BAVARIAN JOURNAL OF APPLIED SCIENCES®

Open Science Dynamics and Science-Technology Linkages in Africa

Roberto Cruz Romero[1] **Stephan Stahlschmidt**[1] **Jian Wang**[2]

https://doi.org/10.25929/dg7ex355
https://doi.org/10.53974/unza.zajlis.10.2.237

ABSTRACT

This study examines how open access (OA) shapes innovation activity in Africa by analysing the scientific literature cited in patents. Using open bibliometric data from OpenAlex and global patent-to-paper citation data from Marx and Fuegi (2022), we explore how OA has been involved in the scientific literature referenced by African patents, as well as characteristics of co-authorship, topic, interdisciplinarity. Our analysis focuses on a qualitative exploration of the cited literature in patents followed by a quantitative study of the influence of OA in the science-patents link. In the qualitative study we find that the African science system is well interconnected with research-intensive partners across the globe and that the applied nature of the research produced aligns with the bloc's established policy goals. Research in the region is mostly focused in key development areas, such as health and engineering sciences; the latter is true across regional and internationally connected research collaborations. OA, on the other hand, seems to still trail behind non-open sources, which may be conditioned by application processes. Furthermore, in the quantitative study we explore whether OA facilitates the knowledge flow from science to patents and its specific effect for African science and African patents. We implement a matched regression analysis based on papers published in the same journal issue (with and without OA), and we find that, compared with a corresponding closed-access paper in the same journal issue, OA papers' probability of being cited by patents is approximately 4 percentage points higher ($p < 0.001$). This effect is significantly larger for African-affiliated papers. When focusing on citations by patents from African applicants, the OA effect is substantially larger for USPTO (about 10 percentage points) but is not statistically significant for EPO, despite a similarly elevated point estimate. Within these citations, however, the OA advantage no longer differs significantly by the cited paper's own regional affiliation. These findings suggest that while OA serves as a crucial mechanism for elevating the global visibility of African science into technological innovation, its influence is mediated by the existing search and citation practices of inventors. We thus offer nuanced evidence for the development and consolidation of a dynamic open science environment; promoting OA can amplify the contribution of African research to global technological development, but strengthening the direct science-patent pipeline requires broader strategic policy interventions beyond access licensing alone.

Die Studie untersucht, wie Open Access (OA) Innovationstätigkeiten in Afrika prägt, indem sie die in Patenten zitierte wissenschaftliche Literatur analysiert. Anhand offener bibliometrischer Daten von OpenAlex und globalen Wissenschafts-Patent-Zitationsdaten

[1] Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) *(German Centre for Higher Education Research and Science Studies [DZHW])*
[2] *Lancaster University Leipzig*

von Marx und Fuegi (2022) untersuchen wir, wie OA in die von afrikanischen Patenten referenzierte wissenschaftliche Literatur eingebunden wurde sowie Merkmale von Koautorenschaft, Thema und Interdisziplinarität. Unsere Analyse konzentriert sich auf eine qualitative Untersuchung der zitierten Literatur in Patenten, gefolgt von einer quantitativen Studie zum Einfluss von OA auf die Verbindung zwischen Wissenschaft und Patenten. In der qualitativen Studie stellen wir fest, dass das afrikanische Wissenschaftssystem gut mit forschungsintensiven Partnern weltweit vernetzt ist und dass die angewandte Natur der Forschung mit den festgelegten politischen Zielen des Blocks übereinstimmt. Die Forschung in der Region konzentriert sich hauptsächlich auf Schlüsselentwicklungsbereiche wie Gesundheits- und Ingenieurwissenschaften; Letzteres gilt für regionale und international verbundene Forschungskooperationen. OA hingegen scheint weiterhin hinter nicht-offenen Quellen zurückzubleiben, die durch Anwendungsprozesse bedingt sein können. Darüber hinaus untersuchen wir in der quantitativen Studie, ob OA den Wissensfluss von Wissenschaft zu Patenten fördert und welche spezifischen Auswirkungen sie auf afrikanische Wissenschaft und afrikanische Patente haben. Wir führen eine gematchte Regressionsanalyse auf Basis von Artikeln aus derselben Fachzeitschrift (mit und ohne OA) und stellen fest, dass im Vergleich zu einem entsprechenden Closed-Access-Artikel derselben Fachzeitschrift die Wahrscheinlichkeit, von Patenten zitiert zu werden, etwa 4 Prozentpunkte höher ist (S. < 0,001). Dieser Effekt ist bei afrikanisch assoziierten Artikeln signifikant größer. Konzentriert man sich auf Zitationen von Patenten afrikanischer Antragsteller, ist der OA-Effekt für USPTO deutlich größer (etwa 10 Prozentpunkte), aber für EPO trotz einer ähnlich erhöhten Punktschätzung nicht signifikant. Innerhalb dieser Zitationen unterscheidet sich der OA-Vorteil jedoch nicht mehr wesentlich durch die regionale Zugehörigkeit des zitierten Artikels. Diese Ergebnisse deuten darauf hin, dass OA zwar ein entscheidender Mechanismus ist, um die globale Sichtbarkeit afrikanischer Wissenschaft in technologische Innovationen zu erhöhen, sein Einfluss jedoch durch die bestehenden Such- und Zitierpraktiken von Erfindern vermittelt wird. Wir bieten somit nuancierte Belege für die Entwicklung und Konsolidierung einer dynamischen Open-Science-Umgebung; die Förderung von OA kann den Beitrag afrikanischer Forschung zur globalen technologischen Entwicklung verstärken, doch die Stärkung der direkten Wissenschaft-Patent-Pipeline erfordert über die reine Zugangslizenzierung hinausgehende strategische politische Interventionen.

KEYWORDS

Open science, patent analysis, open access, innovation, bibliometrics

Open Science, Open Access, Patentanalyse, Innovation, Bibliometrie

1. Introduction

The idea of open science (OS) as a vehicle for innovative research has also been consolidated into policies that range from the supra-national (e.g., UNESCO's *Recommendation on Open Science*, 2022) to the regional (e.g., EU's *Open Science Policy*, or the African Union's *Science, Technology and Innovation Strategy*, 2024) or even the national (e.g., the German Research Foundation's plan for evaluation of proposals; or South Africa's open science policy, 2024). Hence, the advancement of OS and innovation research represents one of the most promising pathways for achieving sustainable development and consolidating global research competitiveness. The latter is an underlying factor of a landscape of intense scientific collaboration, where academic and resource mobilities tend to shape the opportunities for research areas. The examples above serve as reflection of the interconnectedness of different scientific systems. Moreover, many policies have been developed to foster and sustain a collaborative scientific exchange system that incorporates steadily available research funds as well as competitive grants.

For instance, in the context of African academic systems, there is a great variety of funding, infrastructure, or policy directives emanating from national or regional sources.[3] Yet, there appears to be a common understanding of the importance of science and technology as drivers of growth and inclusion. The Science, Technology and Innovation Strategy for Africa (STISA 2034) serves as example of the continental motivation that seeks to "call for governments, academia, civil society and the private sector, to unite in harnessing the power of science, technology and innovation (STI) to unlock the continent's full potential" (African Union, 2024, p. 3).[4] As such, it becomes clear that the path towards establishing robust OS ecosystems hinges upon a delicate balance between public funding schemes, strong university-industry (U+I) programmes and dynamic scholarly networks. However, the continental bloc still contains great disparities to be dealt with, in order to fully motivate the adoption of overarching policies. Hence, it is relevant to address, at least, three dimensions that make up a harmonious landscape for OS and innovation; first, the policy frameworks in place, then the infrastructural conditions and, finally, the collaboration linkages that can support the development of OS.

1.1. Open Science as Policy Norm

Across Africa, the policy landscape for OS remains characterised by promising, yet fragmented initiatives – "although progress has been made in terms of Open Science and Open Data practices, more awareness needs to be created and further advocacy efforts are required" (ASSAf, 2024, p. 4). For example, the African Union's (AU) *Agenda 2063* serves as a blueprint that positions science, technology and innovation as integral elements within Africa's development trajectory. The Agenda explicitly emphasises the development of information and communications technology (ICT) infrastructure for enhancing connectivity and collaboration in science, technology, and innovation. Similarly, the AU's Commission has also developed the *Science, Technology and Innovation Strategy for Africa* (STISA-2034), explicitly focusing on concepts such as open data, open innovation, and collaborative research and innovation frameworks. OS is then framed within the guiding principles of the framework and understood as a "collective benefit" (African Union, 2015, p. 6), upon which knowledge, and the progress derived from it, are commonly shared.

However, the policy landscape still reveals significant challenges in terms of specificity and scope. Many policies lack the detailed implementation mechanisms and resource allocations necessary for effective operationalisation (Dosso et al., 2022; cfr. Omini, 2019). Resources are crucial when considering the costs of publishing OA articles in specific journals from specific fields. This tension between openness and intellectual property protection extends beyond publishing costs to patenting itself, where faculty and institutional patenting practices have long had to balance commercialisation incentives against open science norms (Fabrizio & Di Minin, 2008). Generally, costs for accessing published materials are often a hurdle that declarative norms usually overlook (Bezuidenhout et al., 2017), as this tends to be a problematic dimension of OS in terms of material access to science. Issues of institutional support for OS, for example regarding university libraries, highlights the key gaps between policy and practice. Further dimensions of access are left unexplored and, thus, omitted from the potential benefits into the scientific, economic, and technological (Wang et al., 2025).

Policy alignment is therefore crucial to overcome the atomisation of academic systems, particularly where potential for greater interlinking is present (ASSAf, 2019). *The African Open Science Platform* (AOSP), for example, is a precise indication of the already existing efforts to drive OS as a key component of development and equity through access to knowledge (Participants of African Open Science Platform Stakeholder Workshop et al., 2018).[5] Whether this objective can be achieved

[3] See the African Union's Agenda 2063 – a comprehensive set of policy aspirations that set out to influence the continent's policy-making contexts into adopting more or less harmonious steps towards sustainable development. The Agenda is designed along the lines of 10-year plans and, in the case of science and technology – as with all other fields of political-economic relevance – it is set to overhaul the system and "[b]uild and expand an African knowledge society through transformation and investments in universities, science, technology, research and innovation" (2015, p. 15)

[4] STISA 2034 is the second presentation of such policy framework. It was formulated as an action plan in accompaniment of the already mentioned Agenda 2063; the STISA 2034 represents thus the second 10-year iteration of the policy intended to incentivise science and technology development in the coming years.

[5] See https://openscience.africa/ for a complete overview of several initiatives for OS in the African continent.

depends on the coordination strategies of governments and regional entities, as they can promote the construction and/or consolidation of necessary technical infrastructure. The policy landscape described above provides the motivating context for our empirical analysis, although we do not operationalize specific policy instruments directly. Instead, our empirical approach relies on OA status and regional (African, USA, EPO) affiliation as observable proxies for the underlying access and infrastructural disparities that these policy debates address, built on the premise that some degree of coordination benefits the knowledge flows from science to innovation, facilitated by the policies on OA and collaboration.

1.2. Collaboration and Partnerships as Drivers for Innovation

International collaboration has emerged as both a critical enabler and a complex challenge in Africa's pursuit of OS consolidation. The patterns of collaboration between African institutions and international partners, particularly European universities and research organisations, demonstrate the transformative potential of strategic partnerships whilst highlighting the need for more balanced and sustainable collaborative frameworks. Similar to the detrimental effects of *brain drain*, Mwelva et al. (2020, p. 10) problematise some of the efforts for North ↔ South cooperation as "helicopter science". The latter means that "collaboration with global north partners, funded by northern agencies, are frequently dominated by northern scientists, who fly in, collect data from their African partners, then fly out". This framing sets the stage for an extractivist approach to research and shuts down many pathways for effective collaboration.

However, numerous initiatives are in place to bolster international collaboration in a matter that acts as a structural support and not dependence mechanism. For instance, both AU and the European Union (EU) issued a so-called *Joint Vision for 2030* in early 2022. The declaration explicitly recognises the importance of research and innovation in the bilateral relationship. The policy seeks to foster capacity building in relation to the design and implementation of ICT solutions as one of the most crucial areas of policy action acknowledged by most assessments (see ASSAf, 2019; Vallejo et al., 2019).[6] Additionally, the consolidation of a strong intellectual property regime that recognises the role of local research bodies that facilitate or indeed carry out relevant parts of data gathering or analysis processes.

A more symbolic advantage gained from intense collaboration networks is seen in the knowledge production and its later recognition. As per other regional and global trends, co-authored works (particularly those with inter-regional collaboration) tend to be cited increasingly often (Huang et al., 2024; cfr. Iyandemye & Thomas, 2019). While OS encompasses a broad range of practices, including open access to publications, open data, open methods, and open peer review, our analysis turns specifically to open access (OA): the free, immediate availability of published research articles. OA increases the visibility of scholarly works to broader audiences (Nishioka et al., 2020); for instance, considering the visibility gains in the context of innovation and applied research, leading to increased citations in patents (Jia et al., 2026; Maddi, 2024). Considering the framework to bolster OS in the African contexts, the dissemination of OA content signifies a great tool for addressing the (still unresolved technical and financial) inequalities of the academic landscape (see Bryan & Ozcan, 2021). So, for illustration, cases like the Scielo repository in the Latin American case, may also function as a reference point from which to guide specific policy implementations. This focus may lead to offset the otherwise positive concentration of partnerships with European institutions, reinforcing instead South ↔ South collaborations.

1.3. Research Aim

Following the theme of innovation and the surrounding elements that determine a successful OS environment, the research seeks to connect the corresponding OA dimension, in terms of published research articles, with an innovation-oriented perspective. We examine how OA is associated with knowledge diffusion from science to technological innovation, and how these associations are contingent on region, particularly Africa. Thus, this contribution focuses not only on the OS environment conditions, but also on the political-economic implications that give way to the uptake in OS practices that push innovation. We therefore

[6] In addition to the already mentioned Recommendations on Open Science, UNESCO has also elaborated guidelines on Open Educational Resources (OER) and on Science, Technology and Innovation Policy Instruments.

also focus on how open innovation (OI) – the intertwinement of OS practices within a dynamic innovation environment – represents the specific political-economic (structural) conditions that determine the intensity of public-private co-production in innovation activities (e.g., through co-authorship or funding).

Recent work has begun to map how scientific research feeds into technological innovation through patent citations more broadly (Dorta-González et al., 2025), and we extend this line of inquiry by centring the OA dimension of this science–patent link within the African context. Through a coupling of bibliometric and patent data sources, we look to analyse the knowledge flows from scholarly literature into the knowledge structure of patents. Particularly, we are interested in assessing the bridging dynamics that OA literature may have in relation to patents' references. From this perspective, we analyse how the OA transformation has affected innovation processes in the diverse African countries and, from this analysis, highlight the enabling factors that socio-economically or from an academic regulatory perspective influence the knowledge flows, e.g., collaboration practices. These dimensions inform policy action from both academia and public sector officers, providing evidence for objective-oriented initiatives in the research and development (R&D) and innovation sectors.

2. Data and Methods

The empirical approach follows a descriptive-inferential coupling of the scholarly and patent data. At first instance, we conduct an exploratory analysis of different data sources in order to fully characterise the knowledge flows and their scholarly nature. Whilst, at a second stage, we conduct a matched regression analysis based on a stylised model of publication data focusing on the differences between papers published in hybrid journals. Together, these two lenses offer a complementary outlook into the trends and conditions of an expanding OS environment, namely scientific publishing and innovative research.

Regarding the descriptive exploration, we use open bibliometric data from the OpenAlex (OAL) database and open patent citation data from the Marx and Fuegi (Marx, 2024; Marx & Fuegi, 2020; see also Marx & Scharfmann, 2024; Abbasiantaeb et al., 2025). We complement these data with additional patent data from PATSTAT with affiliation data for each patent applicant, offering more fine-grained information on affiliation patterns. Further, we focus on the literature being cited in patents, whether front page or in-text, and its metadata. To this regard, we use this distinction only for the descriptive section. Because we are interested in identifying the OA-based transfer of scholarship through patents' knowledge structure, we aim to characterise the extent to which affiliations in African countries are present in this dataset, alongside collaboration patterns and disciplinary distributions. The affiliation metadata in both patent and OAL datasets allow us to filter to only African countries in the patents' applicants list and, thus, to focus only on the scholarly items cited in these patents. This includes all listed applicants; that is, regardless of their legal persona (individual or companies) – in this sense, we capture all effects in the patent landscape.

We thus matched these two datasets through the OAL (unique) identifiers and obtained a sub-set with complementary information about patent applicants and academic researchers. Given that OAL item metadata also includes affiliation information, we are also able to complement the data and create a patent-to-paper link on two levels: a) one describing the affiliation and collaboration network of the patent applicants, and b) another describing the same features for scientific articles. The filtering of the patent data is done according to the applicants' country data, and the scholarly items are then matched through the citation keys. The dataset comprises patent publications published between 2000 and 2023 citing scholarly articles published as early as 1907 and up to 2021.

We narrow down further to focus on a specific type of OA licensing, namely the hybrid type. Hence, we group match the journal and issue metadata and extract all items from journal issues where at least one item labelled as hybrid was published (that is, an OA item published in a subscription-based journal). So, we keep all patent-cited articles, and then for each patent-cited article we randomly sample a non-cited article in the same journal issue, to construct an analytical (matched) dataset that is balanced in terms of predicted outcome, i.e., being cited by patents. The rationale for this step is that being cited by patents is rare, which makes standard statistical analysis problematic.

This approach allows us to identify the effect of OA publishing, i.e., the unrestricted accessibility of scientific publications, on transfer of the respective content into patents while controlling for other confounding factors, such as fields of study. Figure 1 visually exemplifies the features for each level of analysis and the structure of the data.

Thus, the a) qualitative exploration is based only on papers cited in African patents,[7] which have been published in the hybrid journals. This intersection allows us to characterise the African innovative R&D landscape. The b) quantitative study, on the other hand, uses the matched hybrid journals dataset, i.e., all papers cited in the patent dataset (regardless of the patents' location). The statistical estimation seeks to provide some evidence regarding the affiliation of authors, the patent offices (i.e., USPTO [United States Patent and Trademark Office] or EPO [European Patent Office]), and the interaction of OA and institutional affiliations (grouped at regional level, controlling for the author affiliation in, e.g., African institutions). More on this below.

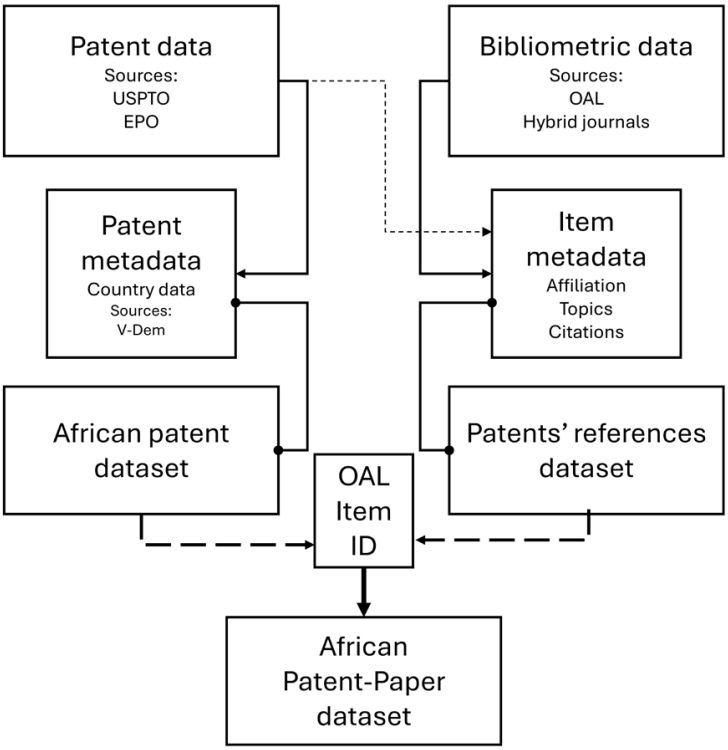

Figure 1: Patent and bibliometric data workflow.

Table 1 shows a quick summary of the filtered data on African patents (i.e., with at least one African institutional affiliation) based on the corresponding patent office, as well as the access type for the cited papers. The access status is given according to OAL's classification, distinguishing between closed and open (and from the latter we highlight the share of hybrid items as they represent our feature of interest for the logistic regression below).

Category	n	%
Patents	2239	100.00
USPTO	*1596*	*71.28*
EPO	*643*	*28.72*
Papers cited	10945	100.00
Closed	*6650*	*60.76*
Open	*4295*	*39.24*
Hybrid*	***546***	***4.99***

Table 1: Summary count of patent and papers (from patents with at least one African affiliation)

[7] By African patent publications we refer to the patents where, grouped by each patent number, at least one listed applicant's country affiliation matches any country in the African continent. Given the format of the country data (ISO-3166-2 two-lettered country codes), for the specific case of Namibia (country code "NA") there was a manual check and a character string matching to avoid false positives (as with no data NA).

2.1. Data Description

Marx and Fuegi's dataset includes data from many patent offices around the world, yet we filtered and matched only the USPTO and the EPO offices due to the concentration of counts for these two entities. Patents are identified by their filing/granting office, coded as US (USPTO) and EP (EPO) in both Marx and Fuegi's dataset and in PATSTAT, from which we draw applicant affiliation data. We describe the location data of the references cited in each patent; i.e., if they are cited in the front page or in the body of the text (or both). Following our approach, we then filter to only the patents with an African(-based) applicant, looking to further delineate the role of affiliation in relation to OA citation. Hence, our dataset for the *qualitative* exploration comprises nearly 60 thousand patent-paper reference links, from 2,239 unique patents and 10,945 unique cited papers.

For the *quantitative* study, based on the matched regression analysis (subsequent section), we make use of the matched hybrid subset, which represents a total of 2.6 million cited articles. These cited items make up a further filter, as we are interested in matching cited items with non-cited items from the same journal issues – by construction only hybrid journals. Thus, after matching complete cases from the respective journals, we come up with a balanced sample of 4.4 million citation links (more below). The matching is done on the basis of the closed/open binary, with the hybrid option providing the control filter. That means that the estimation approach only distinguishes if a published item is OA or not, in relation to the matched sample of articles published in hybrid journals.

Figure 2, an explorative approach of the qualitative dataset, shows a summary of different temporal dynamics for each group (patents and paper citations), and within each of the categories of citation (only in the front page, only in the text, or in both). We then couple these features with the access category for each cited paper. The provenance of front-page citations differs by patent office: at the EPO, front-page references are added exclusively by examiners during the search report process, whereas at the USPTO front-page references are primarily supplied by applicants via the Information Disclosure Statement, with examiners adding further citations during examination. Given this cross-office heterogeneity in who supplies front-page citations, we pool both citation locations within each office-specific regression model, as our interest lies in the overall likelihood of a paper being cited in a patent's knowledge structure rather than in attributing citation behaviour to a single actor. At first glance, a small lag between item citations and patent publication is observable, which speaks of a preferential tendency for more recent literature (despite the long tail of older papers). Moreover, following up on the shares of OA papers cited in general, despite the overall predominance of non-OA papers cited, the trend for OA papers increases with time (inversely: there is an apparent down take in citing non-OA references).

Although our interest lies in the internal comparison of hybrid journals,[8] in terms of their OA publications, and their influence in patent citations we nonetheless call to attention the OA developments after the year 2000.[9] In part, this is relevant to observe the patent trends (increasing from 25 in the year 2000 to over 250 by the year 2010), with a similar pattern following the total number of patent citations as well as the share of OA articles cited (increasing from a share of around 30% OA citations in 2003 to over 40% after 2010). The lower panels (e) also reflect the distribution of OA references (vis-à-vis non-OA) in each reference location. Though this trend is rather influenced by the role of patent examiners in accessing and incorporating literature (e.g., the difference between front page and in-text shares – see Bryan & Ozcan, 2021). The latter may reflect a preference for more consolidated research and/or consolidated journals – OA would represent an afterthought. Though, as shown below, this trend is correlated with the time dimension, as the usage of OA references has grown in the last years of the sampled data (where in-between access licenses are still the most common; i.e., *green* and bronze *models*).

As for visibility and ease of *findability*, front page literature is assumed to hold more general research meant to show the innovation potential (novelty) of applications, whereas in-text reference may contribute to specifics of the patent process (Bryan & Ozcan, 2021; cfr. Wang & Verberne, 2024). This could be the reason for

[8] The number of hybrid items (i.e., published in hybrid journals) reflects an under-counting, evidencing how patent publications from teams with at least one African institutional affiliation opt for research published mainly in non-OA, subscription-based journals.

[9] The latter is meant to reflect the influence of policies, in particular that of the Budapest Open Access Initiative from 2002, where signatories sought to push OA to scientific publications through self-archiving and OA journals.

the positive tendency towards the use of more OA references in the front pages of patents than in the text body (*e*). The ratio for references cited in *both* locations appears more balanced, in addition to cover a small overlapping share, as shown by the blank periods in this category. In general, there appear to be considerable differences between front-page and in-text referencing practices and the type of referenced sources. And, whilst non-OA seems to still be the largely dominant type of cited literature, intermediate licensing models seem to gain terrain.

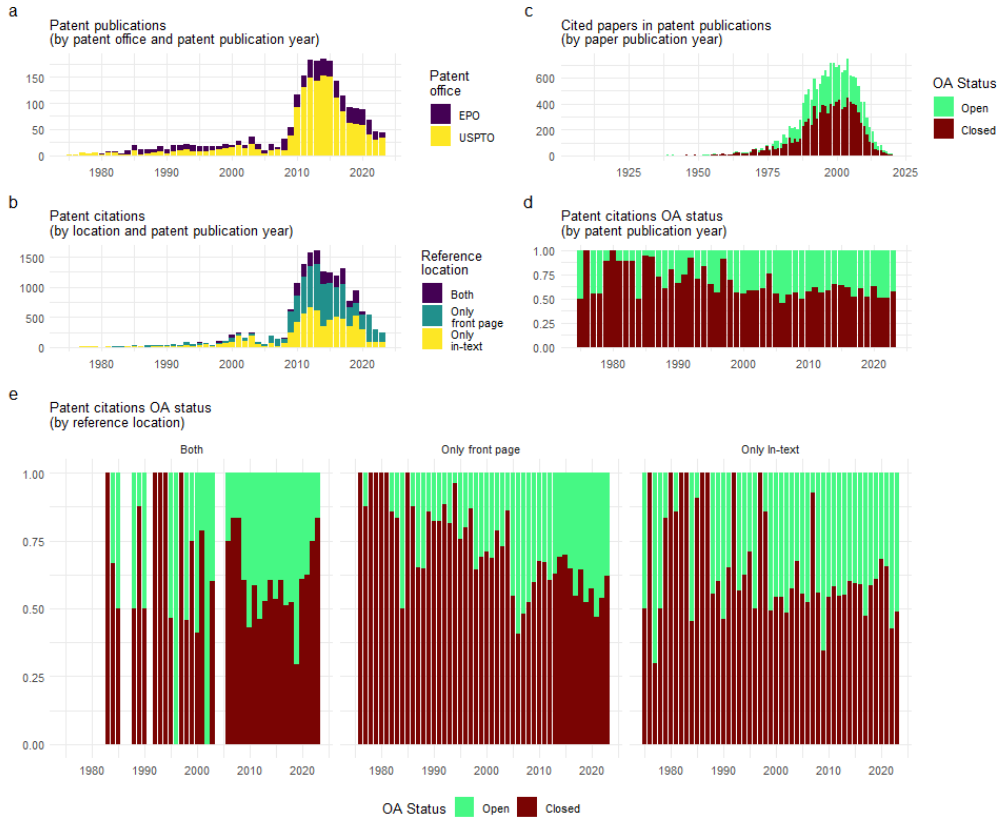

Figure 2: Patent-paper dynamics (by publication year, reference location, and reference access type, from patents with at least one African affiliation).

3. Results

3.1. Qualitative Insights

We now focus on a qualitative exploration of the main features of the scientific contribution of scholarly works cited in patents. To reiterate, the data contains the nearly 11 thousand papers cited in the over two thousand patents from African institutional affiliations (or at least one given multiple applicants). First, we narrow down the affiliations of the papers' authors in order to assess the extent to which collaboration patterns appear. Expectedly, the USA stands out as the predominant host of scholars whose work is cited in patents. With nearly 45% of all authors based in the USA,[10] the landscape seems skewed towards this scientific system, where grant searching and patents are based on highly competitive paradigm (Frakes & Wasserman, 2021; Graham et al., 2015). Moreover, when compared according to the authorship composition (single- vs. co-authored), a similar trend is maintained, reflecting the dynamics of cooperation that endure. Notable shifts occur in some positions regarding the number of collaborators; it may be indicative of national levels of collaboration based on specific working groups. USA-affiliated scholars show a relative advantage, as their work gets national and international attention given the language feature.

[10] For comparison, this proportion is much higher than the USA share for overall publications in the WoS or Scopus, for instance. For a detailed breakdown, see: https://bibliometrie.info/forschung/arbeitsberichte/.

Furthermore, there is approximately an average of four co-authors per paper in the entire dataset, showing a tendency towards intense collaboration from the academic side. In general, around 10% of the cited papers are single-authored, with a large majority of multi-authored papers (and a few with quite high author counts, with some few over 100). Similarly, the affiliation per organisations (universities or research institutions) follows a similar pattern, with at least one paper with ~60 authors affiliated in respectively different institutions. There is a wide variety of individual affiliations, which goes to show the interaction of regions and disciplines. So, in order to make sense of the affiliation patterns in collaborative works, we group countries by their geographic region (in order to avoid a very clear over-representation of the USA given their weight in geopolitical distributions) and match the co-authorship links for each cited paper.

Table 2 shows how the regions relate to each other in terms of their conditional probability for collaboration. That is, for each row, we show the normalised counts of fractional authorships for papers with more than one author (around 90% of cited papers). By estimating the conditional probability, we see the likelihood of one region collaborating with others (e.g., how likely is that an author affiliated in an institution in Western Africa appears in a paper also co-authored by a scholar affiliated in Eastern Africa – in our data, the probability is close to zero). Different to a cosine similarity measure, the conditional probability indicates the relative strength of that co-appearance. In addition, the table also gives indication to our assumption of intense internal collaboration, as regions collaborate more strongly inwards, suggesting a strong spatial determination to scientific work relevant for patents; in particular if that work serves an innovative purpose that would feed into innovation realisation processes (i.e., patents). This phenomenon offers contradicting evidence for the robust policy frameworks in place in

Region	Northern Africa	Middle Africa	Western Africa	Eastern Africa	Southern Africa	North America	Central America	Caribbean	South America	Northern Europe	Western Europe	Eastern Europe	Southern Europe	Eastern Asia	South-Eastern Asia	Central Asia	Southern Asia	Western Asia	Oceania
Northern Africa	*0.62*	*0.00*	*0.00*	*0.00*	*0.00*	*0.07*	*0.00*	*0.00*	*0.00*	*0.01*	*0.23*	*0.00*	*0.03*	*0.00*	*0.00*	*0.00*	*0.00*	*0.01*	*0.01*
Middle Africa	*0.00*	*0.48*	*0.01*	*0.00*	*0.05*	*0.08*	*0.00*	*0.00*	*0.00*	*0.09*	*0.25*	*0.00*	*0.04*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Western Africa	*0.00*	*0.01*	*0.54*	*0.00*	*0.02*	*0.19*	*0.00*	*0.01*	*0.00*	*0.07*	*0.12*	*0.00*	*0.03*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
Eastern Africa	*0.00*	*0.00*	*0.00*	*0.47*	*0.06*	*0.17*	*0.00*	*0.00*	*0.00*	*0.13*	*0.13*	*0.00*	*0.01*	*0.01*	*0.00*	*0.00*	*0.01*	*0.00*	*0.00*
Southern Africa	*0.00*	*0.00*	*0.00*	*0.01*	*0.73*	*0.08*	*0.00*	*0.00*	*0.00*	*0.07*	*0.06*	*0.01*	*0.01*	*0.00*	*0.00*	*0.00*	*0.01*	*0.01*	*0.01*
North America	0.00	0.00	0.00	0.00	0.01	0.89	0.00	0.00	0.00	0.02	0.03	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00
Central America	0.00	0.00	0.00	0.00	0.00	0.13	0.66	0.00	0.09	0.02	0.04	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.02
Caribbean	0.00	0.00	0.03	0.00	0.00	0.11	0.00	0.72	0.04	0.04	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
South America	0.00	0.00	0.00	0.00	0.01	0.09	0.03	0.00	0.73	0.04	0.04	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.01
Northern Europe	0.00	0.00	0.00	0.00	0.02	0.08	0.00	0.00	0.00	0.79	0.05	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01
Western Europe	0.01	0.00	0.00	0.00	0.01	0.07	0.00	0.00	0.00	0.03	0.83	0.01	0.02	0.01	0.00	0.00	0.00	0.01	0.00
Eastern Europe	0.00	0.00	0.00	0.00	0.01	0.13	0.00	0.00	0.00	0.03	0.10	0.67	0.02	0.01	0.00	0.01	0.00	0.01	0.00
Southern Europe	0.01	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.01	0.03	0.07	0.01	0.78	0.01	0.00	0.00	0.00	0.00	0.00
Eastern Asia	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.88	0.01	0.00	0.01	0.00	0.00
South-Eastern Asia	0.00	0.00	0.00	0.00	0.01	0.10	0.00	0.00	0.00	0.03	0.05	0.00	0.01	0.08	0.67	0.00	0.03	0.00	0.01
Central Asia	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.48	0.00	0.00	0.00	0.48	0.00	0.00	0.00
Southern Asia	0.00	0.00	0.00	0.00	0.02	0.07	0.00	0.00	0.00	0.03	0.04	0.00	0.00	0.04	0.01	0.00	0.77	0.00	0.01
Western Asia	0.01	0.00	0.00	0.00	0.02	0.13	0.00	0.00	0.00	0.04	0.10	0.01	0.01	0.01	0.00	0.00	0.00	0.65	0.00
Oceania	0.00	0.00	0.00	0.00	0.01	0.08	0.00	0.00	0.00	0.04	0.02	0.00	0.01	0.02	0.00	0.00	0.01	0.00	0.80

Table 2: Conditional probability matrix for co-authored papers

the African bloc for incentivising science and technology advancements.

Moreover, in order to visually understand the flows of cooperation links that make up the literature in this sub-sample, we build a network of authorship co-occurrence in the papers. This relationship counts the joint share of participations per the co-authorship entities named. In this case, we group the affiliation data from all papers by countries and then by the respective geographic region, adding up the

fractional counts of author attribution in each one. So, we estimate the paper-region matrix (M, where $X_t X$ is the matrix) and then the fractional shares assigned to each region by summing the products of the shares pairings ($M_{i,j} = \sum pX_{p,i} \times X_{p,j}$, where $X_{p,i}$ represents the fractional share of paper p for region i). Further, we overlay the network onto a world map projection so as to better communicate the relationships – the nodes in the figure represent each region's centroid weighted by the participation share (i.e., the number of co-authorships).

Complementary to the conditional probabilities, the network visualisation offers a more intuitive approach to understanding collaboration patterns, as it shows the connections between regions. The nodes are representative of each region, whilst the width of the edges connecting them is indicative of the strength of the estimated links. This visualisation serves also as evidence of the strong centrality of the so-called Global North regions (North America and Western Europe represent the biggest nodes in this network arrangement). However, given the scope of the sample analysed, we see a relative importance of Eastern and Middle African regions in terms of collaborations. These regions not only collaborate intensely within, but also between regions. For instance, plot **b** of the figure below shows only the flows coming from or directed to African institutions (affiliations). As such, the African continent has an important relationship with the most research-intensive regions; in particular, the Southern African region where, admittedly, the nation of South Africa takes most of the credit, as being one of the most research-intensive countries in the region and in the continent. Notwithstanding, there appears to be a knowledge link between the collaborative work in the different African regions and the patents. From a policy perspective, this entails

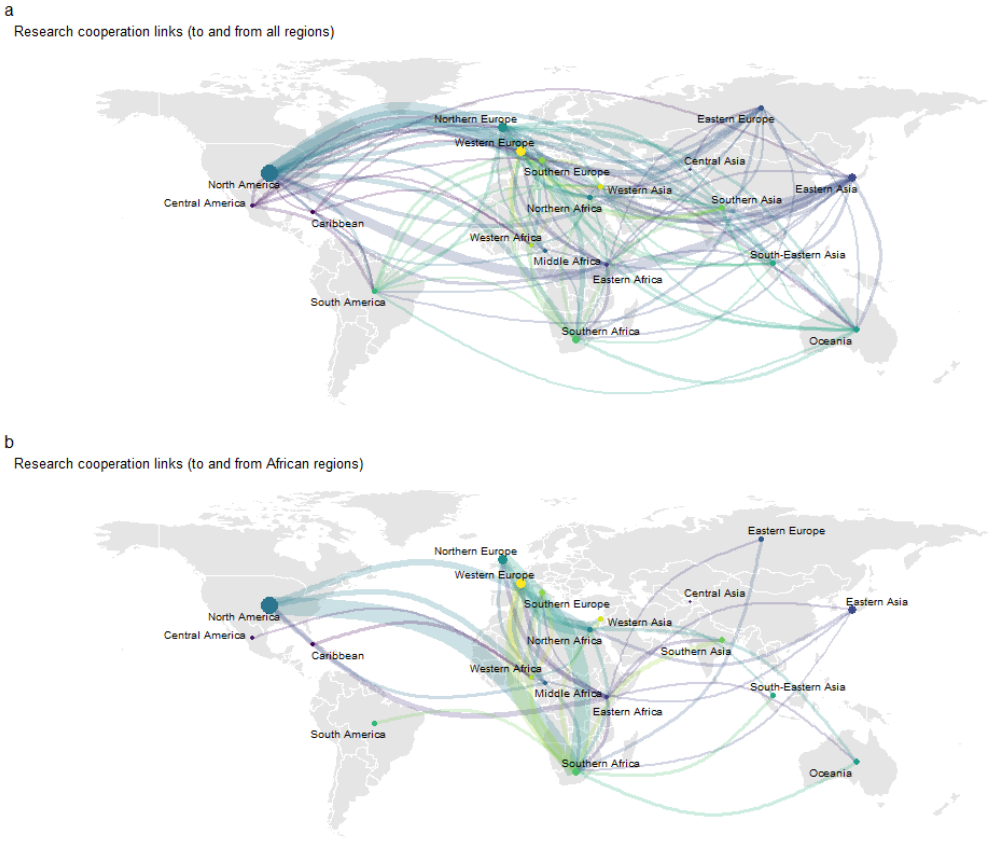

Figure 3: Cooperation patterns in papers cited in African patents.

great promise, as it poses the African context in a strategically relevant position of research collaboration.

3.1.1. Topics

Notably, a thematic subdivision of the cited papers offers indication of specific thematic directions, as a substantial part of the affiliate institutions show a tendency towards the technological and health sciences fields. One of these cases in best exemplified with the *Centers for Disease Control and Prevention* (CDCP) from Kenya, or the *Council for Scientific and Industrial Research* (CSIR) from Ghana, with numerous collaborations in the cited papers. Linking this back to the context of African science and technology policy, the CDCP is an example of a medical and health sciences institution that is also a reflection of one of the key policy areas sketched in the *Agenda 2063*, namely, public health.

Following this intuition, we look at the field distribution in the cited papers from African patents. This is relevant twofold: a) these are the papers cited directly by the patents, and b) these citations show the multidisciplinary basis underpinning specific patenting processes. For instance, organisational studies, psychological assessments, or even economic analyses can be part of the prior knowledge for a patent and, thus, we emphasise the need to comprehensively assess the broad variety of these sources. For that, we group the papers per patent (the number of papers cited in each patent) according to the OAL topic classification (as per OAL's domain and field names). In addition, since we are parting from this perspective (science flows into patents), we underline the value of the academic classifications over the structured patent typologies. We seek to better understand what types of topics are entering the patents' knowledge base regardless of where the patent is classified or what type of innovation is being filed.

Figure 4 shows that a strong indication of topics directed towards health and medicine sciences dominate, both in magnitude and the distribution of counts. The general field of medicine is followed by biotechnology and then by immunology; fields that correspond to an emphasis on public health and cost-effective health solutions. The case of e-medicine comes to mind as an example of an area that can profit from this thematic emphasis (e.g., the case of telemedicine in the wake of COVID-19, Dodoo et al., 2021). Moreover, the prevalence of medicine-focused research remains even when separating the sample into the OA and non-OA sides.

Likewise, in the fields of engineering and agricultural sciences, the energy sector has seen efforts made to advance the research & development (R&D) and implementations relating to efficiency and sustainable development (Ahodode, 2023). In this same line, this pattern follows the objectives stated in the STISA, as the push for collaboration in key areas of science and technology seems to be the driver of academic output. Similarly, the OA feature does not seem to greatly dilute the distribution, as it could be said, that it is the preponderance of OA literature that pushes the distribution. OA papers seem to be cited even more in the key fields named before, although the non-OA cited papers come from a wider variety of disciplinary fields. As for African patents, the topic distribution does not differ greatly, with the large majority in medicine and the life sciences. The social sciences remain far behind in the disciplinary contribution to

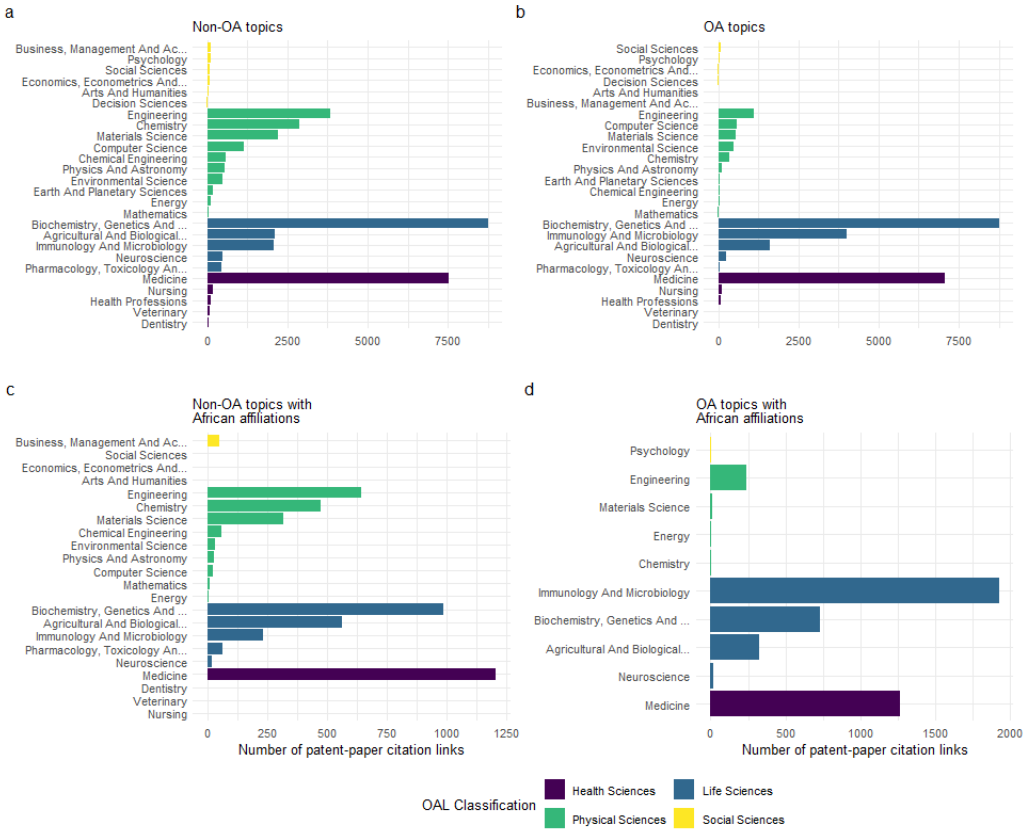

Figure 4: Topic distribution in papers cited..

patents, however, the order and proportion of fields of study seen for African-affiliated researchers differs slightly, which may suggest some more precise objects of study.

3.1.2. Open Access

Moreover, carrying on to the specific OA characteristic of the literature cited (central to the argument for OS), we find that it remains largely closed. This presents the full picture as to what literature is mostly found and used in specific fields; as noted above, a structural predisposition for non-OA may play a role for patent applicants and examiners alike. Yet, the OA dimension is relevant for fields of medicine and biotechnology, arguably the largest most research intensive and funded fields. Despite a broad trend for non-OA literature, sub-fields such as immunology and biotechnology – key political-economic areas of development for the African regions – already show an overtake of OA in patents' cited sources. We do not hypothesise on this regard but highlight the disciplinary differences that underpin our approach. That is, in most fields of study, a large share of the knowledge produced and, consequently, being cited in patents remains closed (under subscription or pay-to-read licenses). In terms of our argument, this is preliminary proof that policy declarations are still somewhat decoupled from effective OS results. Further, we only show OA vs. non-OA groups in order to more clearly present these trends.

The plots below (Figure 5) reflect the total and yearly distributions of the shares of OA papers being cited in African patents. Closed (non-OA license) is the most common access model found in the sample. It seems, however, that in recent years there has been an inverse trend where, proportionally, OA formats catch up with the non-OA – this is more clearly seen in the health sciences and in the life sciences, where field-specific journals transition into broader OA licensing schemes (with gold or hybrid formats – not shown separately; similar to the trends observed in more general contexts – see Cruz Romero & Stahlschmidt, 2025). This trend

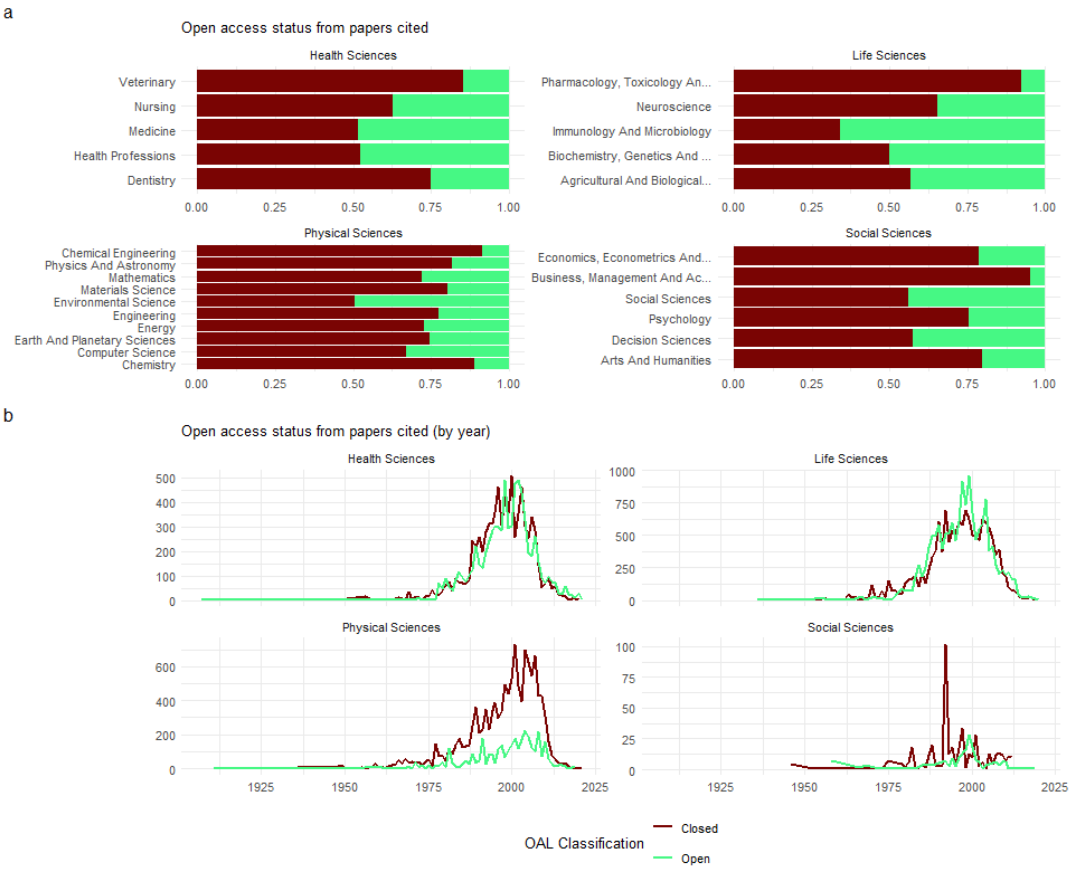

Figure 5: Open access information in papers cited.

is part of a global movement driven partly by *transformative* agreements (e.g., DEAL) and, in the case of many African countries, with the transfer effect of initiatives like Research4Life.

3.1.3. Interdisciplinarity

Now, moving to the relative difference in these fields' distribution, we focus on the concentration of fields per patent. To address this, we estimate an interdisciplinary distance measure, which takes a normalised cosine similarity and maps it on to a matrix. Fundamentally, we set this as the difference between each field's appearance in the topic distribution for each patent's cited papers (it takes the form of *Dist = 1 - norm(Cosine Similarity*) – again, as per OAL's topic classification). The goal is to estimate the degree of disciplinary differences encompassing each patent's knowledge base. The patents in the data have approximately 51 citations on average (the minimum being one and the maximum being 249, with an interquartile range of 56). This means that an important proportion of the patents cites just two or fewer papers (1,187 – over half of the patents), and around two-thirds of the patents cited ten (10) or less papers (2,185).

Interestingly, there are notable patterns relating to a clustering effect of specific subfields. For example, chemical engineering, business and accounting, and health profession seem to be the most distanced in terms of being cited within a same patent. That is, these disciplines are mostly cited alongside papers with a similar disciplinary classification. The latter hints at a strong intra-disciplinary co-citation. On the contrary, biochemistry, genetics and molecular biology appear rather close to agricultural and biological sciences, suggesting a close co-citation of papers within these fields. A similar behaviour characterises the life sciences and medicine fields. Noteworthy though, the social sciences seem to be closely related to most other fields; this may be due to misclassification or, as seen from above, business or economics papers being cited in common with medicine or

engineering papers.

Furthermore, Table 3 summarises different scores of diversity and/or concentration in the scholarly papers cited by (African) patents. We first estimate a Herfindahl–Hirschman Index (HHI) for measuring the concentration of topics across the overall distribution. Secondly, we plainly invert the HHI to link it to diversity (as per Gini-Simpson). Finally, we estimate a Rao-Stirling (RS) diversity score for the topic distribution (Stirling, 2007; cfr. Wang et al., 2015; Wang & Verberne, 2024). Both the HHI and RS scores are meant to indicate how diverse or not topics are spread across fields and, in this case, in their citations in patents. All values range 0–1. The results indicate a relative diversity; whereas the HHI (and the inverse HHI) signal to a high diversity of topics in the papers cited by patents, this must be cautiously acknowledged based on the previous insight. That is, a few fields seem to be cited almost exclusively (chemical engineering), whilst others – rather dissimilar – appear to be more closely linked. In this sense, the RS score is indicative of such phenomenon. This suggests that there is a moderate diversity possibly weighted by variety of fields (thus the apparent closeness between humanities and business and accounting, for example). Then, it seems that topic citations in

HHI	Inverse HHI	RS
0,17	0,83	0,30

Table 3: Diversity scores in patent citations

patents may be interdisciplinary and distributed across multiple fields, but those fields tend to be more related/adjacent rather than spanning highly dissimilar domains

3.2. Quantitative Insights

The second stage of our analysis is based on an inferential approach to the patent-to-paper data. This set of analysis is not restricted to African patents and their references but instead examines linkages between all scientific articles and all patents across geographic regions. Again, the empirical base for this approach is a large paper-patent citation link dataset, based on the matching for OA and non-OA articles in hybrid journals. Thus, we are interested in whether OA facilitates the knowledge flow from science to technological innovation, for example by facilitating the diffusion of open publications in general and/or by making the open publications more visible when applicants or examiners search for relevant prior art at various stages of the innovation and patenting process.

We aim to examine (1) whether OA increases the likelihood of a scientific paper being cited by patents and, more importantly, (2) how this association is stronger (or weaker) for African papers, and finally (3) how this association is stronger (or weaker) for African patents. One key feature of our empirical strategy to incorporate journal issue fixed effects. In other words, we are comparing open vs. closed access articles in the same journal issue, ruling out any differences in terms of topic, timing, journal visibility at the fine-grained journal-issue level. Following this empirical strategy, by construction, our sample restricts to hybrid journal issues where both open and closed articles are present.

Moreover, being cited by patents is a rare event, which may lead to biased estimates in traditional statistical models. Accordingly, we keep all patent-cited articles as the treatment group, and for each of them we randomly sample a non-cited one in the same journal issue for the control group, yielding a balanced dataset for analysis. As noted, this analysis is restricted to hybrid journal issues where there are both open and closed access article present. The final analytical sample consists of approximates 4.4 million articles.

3.2.1. Variables and Descriptive Statistics

Table 4 reports the descriptive statistics for each variable included in our model. The first four variables are our dependent variables; these are binary variables and represent whether a paper is cited by any patents in the USPTO or the EPO offices: *I(USPTO), I(EPO)*, and if they are cited by patents from African applicants in USPTO or in EPO patents: *I(African patents in USPTO), I(African patents in EPO)*, respectively. In our setting, Open Access is the key explanatory variable, more specifically, it indicates whether

a paper is open or not. We track whether a paper includes affiliations from Africa, EPO member states, or the United States: *I(African affiliation), I(US affiliation), I(European affiliation)*.[11] We also control for the number of co-authors, countries, and references, which take natural logarithm for regression analysis in order to control for large absolute differences. The specification is as follows:

$$I(Cited\ by\ patents) = \beta_0 + \beta_1 \cdot Open_Access + \beta_2 \cdot log(Countries) + \beta_3 \cdot African_Affiliation + \beta_4 \cdot US_affiliation + \beta_5 \cdot European_Affiliation + \beta_6 \cdot (Open_Access \times African_Affiliation) + \beta_7 \cdot (Open_Access \times US_affiliation) + \beta_8 \cdot (Open_Access \times European_Affiliation) + \beta_9 \cdot log(References) + \beta_{10} \cdot log(Authors) + \alpha_i + \varepsilon_i$$

Variable	Description	Obs	Mean	Std. dev.	Min	Max
I(USPTO)	Cited by USPTO patents	4425433	0.4459853	0.4970739	0	1
I(EPO)	Cited by EPO patents	4425433	0.1577769	0.3645317	0	1
I(African patents in USPTO)	Cited by USPTO patents from African applicants	4425433	0.0015551	0.0394041	0	1
I(African patents in EPO)	Cited by EPO patents from African applicants	4425433	0.0002935	0.0171302	0	1
Open Access	Open Access	4425433	0.3333807	0.4714213	0	1
I(African affiliation)	Paper with author affiliation in Africa	4425433	0.0082584	0.0904997	0	1
I(US affiliation)	Paper with author affiliation in the US	4425433	0.3559636	0.4788043	0	1
I(European affiliation)	Paper with author affiliation in Europe	4425433	0.3225750	0.4674617	0	1
Authors	Number of authors	4231022	4.8138420	7.6734200	1	2367
Countries	Number of countries	3827539	1.2059520	0.6167942	1	72
References	Number of references	4425433	30.1925200	37.3733500	0	11399

Table 4: Descriptive statistics

3.2.2. Model Specification and Regression Estimates

Table 5 shows the estimation results for 12 different specifications, three for each of the dependent variables. Each of these sub-sets include a base estimation, which only accounts for the number of references, the number of authors and the OA status. The second specification is expanded with affiliation variables, capturing the effects of researchers' geographic location. Then, the full specification includes interaction effects to test whether OA and regional affiliation have a compounded effect on the probability of being cited by patents. These different estimations make up our modelling approach, by which we try to identify the relationship that increased visibility (OA) has on a set of otherwise comparable articles (matched from the same journal issues).

In this sense, columns 1–3 show specific effects for papers being cited in USPTO patents. Because the dependent variable is binary and the model is linear, each coefficient represents the change in the probability (in percentage points) of a paper being cited, associated with a one-unit change in the corresponding independent variable, holding other covariates and journal-issue fixed effects constant. The results indicate that papers with a greater number of authors and references are more likely to be cited by USPTO patents (*Column 1*). Column 2 adds the number of countries and regional dummies. Papers originating from the United States are the most likely cited, while those from Africa are the least; papers from Europe and the rest of the world fall in between. OA generally

[11] As a short-hand, we refer to African papers (as well as USA and European papers) to those papers with at least one such regional affiliation in its author list. The same applies when we refer to patents; e.g., African patent means patent publications where the affiliation information (individual or organizations) includes an African address.

increases the likelihood of being cited by patents, even when comparing papers published in the same journal issue and controlling for the number of countries, authors and references, by approximately 4.1 percentage points ($\beta = 0.041$, $p < 0.001$). Column 3 further interacts open access with regional dummies. it shows that the positive effect of open access is particularly strong for African papers (1.5 percentage points higher than for the reference group) but weaker for US papers (1.5 percentage points lower than for the reference group), suggesting that OA enhances the translation or visibility of African science into patents.

On the specific effects for EPO patent citations, columns 4–6 reveal similar patterns. Both the number of authors and references are positively associated with the likelihood of being cited by EPO patents (Column 4–6). Column 5 adds the number of countries and regional dummies: As with USPTO citations, papers with a USA institutional affiliation remain the most likely to be cited, while papers with an African institutional affiliation remain the least. One difference is that European-affiliated papers are more likely to be cited than those from the rest of the world. The main effect of OA on EPO citation likelihood is similarly positive, increasing the probability of citation by approximately 3.8 percentage points ($\beta = 0.038$, $p < 0.001$; Column 5). Column 6 further interacts open access with regional dummies: again, the positive effect of OA is significantly stronger for African papers (3.1 percentage points higher than the reference group) and weaker for USA papers (1.8 percentage points lower than the reference group).

Further, we then turn to USPTO citations from African applicants (columns 7–9). Column 8 shows that African papers are the most likely to be cited, followed by USA papers, with European and other international papers being the least likely to receive citations. This finding reflects localized knowledge diffusion, where knowledge diffusion is stronger with geographic proximity (Jaffe et al., 1993). The positive effect of OA is smaller for USA papers, while for African, European, and other regions' papers, the effect is not significantly different. This suggests that papers with institutional affiliations in the USA already enjoy high visibility, and therefore the additional visibility conferred by OA is more limited.

Finally, results for EPO citations from African applicants (Columns 10–12) show a slightly different pattern. Column 11 shows that African papers remain the most likely to be cited, but no significant differences emerge among papers from USA, European, and the rest of the world.

	I(USPTO)			*I(EPO)*			*I(African patents in USPTO)*			*I(African patents in EPO)*		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Open access	.044***	.041***	.047***	.034***	.038***	.047***	.090***	.101***	.150***	.035	.088	.147+
	(.001)	(.001)	(.002)	(.002)	(.002)	(.003)	(.022)	(.026)	(.042)	(.052)	(.058)	(.085)
I(African affiliation)		-.072***	-.077***		-.070***	-.081***		.672***	.707***		.883***	.888***
		(.003)	(.004)		(.006)	(.008)		(.037)	(.043)		(.051)	(.061)
I(US affiliation)		.115***	.120***		.066***	.073***		.154***	.184***		.037	.071
		(.001)	(.001)		(.001)	(.002)		(.023)	(.027)		(.045)	(.057)
I(European affiliation		.004***	.002*		.028***	.029***		.001	.006		.009	.031
		(.001)	(.001)		(.001)	(.002)		(.022)	(.026)		(.042)	(.050)
Open access x			.015*			.031*			-.133			-.017
I(African affiliation)			(.007)			(.013)			(.084)			(.096)
Open access x I(US			-.015***			-.018***			-.084+			-.078
affiliation)			(.002)			(.003)			(.044)			(.085)
Open access x			.003+			-.004			-.019			-.063
I(European affiliation)			(.002)			(.003)			(.043)			(.083)
ln(Countries)		-.055***	-.055***		-.052***	-.051***		-.118***	-.116***		-.048	-.041
		(.001)	(.001)		(.002)	(.002)		(.033)	(.033)		(.067)	(.068)
ln(Authors)	.072***	.079***	.079***	.091***	.094***	.094***	.102***	.106***	.107***	.074**	.085**	.088**
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.012)	(.015)	(.015)	(.027)	(.032)	(.032)
ln(References)	.114***	.106***	.106***	.116***	.110***	.111***	.205***	.209***	.209***	.136***	.114***	.115***
	(.000)	(.001)	(.001)	(.001)	(.001)	(.001)	(.006)	(.008)	(.008)	(.015)	(.021)	(.021)
Journal issue fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-sq (within)	.052	.046	.046	.058	.043	.043	.133	.137	.138	.085	.176	.177
Obs	3743501	3398922	3398922	1342655	1217219	1217219	13060	11746	11746	7128	6308	6308
Journal issues	442984	427910	427910	249667	242304	242304	6332	6153	6153	6087	5449	5449
F	35292***	14315***	10032***	18957***	5329***	3729***	618***	192***	136***	47***	67***	50***

Standard errors shown in parentheses.. + p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table 5: Matched sample fixed effects linear models

The main effect of OA here is similarly large in magnitude (β = 0.088) but does not reach statistical significance. Column 12 further interacts open access with regional dummies: there are no significant regional differences in the effect of OA.

4. Discussion

Several policy norms drive the development of OS in the African continent (the African Union's *Agenda 2063* and the accompanying STISA policy as key examples). These frameworks aim to further develop the technical infrastructure needed for consolidating a vibrant OS environment – which is based on OA as its key socio-technical dimension (e.g., the *African Open Science Platform* – see Participants of African Open Science Platform Stakeholder Workshop et al., 2018). We part from this backdrop in order to characterise the innovative landscape observable through patent citations, and we argue that the knowledge transfer from scholarly articles into patents' citation structures is facilitated by a policy-driven increase in visibility and access.

From this coupling of observed trends and established norms, we observe that OA does play a role in the uptake of citations in patents. Yet, despite the underpinning comprehensive policy framework, the overall effect of increased visibility for being cited in patents becomes limited. This mismatch raises questions regarding the role of OA for the consolidation of an OS environment, particularly in a context where access to knowledge and collaborative practices could help boost and consolidate open innovation practices.

Furthermore, we assess the specific thematic directions of research articles cited in patents from African applicants and we see that these are indicative of the policy problems already acknowledged by national, regional, and supra-national entities. The collaborative efforts with other regions similarly reflect the motivations of African-affiliated scholars to interact with a global scientific system (cfr. Mwelva et al., 2020). However, problems with distribution of resources and the (non-)readiness of access infrastructure may be undermining factors (ASSAf, 2019). Despite African-based scientific articles gaining more attention for patent citations, non-OA sources remain the most cited references in patents. This is nearly cross-cutting in relation to which field is being cited, from health sciences to social sciences. Our inferential analysis also sheds light on this issue.

The incentives for the scientific production and communication (in terms of access licensing) seem to differ from those relating to the innovative R&D framework (Ahodode, 2023, cfr. Gebremariam, n.d.). That is, when we shift the citation outcome to citations from African patent applicants specifically, the OA advantage for African science disappears. African inventors cite African research regardless of OA. This pattern suggests two co-existing dynamics; first, African patent applicants already operate with heightened awareness of African research, meaning that locality and context-specific research is already important in this regard. Second, and more critically, this null finding points to the structural decoupling anticipated initially, in that the incentives governing patent-based innovation do not automatically align with the incentives governing academic OA policy (for example, see Wang et al., 2017). The normative and practical sides of OS policies enter here into a tension, which magnified by the dynamism and fast-paced rhythm of innovative research. Notably, patents also originate from a variety of research backgrounds (R&D teams, specialised offices, and small-, mid-, and large companies); these origins may contribute to an even more dispersed incentive structure, OA makes African research more visible to global inventors but does not, by itself, bridge the science–innovation pipeline within the continent. We do not delve deeper into the specific traits that make up these patterns, but our findings, both descriptive and inferential, are indicative of an underlying structure of practices.

On a more specific note, the matched regression analysis also sheds light on the academic authorship, and how it generally seems to play different roles as to whether it can make a difference in citations in patents. Our estimations, however, cannot distinguish between specific mechanisms as to how does OA operate (e.g., through lower access costs for inventors, or through enhanced search engine discoverability, to name a few). This distinction would be relevant for policy purposes given that, if the mechanism is primarily cost reduction, then institutional subscriptions and repository infrastructure (the AOSP model, for instance) could be framed as direct solutions. Yet, if the

mechanism is primarily discoverability, then metadata standards, indexing practices, and search optimisation become equally important – which are similarly dependent on infrastructural development.

For instance, in more specific terms, academic authorship in the USA appears to be rather inelastic to the visibility benefits of OA, whilst in African countries the effect suggests contrary directions, depending on the access type of the paper cited. The overarching policy implication is that there is a strong base already part-taking in international science networks and sharing globally. Thus, policy should be aimed at amplifying these ties whilst, as it has been proposed hitherto, promoting and consolidating an OS environment propelled by OA at its core. Taken together, these findings suggest that there are underlying elements pointing to a fracture between expected policy outcomes and effective R&D practices. Our goal of understanding how OA shapes OS practices and, with it, a dynamic OI framework leaves the questions of this decoupling open for future research.

There are clear limitations to our approach, mostly regarding the data itself. We do not distinguish between individual or collective (companies) patent applicants, nor between academic and non-academic sectors more broadly. Prior work has shown that scientific output from non-academic sectors carries distinct citation and impact patterns (McManus et al., 2021), suggesting the academia–industry composition of patent applicants is a relevant avenue for future work given our focus on the science–patent interface. We also omit problematising the cost dimensions of innovative research, neither for OA licenses nor for patenting processes. Similarly, our matched regression design allows us to offer only a limited scope of all cited scholarly literature, thus only permitting us to analyse this specific scope.

5. Conclusions and Outlook

In this paper we set out to characterise the open science dynamics in African patents. To do so, we parted from a general perspective, describing the African open science paradigm from its policy-driven framework. We found that a diverse set of instruments and high-level documents exist and share a common goal of broadening the access, infrastructure, and development to and of OS practices in the continent. Yet, most of the tools and documents recognise the inherent and systemic limitations imposed by a lack of resources and the still existing barriers to infrastructure interoperability.

Our descriptive-inferential approach highlights specific elements relating to the limitations noted above. Whilst, on the one side, African patents are focused primarily on areas of key interest for national and regional development, their knowledge base (citation patterns) reflects citation practices shaped by the formal patenting framework. Put otherwise, whilst OA is pursued from a policy perspective in relation to academia, there is little to connect this pretence with the innovative patent-based research process, where OA is of secondary interest.

From the inferential study we presented, we can directly answer our three research questions. First, OA significantly increases the likelihood of a scientific paper being cited by patents. Second, this effect is significantly stronger for African-affiliated papers in the general citation models. Third, however, when focusing specifically on citations by African patents, OA's effect no longer shows an additional premium for African-affiliated science. Together, these findings indicate that OA's benefit for African science operates primarily through global visibility rather than through strengthening the direct science-to-patent pipeline within Africa. We contribute to the knowledge on patents and the academic transfer into innovative and applied research, as well as on specific area studies focusing on innovation. Open questions on specific effects and mechanisms are left open for future research to explore.

References

Abbasiantaeb, Z., Verberne, S., & Wang, J. (2025). Tracing science-technology-linkages: A machine learning pipeline for extracting and matching patent in-text references to scientific publications. *Information Processing & Management*, 62(6), 104264. https://doi.org/10.1016/j.ipm.2025.104264.

Ahodode, B. G. C. (2023). Evolution of innovation systems, economic growth, and sustainable competitiveness in Africa. *African Journal of Science, Technology, Innovation and Development*, 15(3), 362–375. https://doi.org/10.1080/20421338.2022.2115605.

ASSAf. (2019). *African open science platform part 1*: Landscape study (p. 83). Academy of Science of South Africa (ASSAf). https://doi.org/10.17159/assaf.2019/0047.

ASSAf. (2024). Open science policy (p. 21). Academy of Science of South Africa (ASSAf); Science; Innovation Department, Republic of South Africa. https://www.assaf.org.za/wp-content/uploads/2024/02/ASSAf-Open-Science-Policy.pdf.

Bezuidenhout, L., Kelly, A. H., Leonelli, S., & Rappert, B. (2017). "$100 is not much to you": Open science and neglected accessibilities for scientific research in Africa. *Critical Public Health*, 27(1), 39–49. https://doi.org/10.1080/09581596.2016.1252832.

Bryan, K. A., & Ozcan, Y. (2021). The Impact of Open Access Mandates on Invention. *The Review of Economics and Statistics*, 103(5), 954–967. https://doi.org/10.1162/rest_a_00926.

Cotropia, C. A., Lemley, M. A., & Sampat, B. (2013). Do applicant patent citations matter? *Research Policy*, 42(4), 844–854. https://doi.org/10.1016/j.respol.2013.01.003.

Cruz Romero, R., & Stahlschmidt, S. (2025, July 10). *Structural and institutional determinants of open access publishing: A macro-perspective*. 20th International Conference on Scientometrics & Informetrics (ISSI). https://doi.org/10.51408/issi2025_115.

Dodoo, J. E., Al-Samarraie, H., & Alzahrani, A. I. (2021). Telemedicine use in sub-Saharan Africa: Barriers and policy recommendations for Covid-19 and beyond. *International Journal of Medical Informatics*, 151, 104467. https://doi.org/10.1016/j.ijmedinf.2021.104467.

Dorta-González, P., Rodríguez-Caro, A., & Dorta-González, M. I. (2025). Linking science and industry: influence of scientific research on technological innovation through patent citations. *Research Evaluation*, 34, rvaf019.

Dosso, M., Kleibrink, A., & Matusiak, M. (2022). Smart specialisation strategies in sub-Saharan Africa: Opportunities, challenges and initial mapping for Côte d'Ivoire. *African Journal of Science, Technology, Innovation and Development*, 14(1), 121–134. https://doi.org/10.1080/20421338.2020.1816265.

Fabrizio, K. R., & Di Minin, A. (2008). Commercializing the laboratory: Faculty patenting and the open science environment. Research policy, 37(5), 914-931.

Frakes, M. D., & Wasserman, M. F. (2021). Knowledge spillovers, peer effects, and telecommuting: Evidence from the U.S. Patent office. *Journal of Public Economics*, 198, 104425. https://doi.org/10.1016/j.jpubeco.2021.104425.

Gebremariam, E. B. (n.d.). *Restoring Africa's knowledge sovereignty is key to sustainable development. United nations' office of the special adviser on Africa.* Retrieved September 11, 2025, from https://www.un.org/osaa/news/epistemic-justice-and-knowledge-sovereignty-sustainable-development-africa.

Graham, S., Marco, A., & Miller, R. D. (2015). *The USPTO patent examination research dataset: A window on the process of patent examination* (Working Paper 2015-4; p. 119). USPTO. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2702637.

Hegde, D., Herkenhoff, K., & Zhu, C. (2023). Patent publication and innovation. *Journal of Political Economy,* 131(7), 1845–1903. https://doi.org/10.1086/723636.

Huang, C.-K., Neylon, C., Montgomery, L., Hosking, R., Diprose, J. P., Handcock, R. N., & Wilson, K. (2024). Open access research outputs receive more diverse citations. *Scientometrics,* 129(2), 825–845. https://doi.org/10.1007/s11192-023-04894-0.

Iyandemye, J., & Thomas, M. P. (2019). Low income countries have the highest percentages of open access publication: A systematic computational analysis of the biomedical literature. *PLOS ONE*, 14(7), e0220229. https://doi.org/10.1371/journal.pone.0220229.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics,* 108(3), 577-598. https://doi.org/10.2307/2118401.

Jia, P., Xie, W., Zhang, G., & Wang, X. (2026). Open access and the flow of knowledge into technology: Evidence from journal transitions to full and immediate OA. *Journal of the Association for Information Science and Technology*, 77(2), 429–450. https://doi.org/10.1002/asi.70045.

Maddi, A. (2024, November 22). *Open Science as a Catalyst for Invention: Evidence from Patent Data.* Nordic Workshop on Bibliometrics and Research Policy 2024, Reykjavik, Iceland. https://doi.org/10.5281/zenodo.14205167.

Marx, M. (2024). *Reliance on science* (Version 64). Zenodo. https://doi.org/10.5281/zenodo.11461587.

Marx, M., & Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal,* 41(9), 1572–1594. https://doi.org/10.1002/smj.3145.

Marx, M., & Scharfmann, E. (2024). *Does patenting promote the progress of science?* https://assets.zyrosite.com/YZ97xEPgM7SegrpN/ppp20240531-m5KvaGBojMIza35N.pdf.

McManus, C., Baeta Neves, A. A., & Prata, A. T. (2021). Scientific publications from non-academic sectors and their impact. *Scientometrics*, 126(11), 8887-8911.

Mwelwa, J., Boulton, G., Wafula, J. M., & Loucoubar, C. (2020). Developing open science in africa: Barriers, solutions and opportunities. *Data Science Journal*, 19(1). https://doi.org/10.5334/dsj-2020-031.

Nishioka, C., & Färber, M. (2020, August). Trends of publications' citations and altmetrics based on open access types. In Proceedings of the *ACM/IEEE Joint Conference on Digital Libraries* in 2020 (pp. 503-504).

Omini, E. E. (2019). Institutions, innovation and economic growth patterns in East Asia: Learning curves for sub-Saharan Africa. *Open Journal of Economics and Commerce,* 2(1), 16–24. https://doi.org/10.22259/2638-549X.0201004.

Participants of African Open Science Platform Stakeholder Workshop, S. 2018., Participants of African Open Science Platform Strategy Workshop, M. 2018., Advisory Council, A. O. S. P. P., Technical Advisory Board, A. O. S. P., Boulton, G., Hodson, S., Serageldin, I., Qhobela, M., Mokhele, K., Dakora, F., Veldsman, S., & Wafula, J. (2018). *The future of science and science of the future: Vision and strategy for the african open science platform (v02)*. Zenodo. https://doi.org/10.5281/zenodo.2222418.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of The Royal Society Interface*, 4(15), 707–719. https://doi.org/10.1098/rsif.2007.0213.

UNESCO, & Canadian National Commission for UNESCO. (2022). *An introduction to the UNESCO recommendation on open science.* UNESCO. https://doi.org/10.54677/XOIR1696.

Union, A. (2024). *Science, technology and innovation strategy for Africa* (STISA) 2034 (p. 30). African Union. https://au.int/sites/default/files/documents/45087-doc-AU_STISA_2025-2034_Strategy_ENGLISH.pdf.

Vallejo, B., Oyelaran-Oyeyinka, B., Ozor, N., & Bolo, M. (2019). Open innovation and innovation intermediaries in sub-Saharan Africa. *Sustainability*, 11(2), 392. https://doi.org/10.3390/su11020392.

Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PLOS ONE*, 10(5), e0127298. https://doi.org/10.1371/journal.pone.0127298.

Wang, J., & Verberne, S. (2024). Comparing patent in-text and front-page references to science. Journal of Informetrics, 18(4), 101564. https://doi.org/10.1016/j.joi.2024.101564.

Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8), 1416–1436. https://doi.org/10.1016/j.respol.2017.06.006.

Wang, Z., Chen, Y., & Yang, C. (2025). The role of preprints in open science: Accelerating knowledge transfer from science to technology. *Journal of Informetrics*, 19(2), 101663. https://doi.org/10.1016/j.joi.2025.101663.