Infrastructure Readiness Assessment for Federated Learning in Medical Imaging across African Sites

Mahlet A. Birhanu[1], Aisha Goedhart[1], Lidia Garrucho[2], Hakim C. Achterberg[1], Karim Lekadir[2,4] and Martijn P. A. Starmans[1,3]

[1] Department of Radiology & Nuclear Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands
[2] Barcelona Artificial Intelligence in Medicine Lab (BCN-AIM), Universitat de Barcelona, Barcelona, Spain
[3] Department of Pathology & Bioinformatics, Erasmus University Medical Center, Rotterdam, The Netherlands
[4] Department of Mathematics and Computer Science, Universitat de Barcelona

{m.birhanu, a.goedhart, h.achterberg, m.starmans}@erasmusmc.nl, lgarrucho@ub.edu, karim.lekadir@ub.edu

Abstract. Federated Learning (FL) enables collaborative, privacy-preserving AI model development on distributed medical data, which is a critical capability for equitable AI in heterogeneous, low-resource settings such as Sub-Saharan Africa (SSA). While FL has been shown feasible for medical imaging in Africa [1], infrastructure barriers remain poorly characterized before deployment. We conducted a structured infrastructure assessment across 10 SSA clinical and research sites from 10 countries, prior to data collection, using a questionnaire and 1-on-1 interviews with IT personnel, principal investigators, and data managers. Results reveal substantial heterogeneity: imaging formats span DICOM (50%), non-DICOM (40%), and film (10%); clinical data is digital at 40% of sites, paper-based at 50%, and mixed at 10%; storage ranges from dedicated servers to USB drives; and bandwidth spans 1–50 Mbps to over 1 Gbps. Findings are based on a convenience sample of 10 AFRICAI-RI-affiliated sites and may not generalize beyond this network. To conclude, our assessment provides a categorical baseline of infrastructure readiness that can inform planning for federated AI in SSA contexts and similar low-resource settings with comparable infrastructure challenges.

Keywords: Federated Learning · Medical Imaging · Infrastructure Readiness · Sub-Saharan Africa · Data Harmonization · Anonymization.

1 Introduction

Artificial intelligence (AI) models for medical imaging achieve the greatest clinical impact when trained on diverse, representative datasets. In Sub-Saharan Africa (SSA), this ambition faces a compound challenge: imaging data are distributed across institutions with heterogeneous hardware, fragmented data workflows, variable connectivity, and differing digitization standards. Data sovereignty constraints limit centralized pooling where all data is transferred and stored in a central location. Federated Learning (FL) [2] offers a solution to this by enabling collaborative model training without personally identifiable information (PII)/data sharing.

However, the success of any FL deployment depends not only on algorithmic design but on the readiness of the underlying research infrastructure at each participating site. In heterogeneous, low-resource settings, FL is hindered by bottlenecks ranging from incompatible imaging formats and unstable power supplies to absent data management staff. Despite growing interest in FL for medical imaging in Africa [1,3], infrastructure barriers have only been described retrospectively, after deployment problems emerge. FL on paired medical imaging and clinical data faces significant challenges compared to tabular or text data due to the complexity of handling multimodal information distributed across institutions. Both clinical and imaging data should first be digitized, anonymized (pseudo-anonymized), and then harmonized to ensure consistency. Afterwards, the two data sources should be paired and stored to be remotely accessed by an FL client. A systematic infrastructure assessment is therefore essential to ground FL architectures in local realities, quantify harmonization effort, and enable proactive planning. The closest related work is Asad et al. [3], which highlights that FL for medical imaging in low-resource settings remains hindered by system heterogeneity, connectivity constraints, and weak data governance.

This study provides an inventory of the medical imaging FL infrastructure readiness of low-resource settings by conducting a systematic infrastructure readiness assessment covering the status of hardware, workflows, and standards at 10 SSA sites. The results and findings of this paper can serve as a shared baseline reference for other research groups aiming to deploy federated learning on medical images across SSA sites.

2 Methods

The assessment was conducted across 10 SSA sites in three sequential steps: (1) questionnaire design and distribution, (2) structured 1-on-1 interviews, and (3) quantitative and thematic analysis. The 10 sites were selected as part of the AFRICAI-RI project [4], a pan-African federated imaging research initiative targeting tuberculosis and pneumonia detection, which motivated the focus on Chest X-Ray (CXR) modalities alongside paired clinical records.

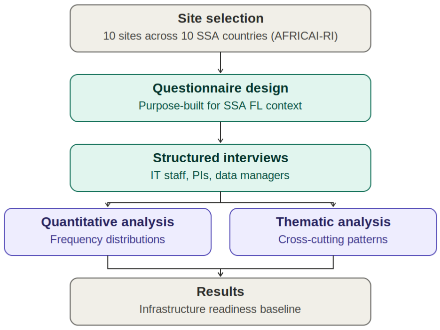

Fig. 1. Methodology workflow, from site selection through convergent quantitative and thematic analysis.

The selected sites are: Université Internationale de Dakar et de Thiès (UIDT), Senegal; Medical Research Council Unit The Gambia (MRCG), The Gambia; Delft Imaging Ghana (DIG), Ghana; Nigerian Institute of Medical Research (NIMR), Nigeria; Centre de Recherches Médicales de Lambaréné (CERMEL), Gabon; Centro de Investigação em Saúde de Manhiça (CISM), Mozambique; Kamuzu University of Health Sciences (KUHES), Malawi; Université de Kinshasa (UNIKIN), DR Congo; Infectious Diseases Research Collaboration (IDRC), Uganda; and Jimma University (JU), Ethiopia. These sites represent varied linguistic, infrastructure, and healthcare ecosystems.

2.1 Questionnaire Design and Distribution

A structured questionnaire was developed to capture infrastructure readiness for FL deployment in the SSA context; domains were purpose-built for this study, with the Cancer Image Europe (EUCAIM) [5] data holder's questionnaire referenced as one input during design. The first domain covered contact information, identifying personnel responsible for data management, IT operations, and scientific coordination at each site. The second domain addressed hardware assessment, including inventorying existing compute and storage resources (servers, graphical processing units (GPUs)), network connectivity and bandwidth, power backup systems (e.g. uninterrupted power supply (UPS), generators), and cooling infrastructure for sites with data centers. The third domain assessed data workflow and standards adoption, covering existing practices for data collection, storage, linkage, quality control, de-identification, and export. This section also checks compliance with imaging standards (DICOM), clinical data models, and controlled vocabularies (SNOMED-CT [6], ICD-10 [7], and LOINC [8]). For reporting purposes, standard image formats (JPG, JPEG, PNG, PDF) are grouped as non-DICOM. Pseudonymization and anonymization are used as defined by the GDPR. Pseudonymization replaces direct identifiers with a code linked to a lookup table and anonymization irreversibly removes all identifying information.

Forms were pre-distributed electronically for self-completion ahead of interviews. The questionnaire is publicly available in Zenodo [11].

2.2 Structured 1-on-1 Interviews

Following questionnaire distribution, a structured 1-on-1 interview was conducted by two research software engineers and key informants at each of the 10 partner sites: the site's principal investigator, data manager, and/or IT staff, depending on availability. Where questionnaire forms had been pre-filled, the interview served to review, clarify, and expand the submitted responses. Where forms had not yet been completed, the interview was used to collect responses directly. The interviews were semi-structured and used a predefined set of questions, but follow-up questions were asked when needed to clarify ambiguities and document constraints not covered by the questionnaire. Interview notes were recorded in writing during or immediately after each session.

2.3 Analysis

Responses are analyzed using two complementary approaches. Quantitative analysis produces frequency distributions across key infrastructure variables to characterize the infrastructure landscape across the consortium. Thematic analysis draws on qualitative summaries from interviews to identify cross-cutting patterns, common challenges, and site-specific constraints not captured by frequency counts, such as conditional infrastructure dependencies and critical data loss risks.

3 Results

Results are organized into three subsections: hardware infrastructure (Sect. 3.1), imaging data (Sect. 3.2), and clinical data and workflow (Sect. 3.3). Fig. 2 summarizes the distribution of the most informative variables across all 10 sites: imaging format (Fig. 2a), clinical data type (Fig. 2b), and storage infrastructure (Fig. 2c). Full per-site details are provided in Tables 1–3.

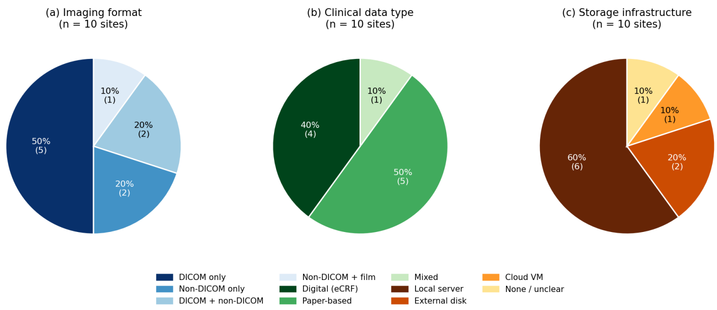

Fig. 2. Distribution of key infrastructure variables across the 10 SSA sites (n = 10). (a) Imaging format. (b) Clinical data type. (c) Storage infrastructure. CRF: Case Record Form, VM: Virtual Machine.

3.1 Hardware Infrastructure

Table 1 provides a full per-site overview of hardware infrastructure. Hardware maturity varied substantially across sites. Network bandwidth is a critical determinant of FL model exchange feasibility ranging from 1–50 Mbps in 2 sites (20%) to above 1 Gbps in 1 site (10%). Dedicated GPU workstations are available at only 3 sites (30%). Power supply was stable at most sites, though two reported weekly outages, and one reported daily outage without UPS. Regarding storage, most sites (60%) rely on local servers, while two (20%) use external disks, one (10%) uses a cloud virtual machine (VM), and one (10%) has no dedicated storage infrastructure (Fig. 2c). Most sites (80%) confirmed contacts with local hardware resellers, which is relevant for procurement planning.

Table 1. Hardware infrastructure per site. GPU: Graphics Processing Unit; PACS: Picture Archiving and Communication System; HW: Hardware provider; N/A: data not available or known.

Site	Country	Storage	Bandwidth	Power Stability	GPU	Local HW Contacts
MRCG	The Gambia	Local server	>200 Mbps	Stable	Yes	Yes
JU	Ethiopia	Local server	>1 Gbps	Stable	Yes	Yes
DIG	Ghana	Cloud Virtual Machine	N/A	Stable	No	Yes
KUHES	Malawi	External disk	50–100 Mbps	Weekly outages	No	No
CISM	Mozambique	Local server	50–100 Mbps	Stable	Yes	Yes
UNIKIN	DR Congo	Local server	200–1 Gbps	Weekly outages	No	Yes
IDRC	Uganda	Local server	1–50 Mbps	Fairly stable	No	Yes
CERMEL	Gabon	Local server	200–1 Gbps	Stable	No	No
NIMR	Nigeria	External disk	1–50 Mbps	Daily outages	No	Yes

UIDT	Senegal	N/A	N/A	N/A	No	Yes

3.2 Imaging Data

Table 2 provides a full per-site overview of imaging data characteristics. Nine sites (90%) exclusively produce digital images; one (10%) retains analog film archives requiring scanning before use. Five sites (50%) store data in DICOM format; two (20%) use non-DICOM image formats, which are often direct scanner exports or photographs of film; two sites (20%) use a combination of DICOM and non-DICOM image formats; and one site (10%) operates with a combination of film and digital images. Scanner models differ substantially across sites, reflecting heterogeneity in acquisition hardware, which is a relevant factor for AI model generalizability. Such cross-scanner heterogeneity is a documented driver of performance degradation in deep learning-based chest radiograph classification when models are evaluated outside their training domain [10].

Table 2. Imaging data characteristics per site. CXR: Chest X-Ray; PACS: Picture Archiving and Communication System; N/A: data not available.

Site	**Country**	**Imaging Format**	**Modality**	**PACS Extracted**	**Scanner Model(s)**	**Film Archives**
MRCG	The Gambia	DICOM	CXR	Yes	Delft Light / EasyDR	No
JU	Ethiopia	DICOM	CXR	Yes	N/A	No
DIG	Ghana	DICOM	CXR	Yes	Delft Light; Clarius/Telemed/Leltek US	No
KUHES	Malawi	DICOM	CXR	Yes	MinXray HF100H+; United uDR 780i; PERLOVE	No
CISM	Mozambique	DICOM + non-DICOM	CXR	Partial	N/A	No
UNIKIN	DR Congo	DICOM	CXR	Yes	Delft Light + CAD4TB	No
IDRC	Uganda	Non-DICOM	CXR	No	N/A	No
CERMEL	Gabon	Non-DICOM	CXR	No	N/A	No
NIMR	Nigeria	Non-DICOM + Film	CXR	No	N/A	Yes
UIDT	Senegal	DICOM + non-DICOM	CXR	No	N/A	No

3.3 Clinical Data and Workflow

Table 3 provides a full per-site overview of clinical data and workflow. Clinical data management is digital or already recorded in a Case Record Form (CRF) in 4 sites (40%), entirely paper-based in 5 sites (50%), and mixed at 1 site (10%). Digital platforms commonly used to manage clinical data include REDCap [9], Microsoft Access, and Excel. At paper-based sites, data managers will manually transcribe CRF entries into electronic formats prior to linkage to medical images. One site (UIDT) plans Optical Character Recognition (OCR)-based digitization of handwritten records, introducing additional quality control requirements. Clinical-to-imaging data linkage is currently manual at most sites, relying on Patient IDs, Medical Registration Numbers, or Study IDs matched by data managers. The naming convention of identifiers also varies per site. Two sites (DIG, CISM) use a lookup-table identifier rather than a direct patient ID for linkage, which reduces but does not eliminate manual matching effort; no site reported a fully automated linkage pipeline via integrated health records or PACS. Anonymization approaches range from GDPR/HIPAA-compliant procedures (KUHES) to pseudonymization with lookup tables (DIG, CISM) or direct patient IDs (MRCG, UNIKIN), through basic name or study ID removal (JU, IDRC, NIMR, UIDT), to no formal procedure (CERMEL). Standards vocabulary adoption is limited; only UNIKIN referenced SNOMED-CT, ICD, and LOINC compliance. JU reports diagnosis based on the ICD-10 standard.

Table 3. Clinical data and workflow per site. CRF: Case Record Form; EMRS: Electronic Medical Record System; GDPR: General Data Protection Regulation; HIPAA: Health Insurance Portability and Accountability Act; OCR: Optical Character Recognition; PACS: Picture Archiving and Communication System; N/A: not available or none declared.

Site	Country	Clinical Data	eCRF Platform	Linkage Method	Anonymization	Vocab. Standards
MRCG	The Gambia	Digital	REDCap / EMRS	Direct Patient ID	Pseudonymization	N/A
JU	Ethiopia	Paper	REDCap	Medical Reg. No.	Study IDs & name removal	ICD-10
DIG	Ghana	Digital	In-house dev.	Lookup table ID	Pseudonymization	N/A
KUHES	Malawi	Digital	REDCap / ODK	Direct Patient ID	GDPR/HIPAA compliant	N/A
CISM	Mozambique	Digital + Paper	REDCap	Lookup table ID	Pseudonymization, anonymization	N/A
UNIKIN	DR Congo	Digital	REDCap	Direct Patient ID	Pseudonymization	SNOMED-CT, ICD, LOINC
IDRC	Uganda	Paper	Microsoft Access	Direct Patient ID	Study ID removal	N/A
CERMEL	Gabon	Paper	REDCap	Direct Patient ID	None	N/A

NIMR	Nigeria	Paper	REDCap	Direct Patient ID	Name removal	N/A
UIDT	Senegal	Paper	None	Direct Patient ID	Name removal	N/A

4 Discussion

Our systematic assessment for infrastructure FL readiness revealed substantial heterogeneity across 10 SSA sites in imaging formats, clinical data management, storage infrastructure, connectivity, and standards compliance. This heterogeneity is not an anomaly, but a defining characteristic of real-world research infrastructure in SSA and similar low-resource settings, with direct implications for federated platform design. The specific variations provide insight into the bottlenecks that should be considered when designing FL systems in low-resource settings.

Several cross-cutting challenges emerged. First, manual data entry and linkage are very common practices. Clinical staff transcribe paper CRFs into eCRFs, and data managers manually match patient IDs, which are error-prone steps. Second, digitization of legacy data using film scanning and OCR may introduce noise requiring robust quality checking pipelines. Third, infrastructure instability like weekly power outages and slow bandwidth may impose hard constraints on synchronous FL model exchange. Additionally, specific non-standard ports might need to be opened for accessing a message queue of a federated learning orchestrator. The legal and privacy procedures to do so differ in different countries and sites. Fourth, limited dedicated data management capacity at several sites risks under-resourced curation and validation of workflows.

The bandwidth heterogeneity observed here (Table 1) has direct implications for FL protocol design: sites reporting 1–50 Mbps versus >1 Gbps will experience substantially different per-round model transfer times under a synchronous protocol, likely making the slowest sites the bottleneck. We defer empirical measurement of transfer time and convergence behavior to the forthcoming pilot study below, and flag asynchronous or bandwidth-adaptive aggregation as a concrete design direction.

Although this assessment gathered information on the most important variables, some sites reported incomplete information on the questionnaire that was clarified on interviews. This was partly caused by gaps in the questionnaire as well as lack of concrete information from interviewees. Language barrier was also occasionally an issue during conducting interviews as English was a tertiary language to local IT experts. Additionally, the 10 sites represent one network of SSA institutions; however, generalizability to other regions should be confirmed by additional assessments using the same methodology by using the findings as a baseline.

Several further limitations remain: no composite readiness index is synthesized (premature given n = 10), the questionnaire did not systematically capture governance readiness or security/adversarial readiness, and the assessment is pre-deployment, not yet validated against an actual FL pilot; the forthcoming pilot study below is intended to address the latter.

For future research, the identified bottlenecks define a concrete research agenda. Automated format conversion and non-DICOM anonymization pipelines are needed for sites lacking DICOM infrastructure. Lightweight FL protocols tolerant of intermittent connectivity should be prioritized for unstable-network sites. Standardized data management capacity-building programs are also required alongside technical deployment. Absolute GPU compute capacity was not benchmarked, as 7 of 10 sites report no dedicated GPU hardware (Table 1), making a comparative compute baseline infeasible at this stage. A pilot study is being organized to empirically measure GPU compute-hours, bandwidth effects on transfer time, and convergence behavior under real FL rounds, contingent on finalizing legal and data-governance agreements. This assessment thus establishes the categorical baseline; quantitative performance benchmarking is reserved for that follow-up. Within AFRICAI-RI [4], these findings will directly guide the deployment of a federated imaging and clinical data repository, with the aim of enabling the first large-scale multi-site federated respiratory disease AI study across SSA.

5 Conclusion

This infrastructure readiness assessment across 10 heterogeneous SSA research sites provides insights for deploying scalable federated medical imaging platforms in low-resource settings. Findings are based on a convenience sample of AFRICAI-RI-affiliated sites and should be interpreted as indicative rather than representative of SSA infrastructure more broadly. The results reveal a clear pattern, revealing heterogeneous imaging formats and infrastructure setup. The main contribution of this study for the research community is to give FL system designers an empirically grounded, categorical starting point that they can use when scoping multi-site deployments in low-resource settings.

Acknowledgments. This work is supported by the European Union's Horizon Europe programme under Grant Agreement No. 101190731 (AFRICAI-RI). The authors gratefully acknowledge contributions from all 10 participating AFRICAI-RI partner institutions and their IT managers, data managers, and principal investigators who completed the assessment questionnaire and participated in interviews. Additionally, the authors acknowledge all the meaningful comments made by the infrastructure working group of the AFRICAI-RI project.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Fabila, J., Campello, V. M., Martín-Isla, C., Obungoloch, J., Lekadir, K., et al. (2024). Democratizing AI in Africa: Federated Learning for Low-Resource Edge Devices. In MICCAI Meets Africa Workshop (pp. 101-109). Springer. (Also available as arXiv preprint arXiv:2408.17216, submitted August 30, 2024).
2. McMahan, B., Moore, E., Ramage, D., Hampson, S., Agüera y Arcas, B.: Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), pp. 1273–1282. PMLR, Fort Lauderdale (2017)
3. Asad, M., Shaukat, S., Duarte, J., Wang, Z., Bgn, S., Krishnamurthy, S., Chen, J., Bhatt, G., Hu, X.: Federated learning for privacy-preserving medical image analysis. In: ICLR 2022 Workshop on Practical ML for Developing Countries (PML4DC) (2022). https://pml4dc.github.io/iclr2022/pdf/PML4DC_ICLR2022_12.pdf
4. AFRICAI-RI: A Pan-African Research Infrastructure for Collaborative Biomedical Imaging and Artificial Intelligence. https://africai-ri.eu. Accessed June 2026
5. EUCAIM Consortium: European Cancer Imaging Initiative — Data Holder's Infrastructure Assessment Questionnaire. https://cancerimage.eu. Accessed June 2026
6. SNOMED International: SNOMED CT — Systematized Nomenclature of Medicine Clinical Terms. https://www.snomed.org/snomed-ct. Accessed June 2026
7. World Health Organization: International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10). https://icd.who.int. Accessed June 2026
8. Regenstrief Institute: Logical Observation Identifiers Names and Codes (LOINC). https://loinc.org. Accessed June 2026
9. Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., Conde, J.G.: Research electronic data capture (REDCap), a metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inform. 42(2), 377–381 (2009). https://doi.org/10.1016/j.jbmi.2008.08.010
10. Pooch, E.H.P., Ballester, P., Barros, R.C.: Can we trust deep learning-based diagnosis? The impact of domain shift in chest radiograph classification. In: Cardoso, J. (ed.) Thoracic Image Analysis, LNCS, vol. 12502, pp. 74–83. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-62469-9_8
11. Birhanu, M. A., Goedhart, A., Garrucho, L., Achterberg, H. C., Lekadir, K., Starmans, M. P. A.: Infrastructure Readiness Assessment Questionnaire for Federated Learning in Medical Imaging in Africa. Zenodo (2026). https://doi.org/10.5281/zenodo.20747183