

Gouvernance et circulation des données dans les archives ouvertes en Afrique : le cas de DICAMES

Article inédit, mis en ligne le 23 septembre 2026

Alphée Yannick Hilaire NGALLI NGOMO

*Enseignant-chercheur à l'Université de Douala, au sein du Département de Communication. Spécialiste des Sciences de l'Information et de la Communication (SIC), des humanités numériques et du numérique éducatif. Ses travaux portent principalement sur la gouvernance de l'information, la souveraineté numérique et l'endogénéisation des technologies éducatives en contexte africain.
yannickngalli@gmail.com*

PLAN DE L'ARTICLE

Résumé et mots clés en français
Title, abstract and keywords in english
Titulo, resumen y palabras clave en español
Introduction
Cadre théorique et revue de littérature
Méthodologie
Les défis de la gouvernance des données
Conclusion

RÉSUMÉ

Cette étude analyse la collection numérique DICAMES pour comprendre sa gouvernance et la circulation des données. Mobilisant une méthodologie bibliométrique, elle révèle une répartition géographique inégale et des biais structurels dans les métadonnées. L'article préconise une normalisation des données et une double validation pour corriger ces lacunes. Au-delà des aspects techniques, il mobilise les notions de souveraineté informationnelle et de diplomatie scientifique, concluant sur la nécessité de faire de ces archives des leviers stratégiques pour renforcer la position de l'Afrique dans le paysage scientifique mondial.

Mots clés

Collection numérique, archives ouvertes, gouvernance des données, circulation des données, DICAMES, diplomatie des données

TITLE

GOVERNANCE AND CIRCULATION OF DATA IN OPEN ARCHIVES IN AFRICA: THE CASE OF DICAMES

Abstract

This study analyzes the DICAMES digital collection to understand its governance and data

circulation. Utilizing a bibliometric methodology, it reveals an uneven geographical distribution and structural biases within the metadata. The paper advocates for data standardization and double validation to address these shortcomings. Moving beyond technical aspects, it engages with the concepts of informational sovereignty and science diplomacy, concluding that these archives must be leveraged as strategic tools to strengthen Africa's position in the global scientific landscape.

Keywords

Digital collection, open archives, data governance, data circulation, DICAMES, data diplomacy

TITULO

GOBERNANZA Y CIRCULACIÓN DE DATOS EN LOS ARCHIVOS ABIERTOS EN ÁFRICA: EL CASO DE DICAMES

Resumen

Este estudio analiza la colección digital DICAMES para comprender su gobernanza y la circulación de datos. A través de una metodología bibliométrica, revela una distribución geográfica desigual y sesgos estructurales en los metadatos. El artículo aboga por una normalización de los datos y una doble validación para corregir estas deficiencias. Más allá de los aspectos técnicos, moviliza las nociones de soberanía informacional y diplomacia científica, concluyendo en la necesidad de convertir estos archivos en herramientas estratégicas para fortalecer la posición de África en el panorama científico mundial.

Palabras clave

Colección digital, archivos abiertos, gobernanza de datos, circulación de datos, DICAMES, diplomacia de datos

INTRODUCTION

Lancée en 2017 par le CAMES (Conseil Africain et Malgache pour l'Enseignement Supérieur), la plateforme DICAMES est une archive ouverte visant à collecter, préserver et diffuser la production scientifique des 19 pays membres. Elle constitue un outil essentiel pour la visibilité et la circulation des savoirs en Afrique francophone, ouvrant l'accès à des travaux qui n'auraient peut-être pas trouvé d'autre canal de diffusion. Cependant, ces avantages ne se réalisent que si la gouvernance est effective.

Pour comprendre les caractéristiques de cet outil, il est nécessaire de le situer dans son contexte historique et institutionnel. Le CAMES, en tant qu'institution, a traversé plusieurs phases de développement, et DICAMES s'inscrit dans une stratégie de large modernisation et d'adaptation aux enjeux du numérique. Son lancement en 2017 répond à un double objectif : pallier l'absence de visibilité des recherches issues des établissements membres et garantir leur conservation à long terme. Au-delà de son architecture technique, la construction d'une archive ouverte multinationale comme DICAMES soulève des questions fondamentales de gouvernance des données scientifiques. Sa structure, ses règles de contribution et ses biais intrinsèques ne sont pas neutres ; ils reflètent et influencent la circulation des connaissances entre les pays membres et avec le reste du monde. Cet article

propose donc d'analyser DICAMES non seulement comme une collection, mais aussi comme un outil de diplomatie scientifique, en interrogeant ses mécanismes de gouvernance et son potentiel à réduire les déséquilibres dans la production et la diffusion des savoirs en Afrique francophone.

Trois notions structurent notre analyse : la gouvernance des données (processus et règles définissant l'accès et l'usage des données), la souveraineté informationnelle (capacité d'un État à contrôler ses infrastructures informationnelles) et la diplomatie des données (usage stratégique des données pour établir des alliances et influencer les normes internationales, s'appuyant sur le soft power (Nye, 2004).

DICAMES vise à préserver et à promouvoir l'ensemble de la production scientifique des pays africains qui ont le français en partage. Plus précisément, cette collection est principalement conçue pour combattre l'absence de visibilité et le faible accès numérique à la recherche émanant des établissements d'enseignement supérieur membres de la région CAMES, en mettant en place une approche pour garantir sa conservation à long terme dans des conditions idéales. DICAMES rassemble une large typologie de documents : thèses, mémoires, articles (avec ou sans comité scientifique), chapitres de livre, livres, rapports, communications, et autres documents jugés utiles par les gestionnaires.

La création de cette archive s'inscrit dans un mouvement global de promotion qui encourage la diffusion sans restriction des résultats de la recherche (Piron *et al.*, 2017). Cette collection compte aujourd'hui environ 8 000 documents repartis dans le tableau ci-dessous. La prééminence du Cameroun (plus de 6 000 documents) mérite d'être nuancée : elle témoigne autant de stratégies différenciées de valorisation du capital scientifique et de capacités techniques que d'une productivité inégale. DICAMES apparaît ainsi comme un instrument de cartographie des réseaux d'influence académique.

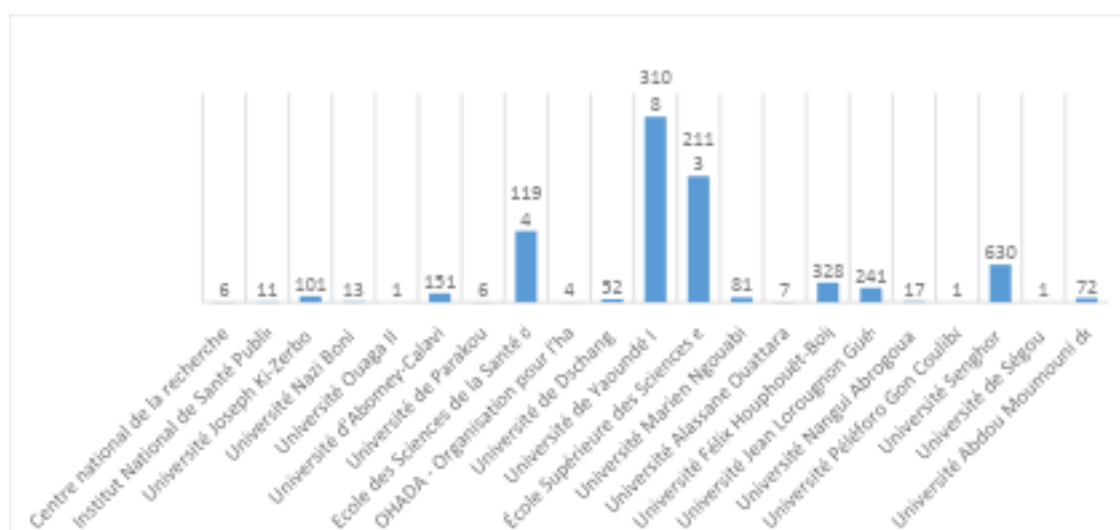


Figure 1. Contribution des établissements d'enseignement supérieur à la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Du point de vue de la gouvernance de l'information, l'analyse de la collection soulève de nombreuses interrogations sur les causes d'une telle disparité. Nous pensons que la méthode choisie pour la construction de ces archives ouvertes ne facilite pas sa mise en œuvre au regard de son caractère multinational et multi-établissements. La coordination et le management d'une telle initiative constitue un enjeu majeur de gouvernance et de circulation des données à l'échelle régionale.

En plus des documents, la base dispose aussi de métadonnées associées à ces documents et c'est cet ensemble qui constitue la collection numérique. En effet, chacun des documents

présents dans la base est rattaché à un certain nombre de métadonnées descriptives. Leur rôle est de faciliter l'accès à la copie numérique dudit document. Les métadonnées choisies par DICAMES sont : titre, auteur(s), mots-clés, date de publication, éditeur résumé, pagination / nombre de pages, URI/URL, *digital object identifier* (DOI), collection(s). Ces métadonnées ne sont pas standardisées puisqu'elles ne correspondent pas aux formats normalisés comme l'ISBD (*International Standard Bibliographic Description*), ni à un encodage standard, tel que MARC (*MACHine-Readable Cataloging*). Ce manque de standardisation peut être un frein à l'interopérabilité des données de DICAMES avec d'autres systèmes et entraver la circulation internationale des données qu'elle contient.

La question de recherche est : comment l'examen des données bibliométriques éclaire-t-il le processus de constitution de DICAMES et ses répercussions sur la gouvernance des données pour l'Afrique francophone ? L'hypothèse est que l'analyse des biais dans les métadonnées révèle les mécanismes de gouvernance et identifie des leviers pour renforcer la souveraineté informationnelle.

CADRE THÉORIQUE ET REVUE DE LITTÉRATURE

Cette recherche s'inscrit dans une approche interdisciplinaire au carrefour des Sciences de l'information et de la Communication (SIC), de la sociologie des sciences et des études postcoloniales sur la circulation des savoirs. Pour analyser DICAMES comme objet de gouvernance, nous mobilisons un cadre théorique articulé autour de deux axes : la sociologie des infrastructures informationnelles et les théories des relations internationales appliquées aux sciences.

Deboin (2015) montre que les techniques et réseaux informatiques constituent le substrat permettant aux chercheurs de s'organiser en communautés, favorisant l'émergence des collections numériques.

Berthaud et al. (2021) identifient trois facteurs dans l'émergence des collections numériques (actions individuelles, mobilisation communautaire, stratégies institutionnelles). Schöpfel et al. (2022) et Schöpfel (2020) montrent, à partir du cas HAL, que les pratiques de dépôt varient selon les disciplines et les environnements institutionnels.

Trzmielewski et Gnoli (2019), quant à eux, ont choisi de réfléchir sur une évaluation de l'*Integrative Levels Classification* (ILC) en vue d'organiser des données ouvertes de la recherche en sciences humaines et sociales. Leur travail s'appuie sur une analyse des propriétés de l'ILC par rapport aux pratiques, usages et contextes de partage et de médiation des données. Les résultats qu'ils ont obtenus présentent l'ILC comme prometteuse pour la construction d'une science ouverte, dynamique et novatrice. Cependant, les travaux centrés sur les usages et les politiques nationales doivent être complétés par une perspective critique sur les enjeux de pouvoir dans la circulation des données.

S'agissant de la collection numérique que nous avons choisie pour notre étude de cas, Piron *et al* (2017) la présentent comme une plateforme ouverte qui vise la conservation et la diffusion en libre accès de toute la production scientifique des pays d'Afrique francophone subsaharienne. Leur travail décrit les principales étapes du processus de création du DICAMES en montrant les difficultés et les défis, mais aussi les solutions trouvées. Au-delà des aspects techniques et organisationnels, la littérature s'intéresse de plus en plus aux questions de gouvernance et de pouvoir dans la circulation des données de la recherche. Borgman (2015) souligne que l'ouverture des données est autant politique qu'infrastructurelle, tandis que Fabre et Gardiès (2008) montrent que l'accès à l'information scientifique est structuré par des rapports de pouvoir reproduisant les hiérarchies mondiales. Dans le contexte africain, ces enjeux sont exacerbés par la fracture numérique, mais les archives ouvertes comme DICAMES peuvent être des outils de souveraineté scientifique, à condition d'une gouvernance robuste.

Nous mobilisons le concept de réseau sociotechnique (Callon, 1991) pour appréhender DICAMES comme un assemblage d'acteurs humains, d'artefacts techniques, de règles et d'enjeux politiques. Ce cadre permet d'analyser la construction des asymétries dans la production et la circulation des savoirs. Par ailleurs, la diplomatie scientifique (Ruffini, 2017) et la diplomatie des données (Morin et al., 2021) fournissent des outils pour comprendre comment DICAMES peut contrer l'invisibilité de la recherche africaine dans les bases internationales. Notre revue de littérature met en évidence que si les aspects techniques des archives ouvertes sont bien documentés, leur dimension géopolitique dans le contexte du Sud global reste peu explorée. C'est en réponse à cette lacune que nos travaux ont été réalisés.

MÉTHODOLOGIE : UNE ANALYSE BIBLIOMÉTRIQUE COMME SONDE DE GOUVERNANCE

Notre approche méthodologique mixte est centrée sur l'analyse bibliométrique des métadonnées de DICAMES. Cette méthode, définie par Archambault et Vignola-Gagné (2004), permet l'évaluation quantitative des sciences et l'identification des biais et disparités révélateurs des mécanismes de gouvernance à l'œuvre.

Objectifs et questions de recherche

En nous référant au cadre méthodologique présenté par Robert K. Yin (2003), cette recherche se penche sur les caractéristiques essentielles des travaux effectués dans les pays membres du CAMES à travers les archives ouvertes DICAMES. Cette analyse poursuit quatre objectifs : cartographier les thématiques privilégiées, identifier les acteurs majeurs, observer l'évolution historique des tendances, et analyser la structure de gouvernance implicite à travers les disparités de contribution.

L'hypothèse principale de cette recherche est que l'examen des données bibliométriques présentes dans la collection numérique DICAMES approfondit la compréhension du processus de gouvernance et de circulation des données de cette collection numérique.

Récupération et préparation des données

La récupération et la préparation des données bibliographiques de la collection numérique DICAMES est une phase cruciale pour réaliser nos objectifs de recherche. Nous avons recours à la technique de *scrapping* en utilisant le langage de programmation Python et la bibliothèque spécialisée connue sous le nom de *BeautifulSoup*. Cette méthode nous a permis d'extraire l'intégralité des métadonnées disponibles sur le site DICAMES. Le script développé est disponible en ligne (https://github.com/YANNHIL2022/travaux_cshn) pour assurer la transparence et la reproductibilité de la recherche. L'extraction des métadonnées a été réalisée par *scrapping* avec Python et BeautifulSoup, en suivant une procédure itérative par communauté, sous-communauté et collection. Les données ont ensuite été nettoyées avec OpenRefine, enrichies via Wikidata pour les coordonnées géographiques, et visualisées avec Khartis, Tableau et Excel. Le script est disponible en ligne pour assurer la reproductibilité (https://github.com/YANNHIL2022/travaux_cshn). Les données collectées ont ensuite été nettoyées et enrichies avec *OpenRefine*, puis réconciliées avec les données de Wikidata pour ajouter des coordonnées géographiques. Les visualisations ont été produites avec Karthis, Tableau et Excel.

Analyse des données

L'analyse des données a consisté en l'examen systématique des métadonnées selon quatre axes : la répartition géographique des contributions, l'évolution temporelle des dépôts, la typologie des documents et l'analyse des auteurs. Pour chaque axe, nous avons cherché à identifier des biais ou des disparités qui pourraient indiquer des mécanismes de gouvernance spécifiques. L'analyse ne s'est pas limitée à une description des graphiques, mais a cherché

à interpréter ces résultats à la lumière du cadre théorique présenté précédemment.

Une fracture géographique dans la contribution

Sur les 19 pays membres, seuls 8 contribuent activement, avec une fragmentation interne. Cette cartographie révèle moins une simple répartition géographique qu'une analyse géopolitique des rapports de force dans la production et le partage des savoirs, distinguant pays actifs (exportateurs de capital scientifique) et pays absents (consommateurs passifs).

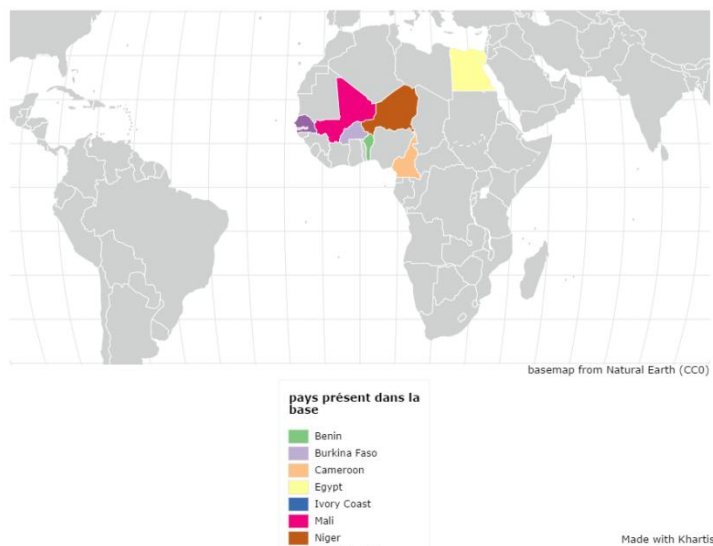


Figure 3. États contributeurs dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

La surreprésentation du Cameroun ne reflète pas nécessairement une domination académique objective, mais peut aussi résulter d'efforts de numérisation ou de politiques incitatives, tandis que les contributions modestes du Mali et du Niger peuvent s'expliquer par l'instabilité géopolitique. La concentration des contributions en Afrique Centrale et de l'Ouest, avec la présence de l'Égypte via l'Université Senghor, dessine un espace d'influence francophone structuré par la langue et les références culturelles. Cette cartographie, moins objective que politique, met en évidence une fracture dans l'accès aux données.

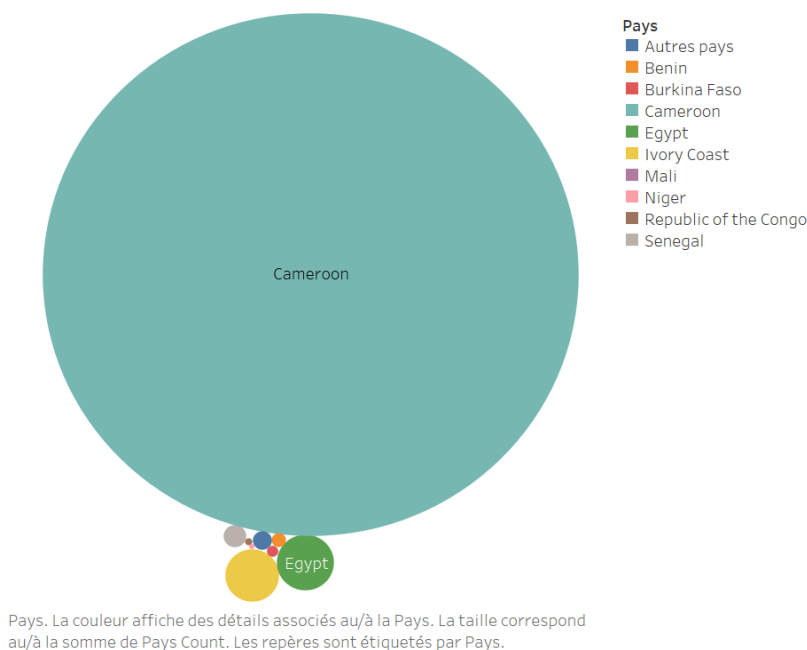


Figure 4. Proportion de la contribution des États dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Ce déséquilibre expose les pays faiblement représentés à un risque de dépendance, réduisant leur souveraineté académique. L'inclusion des données est donc une condition essentielle pour une diplomatie des données authentiquement multilatérale.

Une temporalité de la collection marquée par la date de création de la plateforme

L'ambition de DICAMES d'intégrer rétroactivement la production antérieure se heurte à l'absence de systèmes préexistants de gestion et de numérisation, imposant une reprise manuelle coûteuse. La figure 5 montre une sur-représentation de la décennie 2010-2020, indiquant une faible intégration des travaux antérieurs.

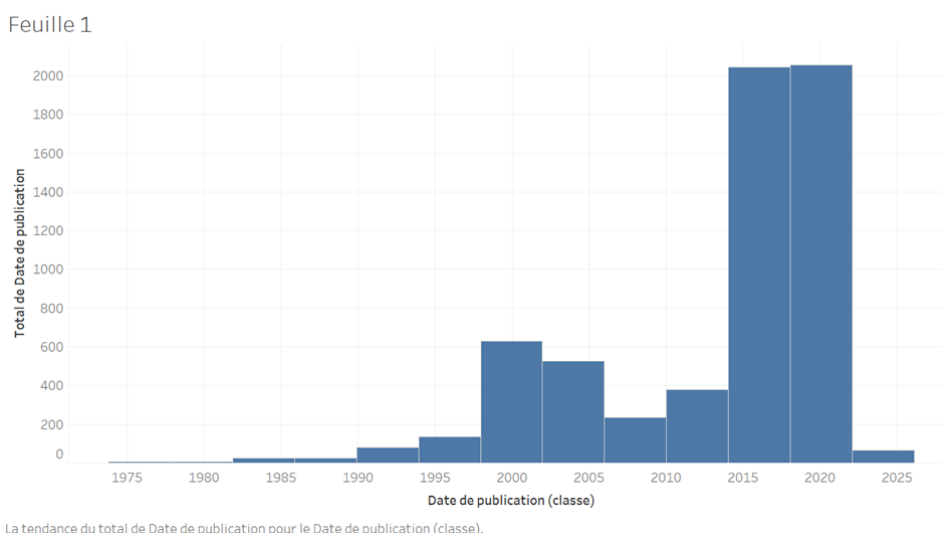


Figure 5. Evolution chronologique des données dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Quatre stratégies de gouvernance sont envisageables : globale (numérisation exhaustive mais coûteuse), ciblée (date limite avec extension progressive), prospective (production post-adhésion, au détriment de l'héritage), ou décentralisée (auto-archivage, plus rapide mais nécessitant un contrôle qualité). Le défi commun reste de renforcer l'autonomie informationnelle et la représentativité des connaissances.

Typologie des documents conservés dans DICAMES

La base de données comprend des thèses, des mémoires, des articles (avec ou sans comité de lecture), des chapitres, des livres, des rapports, des communications, et d'autres documents académiques (ex. : projets de fin d'études professionnelles).

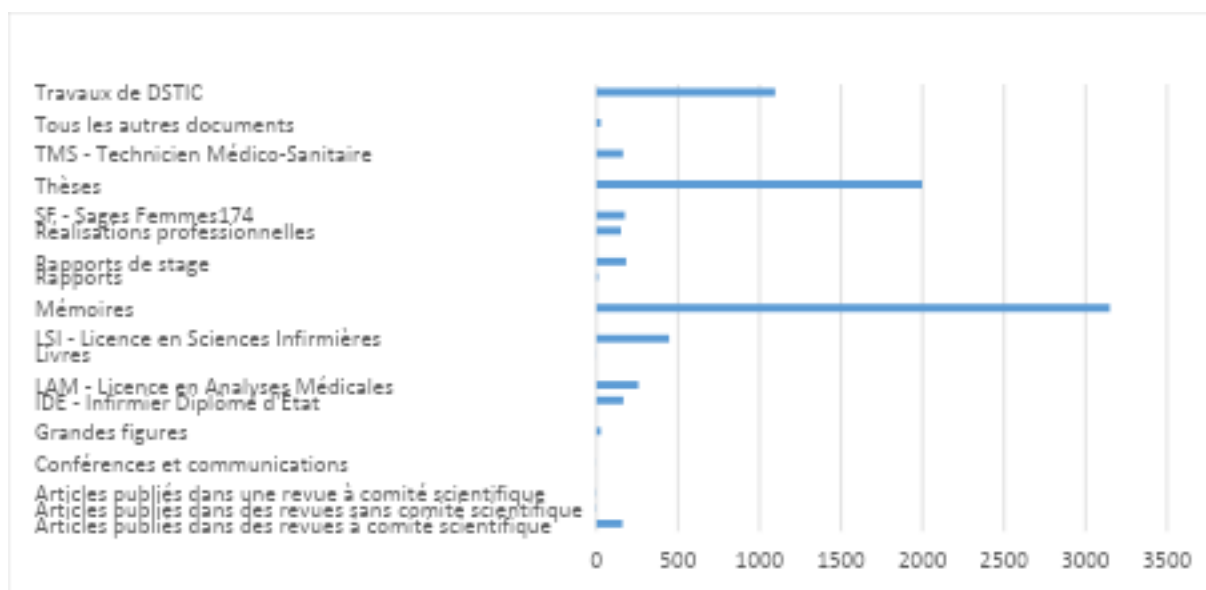


Figure 6. Nombre de documents par type de document dans la collection DICAMES. Données : DICAMES.
Graphique : Yannick Ngalli, 2024

Le graphique ci-dessus estime le volume de chaque sous collection dans la base de données. L'intention qui transparait dans le choix des documents est de regrouper tous les travaux issus de la recherche réalisée dans les établissements cibles. La base fait donc cohabiter les travaux scientifiques comme les thèses, les mémoires, les articles, les livres avec des rapports de fin d'étude de licence qui ne sont parfois que de simples comptes rendus des travaux effectués en stage. Pour bien cerner la constitution de la base en termes de typologie des documents, nous avons choisi d'analyser la catégorie la plus représentée à savoir les mémoires.

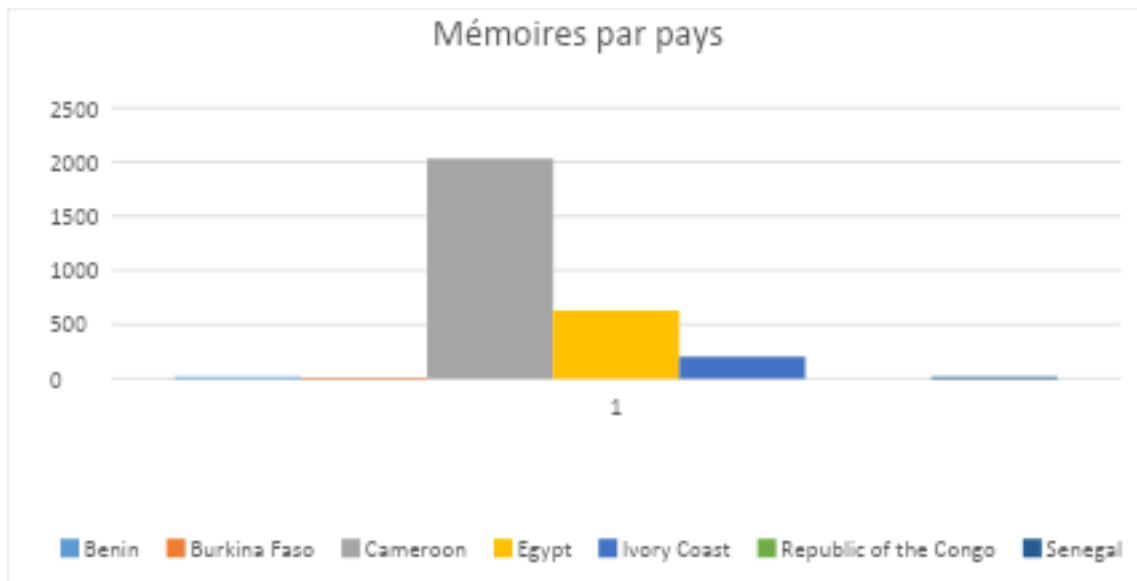


Figure 7. Nombre de mémoire recensés dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

La répartition des mémoires suit la tendance générale (hégémonie camerounaise), mais la position de l'Égypte (2e rang) révèle des contributions ciblées, reflétant les priorités et spécialisations des établissements, et les stratégies différenciées de valorisation du capital scientifique.

On observe ainsi une absence de linéarité entre le poids dataire global d'un pays et son influence dans des sous-ensembles spécialisés. Ce phénomène illustre les logiques de fragmentation et de spécialisation dans la diplomatie des données : certains acteurs peuvent compenser une faible masse critique globale par une contribution ciblée dans des niches à haute valeur académique ou symbolique. Cette situation appelle une gouvernance dataire plus fine au sein de DICAMES, capable de prendre en compte ces disparités thématiques et institutionnelles pour éviter que certains acteurs ne voient leur influence sous-évaluée en raison de biais de catégorisation ou de collecte.

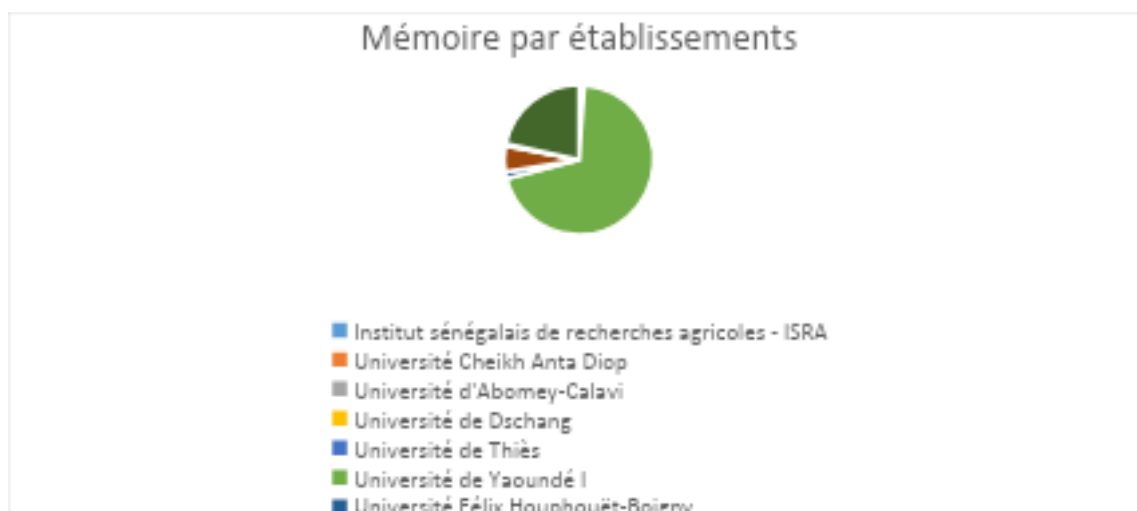


Figure 8. Nombre de mémoires par établissement dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

DICAMES : une diversité encyclopédique

La diversité encyclopédique de DICAMES résulte d'une politique délibérée de non-ségrégation disciplinaire, fondée sur le seul critère d'affiliation à un établissement membre. Si cet atout garantit une représentation équitable des savoirs, il constitue un défi en matière de découvrabilité sans mécanismes de navigation robustes.

Cette approche inclusive traduit une vision stratégique : constituer un bien commun informationnel qui reflète la diversité et la richesse des écosystèmes de connaissance nationaux, sans hiérarchie imposée, affirmant ainsi une autonomie dans la définition de ce qui constitue une production scientifique légitime.

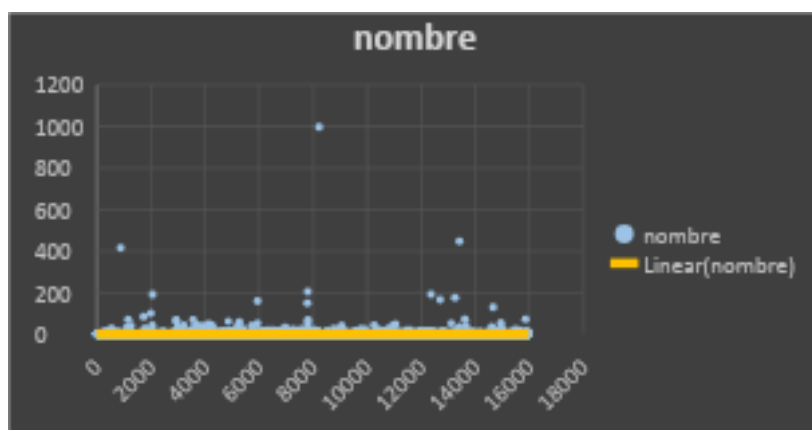


Figure 9. Thématiques dans DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Le graphique ci-dessus est une illustration du caractère encyclopédique de la collection DICAMES. En effet, sur les 16 000 thématiques qui sont dans la base, une seule revient plus de 500 fois, une dizaine reviennent plus de 200 et le reste est inférieur à 200. Il faut néanmoins noter que 80% des thématiques reviennent moins de 10 fois, ce qui démontre le caractère encyclopédique de la base qui couvre un très grand nombre de domaines de connaissance.

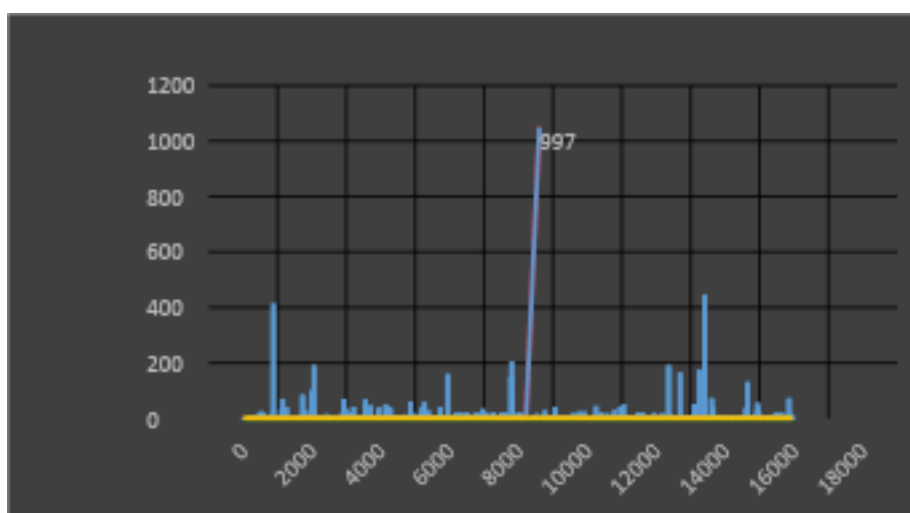


Figure 10. Thématiques ayant été étudiées dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Sur le plan de la gouvernance des données, l'extraordinaire diversité thématique de la collection DICAMES constitue à la fois sa richesse fondamentale et un défi stratégique majeur en matière de découvrabilité et d'accès à l'information.

Enjeu de gouvernance : la politique d'inclusion exhaustive, reflet d'un engagement diplomatique en faveur de la représentation de tous les savoirs, génère un écosystème de data complexe. Sans mécanisme de navigation robuste, ce pluralisme pourrait paradoxalement nuire à l'utilité et à la valorisation du patrimoine scientifique collecté.

Solution de gouvernance : pour transformer ce défi en opportunité, la gouvernance de DICAMES a institué un cadre d'accès multidimensionnel et intuitif. Cette infrastructure technique incarne le principe selon lequel la souveraineté des données implique la capacité à les explorer efficacement.

- Recherche par métadonnées structurées (date, auteur, titre, sujet) : cette approche garantit l'intégrité et la traçabilité de l'information, favorisant une auditable des contenus et une recherche précise, essentielle pour la recherche académique.

- Recherche par arborescence hiérarchique (« communauté et collection ») : cette méthode offre une cartographie navigable du savoir, allant du général au particulier. Elle reflète une vision diplomatique qui organise la connaissance sans la fragmenter, respectant ainsi l'architecture naturelle des disciplines et des institutions.

- Moteur de recherche global : cet outil assure l'exhaustivité de l'exploration basée sur une interrogation transverse de l'ensemble de la collection. Il symbolise l'unité de l'espace CAMES en offrant la possibilité de découvrir des connexions inattendues entre les disciplines.

Cette pluralité de modes d'accès n'est pas une simple fonctionnalité technique ; elle est l'expression opérationnelle d'une gouvernance des données axée sur l'utilisateur. Elle assure que la richesse encyclopédique de DICAMES reste pleinement exploitable, servant ainsi l'objectif diplomatique de diffusion et de valorisation équitable de la production scientifique de l'espace CAMES.

Le problème des auteurs : une granularité sémantique insuffisante

La collection reconnaît les auteurs comme les créateurs primaires du contenu intellectuel, un principe essentiel pour assurer la traçabilité et l'attribution des savoirs. Cependant, la gouvernance actuelle des métadonnées applique une politique d'inclusion extensive en regroupant sous l'entité unique "auteur" des rôles intellectuels distincts : créateurs principaux, directeurs de recherche et membres de jury.

L'analyse des auteurs, (figures 11 à 14) révèle un biais critique dans la gouvernance des métadonnées. La collection regroupe sous l'entité unique « auteur » des rôles intellectuels distincts : créateurs principaux, directeurs de recherche et membres du jury. Cette absence de granularité sémantique a des conséquences importantes. Elle fausse les indicateurs bibliométriques (un directeur de thèse peut apparaître aussi productif que le doctorant) et nuit à la capacité de DICAMES à servir d'outil de cartographie des expertises.

Enjeu de gouvernance des métadonnées, ce choix de modélisation, bien qu'inclusif, introduit un déficit crucial de granularité sémantique. Il engendre une ambiguïté fondamentale sur la nature de la contribution de chaque acteur, portant atteinte à l'intégrité analytique de la base de données.

Conséquences sur la diplomatie des données :

- Biais d'analyse : cette absence de distinction fausse les indicateurs bibliométriques. Un professeur, contributeur indirect en tant que membre de jury, peut apparaître comme plus "productif" qu'un étudiant auteur d'une thèse, biaisant ainsi toute analyse quantitative de la production scientifique réelle.

- Obstacle à la valorisation : cette situation nuit à la capacité de DICAMES à servir d'outil de

diplomatie scientifique. Pour négocier des partenariats ou cartographier les expertises à l'échelle du CAMES, des données fiables et différenciées sur la contribution exacte de chaque chercheur sont indispensables.

- Défi de souveraineté informationnelle : une représentation aussi bruitée des compétences affaiblit la revendication de souveraineté sur la cartographie des savoirs, qui est pourtant un objectif central de ce projet.

Une refonte de la gouvernance des métadonnées visant à normaliser et à différencier les rôles (auteur, directeur, examinateur) n'est pas seulement une amélioration technique. C'est une condition sine qua non pour garantir la qualité, la crédibilité et la valeur diplomatique des données, transformant DICAMES en un instrument fiable pour l'analyse et la promotion de la recherche africaine.

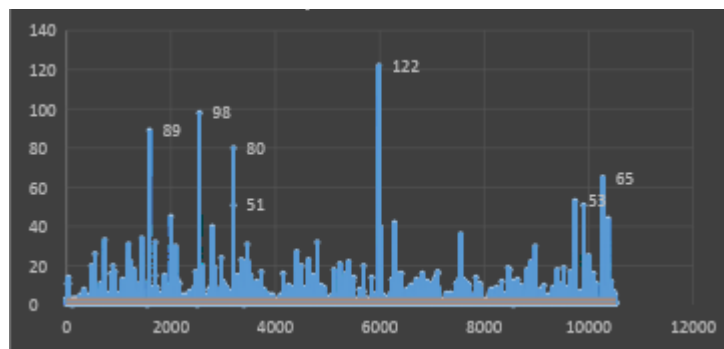


Figure 11. Publication par auteur dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Le faible nombre de documents par auteur s'explique par le fait que c'est une collection numérique des travaux produits à l'université. À l'exception des enseignants chercheurs qui ont l'obligation d'avoir une production scientifique permanente, les autres producteurs ont la seule contrainte de produire leurs travaux de fin d'étude à savoir leurs mémoires et thèses. La plupart des auteurs étudiants qui constituent la majorité des contributeurs auront un, deux ou trois documents dans la base en fonction de leur niveau d'étude alors que les enseignants ont la possibilité d'avoir une présence égale à leur production intellectuelle en plus des travaux qu'ils dirigent et les jurys auxquels ils participent. S'agissant des auteurs par type de document, il est important de noter que les auteurs suivent la tendance générale de la base de données en termes de proportions. C'est ce que traduit le graphique ci-dessous qui présente les auteurs des mémoires par pays.

Du point de vue de la gouvernance et de la diplomatie des données, la distribution des documents par auteur dans DICAMES n'est pas une anomalie statistique, mais le reflet direct de la structure et des finalités de l'écosystème académique qu'elle documente.

Cette caractéristique est gouvernée par deux réalités distinctes :

- La nature institutionnelle des contributeurs : la majorité des auteurs sont des étudiants, pour qui la production est ponctuelle et liée à l'obtention d'un diplôme (mémoire de master, thèse de doctorat). Leur contribution se limite naturellement à un faible nombre de documents.
- Le cadre normatif de la production scientifique : seuls les enseignants-chercheurs évoluent dans un cadre incitatif exigeant une production scientifique permanente et continue. Leur présence dans la base est donc potentiellement plus régulière et diversifiée, incluant non seulement leurs propres recherches mais aussi leur rôle d'encadrement et d'évaluation.

Du point de vue de la gouvernance des données, cette disparité crée une asymétrie fondamentale dans les métadonnées. Elle soulève un défi de granularité et de qualification

des contributions (comme évoqué précédemment avec la confusion des rôles auteur/directeur/jury). Une analyse quantitative brute, sans pondération par le statut et le rôle, générerait des indicateurs biaisés et non représentatifs de la dynamique réelle de production intellectuelle.

Du point de vue de la diplomatie des données, pour servir d'outil de diplomatie scientifique, DICAMES doit distinguer et valoriser ces deux types de contribution :

- La production étudiante qui représente la cartographie de la relève académique et du potentiel futur.
- La production des enseignants-chercheurs qui reflète l'expertise active et établie.

Une gouvernance des métadonnées qui différencie clairement ces apports produirait des analyses bien plus fines et stratégiques pour les décideurs politiques et institutionnels au sein de l'espace CAMES.

Concernant la répartition des auteurs par type de document, les données confirment qu'elle épouse la structure générale de la collection. Le graphique illustrant la nationalité des auteurs de mémoires agit ainsi comme un indicateur de la répartition géographique de la formation académique au sein de l'espace CAMES, un renseignement précieux pour une diplomatie éducative et scientifique ciblée.

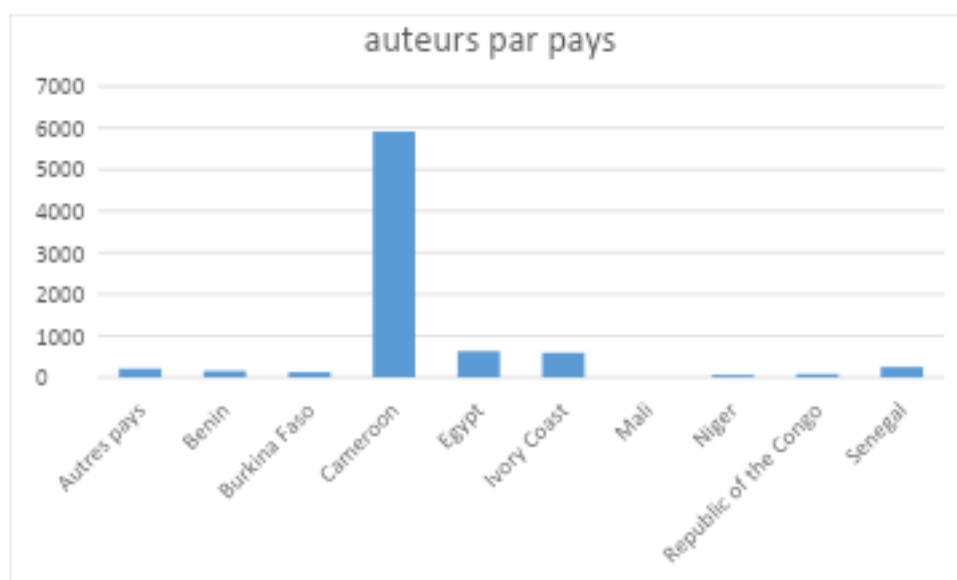


Figure 12. Représentation des auteurs par pays dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

L'histogramme ci-dessus présente les auteurs ayant contribué à la production de mémoires par pays. Le fait d'avoir 6 000 auteurs pour moins de 3 000 mémoires illustre notre propos selon lequel les directeurs et membres du jury sont cités comme auteurs. Cette situation rend les statistiques des auteurs difficiles à analyser parce qu'on y mélange des personnes ayant des fonctions très différentes (candidat, encadreur, membre de jury). Cette difficulté s'applique aussi aux auteurs par établissement comme l'illustre le graphique ci-dessous.

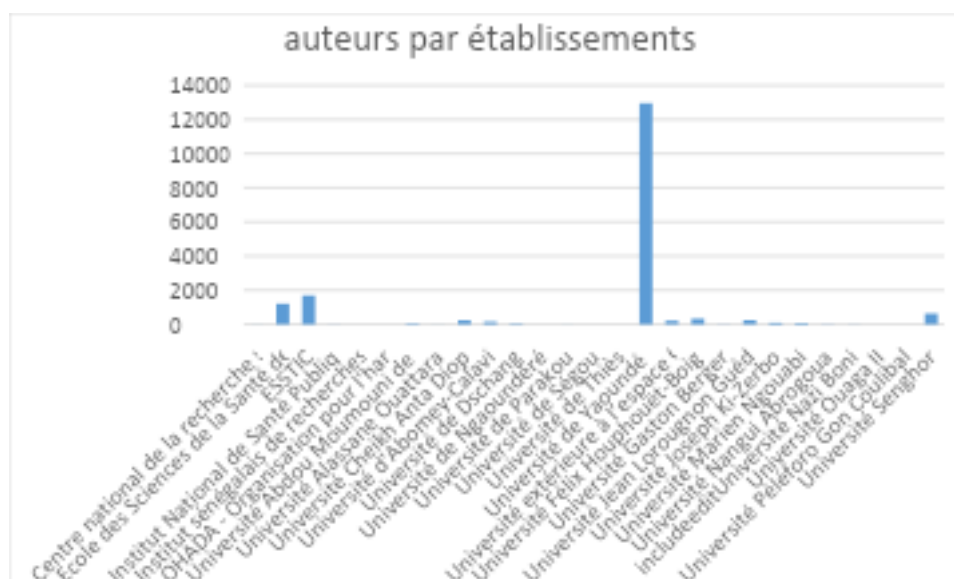


Figure 13. Représentation des auteurs par établissements dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Le nombre d'auteurs par établissement est supérieur aux documents disponibles dans la base. Par exemple, le nombre d'auteurs de l'université de Yaoundé est l'équivalent du double de documents disponibles dans la base. Cette situation indique que le choix des métadonnées (dans ce cas des auteurs) et la façon de les représenter ont une incidence sur la construction de l'information dans une collection numérique.

Ce constat est aussi valable pour l'analyse des auteurs en fonction du type de documents.

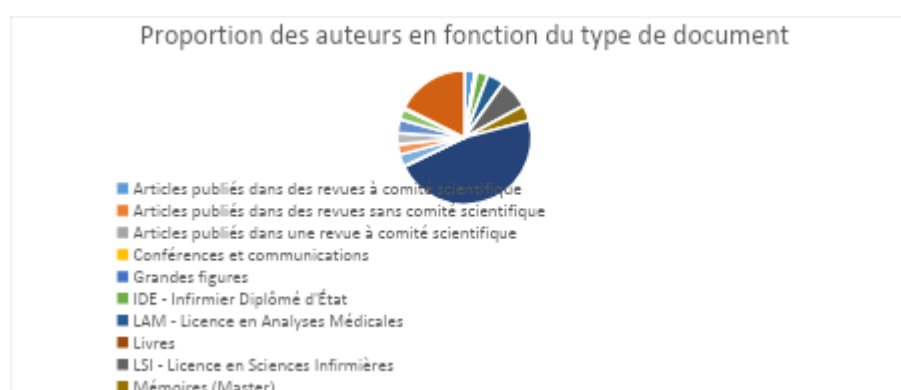


Figure 14. Proportion des auteurs en fonction du type de document dans la collection DICAMES. Données : DICAMES. Graphique : Yannick Ngalli, 2024

Malgré le fait que les auteurs aient représenté indifféremment leurs fonctions, la représentation graphique ci-dessus exprime que la proportion des auteurs par type de documents est proche de la proportion de ces documents dans la collection. La conclusion est que les mêmes critères de choix d'attribution de la qualité d'auteur ont été appliqués à tous les types de documents. Cela ne change pas les équilibres en termes de proportions dans la collection.

LES DÉFIS DE LA GOUVERNANCE DES DONNÉES

Les résultats présentés manifestent des choix de gouvernance, explicites ou implicites, structurant DICAMES. Cette discussion en tire des enseignements sur la gouvernance des données et ses implications géopolitiques.

L'analyse des métadonnées de DICAMES révèle des enjeux critiques qui compromettent son intégrité et son utilité en tant qu'outil de représentation scientifique. Les principaux biais identifiés sont les suivants: (1) la confusion des rôles sous l'entité unique "auteur" (créateur, directeur, jury), rendant les analyses bibliométriques irrecevables ; (2) les erreurs de saisie dans les titres ; (3) l'absence de vocabulaire contrôlé pour les mots-clés (plus de 14 000 occurrences non normalisées) ; (4) les incohérences typologiques (ex. : "thèses soutenues" vs "thèses").

L'absence d'alignement sur des standards internationaux tels que le format l'UNIMARC (UNiversal Machine Readable Cataloging), ou RDA (*Resource Description and Access*) limite fortement l'interopérabilité, la précision et la réutilisabilité des métadonnées. Une telle normalisation aurait permis une structuration fine des informations (données bibliographiques, descriptives, de gestion et d'accès), facilitant ainsi leur échange et leur exploitation à l'échelle internationale. Cette situation peut être interprétée comme un choix de gouvernance qui privilégie la simplicité de la collecte au détriment de la qualité des données et de leur potentiel de valorisation.

Pour garantir la qualité des données, une double validation - humaine et technique - est essentielle. Celle-ci passerait par la mise en place d'outils de contrôle automatisés couplés à une supervision par des experts métier.

L'exhaustivité limitée : une question de souveraineté informationnelle

Avec seulement 8 000 documents pour 19 pays, des centaines d'établissements et plusieurs millions d'étudiants, DICAMES représente une fraction infime de la production scientifique réelle. Ce déficit révèle une fragmentation institutionnelle et un manque de politiques incitatives. Pour devenir un outil diplomatique crédible, DICAMES doit atteindre une masse critique, ce qui suppose des circuits de collecte simplifiés, des campagnes de sensibilisation et des engagements formels avec les établissements.

DICAMES, un outil de diplomatie des données ?

La diplomatie des données (Morin et al., 2021), déclinaison technique de la diplomatie scientifique (Ruffini, 2017), recouvre les négociations et stratégies visant à contrôler et valoriser les flux de données. DICAMES pourrait devenir un instrument stratégique à trois niveaux.

Les trois dimensions diplomatiques potentielles de DICAMES

- Contrôler la narration scientifique : face aux bases internationales aux critères nord-centrés (Piron et al., 2017), DICAMES peut définir localement ce qui compte comme production scientifique (langues, types de documents, thématiques).
- Négocier des partenariats équitables : DICAMES peut servir de levier pour négocier des accords de moissonnage avec OpenAIRE, BASE, CORE, en préservant l'intégrité des métadonnées, la mention de l'origine africaine, et la réciprocité des flux.
- Renforcer la coopération Sud-Sud : DICAMES peut institutionnaliser le rôle de « hubs » informationnels (Cameroun, Côte d'Ivoire) par des accords de mutualisation, de mobilité virtuelle des données, et des communautés thématiques panafricaines.

Les obstacles actuels à une diplomatie des données effective

Quatre obstacles limitent le potentiel diplomatique de DICAMES :

- Technique : l'absence de standardisation (UNIMARC, Dublin Core) empêche l'interopérabilité.
- Politique : seuls 8 pays sur 19 contribuent, révélant priorités divergentes et méfiance envers une gouvernance centralisée.

- Épistémique : l'absence d'indicateurs stratégiques (consultations, collaborations) ne permet pas une diplomatie éclairée.
- Juridique : les législations disparates et l'absence de protocole régional sur les données fragilisent l'ambition diplomatique.

Des prérequis pour une diplomatie des données crédible

Les trois conditions préalables ci-dessous constituent une feuille de route que nous proposons :

Condition	Traduction opérationnelle
Gouvernance technique	Normalisation des métadonnées (Dublin Core, MARC21), mise en place d'une API, validation humaine et automatisée
Gouvernance institutionnelle	Signature d'une charte des données CAMES, création d'un comité de pilotage inter-états, rapports annuels sur l'état de la contribution
Gouvernance stratégique	Production d'indicateurs diplomatiques (cartes des collaborations, indice de visibilité africaine, tableau de bord des téléchargements Sud-Sud)

Tableau 1 : prérequis dans la diplomatie des données

DICAMES peut-il prétendre au soft power académique ?

Le soft power académique (Nye, 2004) désigne la capacité à influencer par l'attraction et l'exemplarité. À ce stade, DICAMES n'en génère pas : sa masse critique (8 000 documents) reste faible face à HAL (1 million), et les contributions limitées réduisent sa légitimité. Cependant, si DICAMES quadruplait ses documents, obtenait une interopérabilité et produisait des analyses bibliométriques originales, il deviendrait un outil de data power.

DICAMES possède les briques élémentaires d'un outil de diplomatie des données (assise institutionnelle, libre accès, diversité encyclopédique), mais sa transformation diplomatique est conditionnée par des choix techniques (normalisation), institutionnels (engagements étatiques) et stratégiques (indicateurs). Sans eux, il restera une archive utile sans rayonnement ; avec eux, il deviendra un modèle africain de data diplomacy.

Limites de l'étude

Cette recherche, bien qu'elle apporte des éclairages inédits sur la gouvernance des données dans les archives ouvertes africaines, comporte plusieurs limites qu'il convient d'explicitier pour situer la portée de ses conclusions. Ces limites sont de nature épistémologique, technique, méthodologique et contextuelle.

- Limites liées à la source des données

L'étude repose exclusivement sur les métadonnées, sans accès aux PDF ni aux logs d'usage, ce qui empêche de vérifier la correspondance métadonnées/contenu, d'évaluer la circulation effective des données, ou de mesurer l'impact réel. L'analyse est donc centrée sur l'offre informationnelle plutôt que sur la demande.

- Limites techniques du *scraping* et de la constitution du corpus

Le *scraping* manuel comporte des risques (omissions liées à l'architecture, absence d'API, nettoyage manuel des caractères spéciaux). La reproductibilité est assurée par le dépôt du script sur GitHub, mais une API officielle offrirait un corpus plus fiable.

- Limites de l'interprétation géopolitique

L'interprétation géopolitique des disparités (surreprésentation du Cameroun, faibles contributions sahéniennes) nécessite une enquête qualitative complémentaire. Des facteurs contextuels non mesurés (instabilité, grèves, absence de politiques nationales) ou l'effet

"institution-proximité" (hébergement initial au Cameroun) peuvent expliquer ces écarts. De plus, les productions en anglais ou langues locales échappent à DICAMES. Des entretiens seraient nécessaires pour distinguer incapacités techniques et choix politiques.

- Absence de comparaison avec d'autres archives ouvertes

L'étude est monographique et ne compare pas DICAMES à d'autres archives (AJOL, IRD, HAL), ce qui limite la généralisation. Une analyse multi-cas serait une extension naturelle. Par ailleurs, l'extraction à un instant T (octobre 2024) offre une photographie et non un film longitudinal, ne permettant pas de mesurer les dynamiques réelles ni l'impact des politiques incitatives.

- Limites temporelles et diachroniques

Les données ont été extraites à un instant T (octobre 2024). Or, DICAMES est une base évolutive. Une analyse longitudinale (plusieurs extractions à intervalles réguliers) contribuerait à mesurer la dynamique réelle des dépôts (et non un simple instantané), identifier des effets de « rattrapage » ou de « décrochage » de certains pays, évaluer l'impact des campagnes de sensibilisation ou des politiques incitatives.

Notre étude offre une photographie, non un film, sans distinction des tendances structurelles des variations conjoncturelles.

CONCLUSION

Cet article a mis en lumière le potentiel de DICAMES comme patrimoine scientifique de l'Afrique francophone, tout en identifiant des défis majeurs (biais dans les métadonnées, disparités géographiques, temporalité limitée) révélateurs d'une gouvernance implicite privilégiant la collecte à la qualité. Pour que DICAMES devienne un outil stratégique de souveraineté scientifique, deux chantiers prioritaires s'imposent : (1) la normalisation des métadonnées via des standards internationaux (UNIMARC, Dublin Core) et un contrôle qualité rigoureux ; (2) l'accroissement de l'exhaustivité par des politiques incitatives ou contraignantes auprès des établissements membres. Cette transformation exige une gouvernance active des données articulée autour de trois piliers : technique (standards et validation automatisée), institutionnelle (mécanismes incitatifs), et stratégique (diplomatie scientifique et collaborations Sud-Sud). Une stratégie de communication ambitieuse est nécessaire pour fédérer les acteurs autour de ce bien commun.

Au-delà des aspects techniques, cet article met en évidence la manière dont l'analyse bibliométrique des métadonnées révèle les mécanismes de gouvernance d'une infrastructure informationnelle, contribuant aux débats sur la souveraineté informationnelle et sur la diplomatie des données dans les pays du Sud.

RÉFÉRENCES BIBLIOGRAPHIQUES

Archambault, Éric ; Vignola-Gagné, Étienne (2004), *The use of bibliometrics in the social sciences and humanities*, Montreal : Social Sciences and Humanities Research Council of Canada (SSHRCC).

Berthaud, C.; Charnay, D.; Fargier, N. (2021). Diffuser et pérenniser le savoir scientifique : 20 ans d'histoire de HAL. Histoire de la recherche contemporaine : la revue du Comité pour l'histoire du CNRS, X(2), p. 139-149. <https://doi.org/10.4000/hrc.6330>

Borgman, Christine L. (2015), *Big Data, Little Data, No Data: Scholarship in the Networked World*, Cambridge : MIT Press.

Callon, Michel (1991), « Réseaux technico-économiques et irréversibilités », *Figures de l'irréversibilité en économie*, vol. 38, n° 1, p. 195-230.

Deboin, Marie-Claude (2015), « Déposer ses publications dans une archive ouverte, en 8 points », [en ligne], consulté le 02 mai 2026, <https://agritrop.cirad.fr/575303/>.

Fabre, Isabelle ; Gardiès, Cécile (2008), « L'accès à l'information scientifique numérique : organisation des savoirs et enjeux de pouvoir dans une communauté scientifique », *Sciences de la société*, n° 75, p. 71-84.

Morin, Jean-Frédéric ; Louafi, Selim ; Orsini, Amandine (2021), « Data Diplomacy: A New Frontier in International Relations », *Global Policy*, [en ligne], consulté le 02 mai 2026, <https://doi.org/10.1111/1758-5899.12948>.

Nye, Joseph S. (2004), *Soft Power: The Means to Success in World Politics*, New York : PublicAffairs.

Piron, Florence ; Diouf, Antonin B. ; Dibounje Madiba, Marie S. ; Mboa Nkoudou, Thomas H. ; Aubierge Ouangré, Zoé ; Tessy, Daphney R. ; ... ; Lire, Zakari (2017), « Le libre accès vu d'Afrique francophone subsaharienne », *Revue française des sciences de l'information et de la communication*, n° 11, [en ligne], consulté le 02 mai 2026, <http://journals.openedition.org/rfsic/3292>.

Ruffini, Pierre-Bruno (2017), *Science and Diplomacy: A New Dimension of International Relations*, Cham : Springer.

Schöpfel, Joachim (2020), « L'usage de la plateforme HAL par des unités de recherche : Le cas de l'Université de Lille », *I2D - Information, données & documents*, vol. 57, n° 3, p. 167-198.

Schöpfel, Joachim ; Kergosien, Éric ; Prost, Hélène ; Barrié, Julien (2022), « Pas si simple que ça... : Une enquête sur l'usage de HAL par les unités de recherche des universités IdEx », *I2D - Information, données et documents*, vol. 59, n° 2, p. 150-183.

Trzmielewski, Marcin ; Gnoli, Claudio (2019), « Une classification interdisciplinaire pour l'échange et la médiation des données ouvertes de la recherche », *Actes du 12ème Colloque international d'ISKO-France, Montpellier, 9-10 octobre 2019*.

Yin, Robert K. (2009), *Case study research: Design and methods*, Los Angeles : Sage (vol. 5).

ANNEXE

Pays contributeurs de DICAMES, les établissements associés à ces pays et le nombre de documents par établissements.

Pays	Etablissement d'enseignement supérieur	Nombre de documents
1-Burkina Faso	Centre national de la recherche scientifique et technologique - CNRST	6
	Institut National de Santé Publique	11
	Université Joseph Ki-Zerbo	101
	Université Nazi Boni	13
	Université Ouaga II	1
2- Bénin	2.1- Université d'Abomey-Calavi	151
	2.2- Université de Parakou	6
3- Cameroun	3.1- Ecole des Sciences de la Santé de l'Université Catholique d'Afrique Centrale - (ESS-UCAC)	1194
	3.2- OHADA - Organisation pour l'harmonisation en Afrique du droit des affaires	4
	3.3- Université de Dschang	52
	3.4- Université de Yaoundé I	3108
	3.5- École Supérieure des Sciences et Techniques de l'Information et de la Communication - ESSTIC	2113
4- Congo	4.1- Université Marien Ngouabi	81
5- Côte d'Ivoire	5.1- Université Alassane Ouattara	7
	5.2- Université Félix Houphouët-Boigny	328
	5.3- Université Jean Lorougnon Guédé	241
	5.4- Université Nangui Abrogoua	17
	5.5- Université Péléforo Gon Coulibaly	1
6- Egypte	6.1- Université Senghor	630
7- Mali	7.1- Université de Ségou	1
8- Niger	8.1- Université Abdou Moumouni de Niamey	72

Traitement : Yannick NGALLI, 2024.