access to a broad pool of specialised expertise and advanced digital infrastructure beyond the capabilities of its core consortium. Depending on the specific needs of beneficiaries, available budgets, and the interest of partner EDIHs in collaboration, DIG-4K may engage experts and leverage facilities from across the network to provide tailored support. This can include specialist guidance in areas such as artificial intelligence, advanced prompt engineering, optimisation of multi-step AI pipelines, cybersecurity, high-performance computing, digital manufacturing, and other emerging technologies, enabling beneficiaries to access complementary knowledge and capabilities through cross-border cooperation.

6. Skilled Personnel and Training Capability

6.1 Team and Roles

The lab is operated by a small, specialised Quantix team, supported by the company's broader engineering capability (spanning AI/ML, cloud-native applications and enterprise software). The core WP4 team is set out in Table 2.

Name	Role	Core competencies
Osman Osmani	Head of Technical Operations	Leads lab infrastructure and AI technical operations; oversees compute, deployment environments and technical delivery.
Albana Zejnullahi	Software Developer (AI products)	Develops and integrates AI/LLM solutions; model deployment, prompt engineering and workflow automation.
Blerta Salihu Krasniqi	Business Analyst	Requirements gathering, business

		process mapping and feasibility/needs analysis for client engagements.

Table 2: Core WP4 lab personnel

6.2 Training and Enablement

As part of T4.3, the lab provides for the training and enablement of MSME users, including onboarding and walkthrough sessions on the lab's tools, focused workshops (for example on prompt engineering and the practical use of AI and automation), and supporting documentation. This capability is established and ready; specialist workshop content can be delivered in collaboration with other EDIHs. The formal scheduling and roll-out of training will follow as engagements commence.

6.3 Shared Ecosystem for Research and Education (T4.4)

Task 4.4 concerns the creation of a shared ecosystem that connects the lab with Kosovo's research and education community. It is distinct from the user training and enablement described in Section 6.2, which falls under technical support and enablement (T4.3): T4.4 is the broader work of building a community and shared-resource base spanning academia, the public sector and industry.

The shared ecosystem is anchored by the consortium's academic and research partners. UBT contributes innovation-campus laboratories, immersive-technology resources (virtual-reality equipment and the EON Reality platform) and access to academic staff and student talent. KREN connects Kosovo's higher-education institutions to one another and to the pan-European GÉANT research community, providing the network backbone, hosting and virtual laboratories on which collaborative research and education activities can run.

Through these partners — and through the wider EDIH network of more than 200 hubs (Section 5.3) — the lab can link MSME and PSO use cases to academic expertise and shared research infrastructure, and channel validated AI use cases toward further research and education collaboration. The formal structuring and roll-out of joint research-and-education activities will progress as engagements commence and partner roles are formalised.

7. Readiness for Service Delivery

7.1 Identified Demand

Demand for the lab's services has been identified through the DIG-4K MSME screening process. Of 41 enterprises screened, 10 (24%) were identified as having artificial-intelligence or analytics-and-automation requirements suitable for the lab — spanning sectors such as media and creative technology, renewable energy, manufacturing, telecommunications and digital services. This provides a concrete pipeline of engagements for the lab to support.

7.2 Scoped Engagements

Two engagements have been scoped and initiated, demonstrating the lab's readiness to begin delivery:

- A Kosovo media-technology SME operating a multi-tool generative-AI content pipeline. The identified support needs are workflow automation to remove manual hand-offs between tools, and advanced prompt-engineering guidance for a multi-step pipeline. Specialist guidance has been sought through EDIH Hungary. Status: scoped; delivery to commence.
- A regional renewable-energy SME with operations across several markets. The enterprise has provided a multi-point AI-integration wish list spanning business intelligence, process automation and AI integration, and a support approach has been prepared. Status: scoped and assessed; delivery to commence.

In line with the public dissemination level of this deliverable and the project's Data Management Plan, individual enterprises are described in anonymized form at this stage.

7.3 Contribution to Work Package 4 Indicators

The lab's readiness positions it to deliver measurable progress against the WP4 key performance indicators for which Quantix is responsible. Table 3 sets out these indicators, their four-year targets and their status at M18. Lab utilisation (O5) is measured against the lab's allocated capacity of four A100 GPUs.

Indicator	Description	4-Year Target	Status at M18
O5	AI Tech Access & Testing Lab utilisation (% of lab capacity utilised)	30%	Infrastructure operational; 4× A100 GPUs allocated to the lab. Ready to deliver against target as engagements commence.
O5b	Collaborative partnerships established enabling lab utilisation	7	Cross-border expert collaboration active (EDIH Hungary); partnership formalisation in progress.
O6	AI solutions tested or developed within the lab	10 solutions	Lab environment ready; engagement pipeline identified. Ready to begin.

Table 3: WP4 key performance indicators — status at M18

8. Gaps and Recommendations

In line with the requirements of this deliverable, this section identifies the remaining gaps between current readiness and full service delivery, and sets out recommendations to address them in the next reporting period. At M18 the lab's infrastructure, technology access, support

capability and personnel are established and operational; the gaps below are therefore forward-looking and relate to the activation and scaling of service delivery rather than to the readiness of the underlying environment.

8.1 Identified Gaps

- Activation of service delivery: with infrastructure ready, the priority is to commence delivery on the identified engagement pipeline and begin recording lab utilisation against indicator O5.
- Formalisation of partnerships: collaborative relationships, including with EDIH Hungary, are active but not yet formalised through signed agreements counting toward indicator O5b.
- Solutions portfolio: no AI solutions have yet been tested or developed in the lab for beneficiaries (indicator O6); the pipeline to populate this is identified.
- Training roll-out: training and enablement capability is ready but not yet scheduled for delivery to project beneficiaries.
- Utilisation measurement: a consistent mechanism for recording and reporting lab capacity utilisation should be operationalised as delivery begins.
- KREN infrastructure scope: KREN's national research-and-education infrastructure is currently dedicated to higher-education institutions and the research-and-education community and cannot, at this stage, be offered to MSMEs or private-sector beneficiaries. MSME testing is therefore served through the Quantix-operated environment, which limits the share of partner infrastructure available for private-sector engagements. Extending KREN access to the private sector would require a change in KREN's usage policy and the approval of the responsible authorities, and is treated here as an external dependency rather than a near-term action.

8.2 Recommendations

- Commence delivery on the two scoped engagements and progressively onboard further enterprises from the identified demand pool.
- Formalise collaborative partnerships through written agreements to build toward the O5b target and strengthen the lab's expert network.
- Establish and apply a utilisation-tracking method (lab GPU-hours used against allocated capacity) to enable consistent reporting against O5.
- Schedule an initial cycle of onboarding sessions and a specialist workshop in collaboration with other EDIHs.
- Document each engagement to build the solutions portfolio that will evidence O6 and inform deliverables D4.2 and D4.3.
- Channel KREN capacity toward eligible public-sector and education engagements: given the current usage scope, direct KREN's connectivity, hosting and virtual-laboratory capacity to the lab's PSO and higher-education engagements, where its use is already

permitted, while the Quantix environment continues to serve MSME and private-sector beneficiaries. This makes full use of KREN's capacity for eligible beneficiaries and maintains a clear, compliant separation between public-sector and education use of KREN and private-sector use of the Quantix lab; the option of later extending KREN to MSMEs can be revisited with KREN and the responsible authorities as the project matures.

8.3 Forward Plan

Over the next reporting period, the lab will move from operational readiness to active service delivery, generating measurable progress against the WP4 indicators. The technologies made available will be documented in deliverable D4.2 (Catalog of AI Technologies Available for Testing, M36), and the support and consultancy services delivered, together with user feedback and case studies, will be compiled in deliverable D4.3 (M48).

9. Conclusion

At Month 18, the AI Tech Access and Testing Lab established by Quantix L.L.C. under Work Package 4 is operationally ready. The lab provides substantial computational power (a dedicated eight-GPU NVIDIA A100 node with capacity allocated to the project), elastic cloud services, a complete AI/ML and automation software stack, dedicated and durable data storage, layered security and monitoring, and isolated virtual environments — supported by a qualified technical team and an active cross-border expert collaboration. It is further complemented by partner infrastructure from UBT (laboratory and immersive-technology resources) and KREN (the national research-and-education network, data centre and virtual laboratories, available for education institutions and PSOs). Demand has been identified and initial engagements scoped. The lab is therefore positioned to begin service delivery and to generate measurable progress against the Work Package 4 indicators in the next reporting period, with the gaps identified in this report addressed through the accompanying recommendations.