



**HAL**  
open science

# Machine Learning in Sub-Saharan Africa: A Critical Review of Selected Research Publications, 2010–2021

Judy Van Biljon

► **To cite this version:**

Judy Van Biljon. Machine Learning in Sub-Saharan Africa: A Critical Review of Selected Research Publications, 2010–2021. 17th International Conference on Social Implications of Computers in Developing Countries (ICT4D), May 2022, Lima, Peru. pp.363-376, 10.1007/978-3-031-19429-0\_22 . hal-04601175

**HAL Id: hal-04601175**

**<https://inria.hal.science/hal-04601175v1>**

Submitted on 5 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

# MACHINE LEARNING IN SUB-SAHARAN AFRICA: A CRITICAL REVIEW OF SELECTED RESEARCH PUBLICATIONS, 2010-2021

Judy van Biljon, University of South Africa, vbilja@unisa.ac.za

**Abstract:** Machine learning, a field of artificial intelligence application, will profoundly impact countries in sub-Saharan Africa (SSA) in the next few decades. The extent of the impact will depend largely on the countries' readiness. Roadmaps and research agendas towards implementing machine learning for development (ML4D) exist, but little is known about the state of ML4D research in sub-Saharan Africa. Guided by an existing ML4D roadmap, we conducted a critical literature review on selected research publications papers published from 2010 to 2021 with the aim of describing the status of SSA ML publications and informing future research. This should be of interest to researchers and funding organizations investigating the potential impact of ML4D projects in the SSA region.

**Keywords:** ML4D, AI4D, machine learning, development, artificial intelligence

## 1. INTRODUCTION

The relationship between information and communication technologies (ICTs) and international development is dynamic and constantly evolving (Heeks, 2020a). The geopolitical discourses on technology adoption have evolved from techno-centric approaches aimed at replacing human capacity, to more holistic design perspectives aimed at augmenting human capacity and enabling meaningful human practices in which technology plays a part. Digitalisation has paved the way for datafication, the masses of digital traces left by people and technologies in online spaces and the proliferation of advanced tools for the integration, analysis, and visualization of data patterns for purposes of decision making and commercialization (Flyverbom et al., 2019). Nonetheless, society still needs to come to terms with artificial intelligence (AI) applications where those applications are used to augment and, increasingly replace human decision making in data intensive, socially sensitive decision processes like loan-approvals, hiring and granting parole to people (Mehrabi et al., 2021). This is especially challenging for sub-Saharan African (SSA) countries, some of which are still coming to terms with issues surrounding the first three industrial revolutions, e.g. problems relating to universal access to electricity, mechanization of production, and industry automation (Butcher et al., 2021).

Against the background of SSA countries having severe challenges in preparing for the optimal and sustainable use of AI, it is imperative to align the efforts of the different stakeholders by having a shared conceptualisation of what is feasible and what is advisable. The focus of this paper is on Machine learning (ML), a subset of AI that refers to the ability of computers to process information relating to a specific task and use the results to adjust their behaviour to maximise their chances of achieving their goals without being explicitly programmed to do so (Mann & Hubert, 2020). There are clear benefits to algorithmic decision-making. Machines can take into account orders of magnitude more factors than humans can and machines do not become tired or bored (Mehrabi et al., 2021). However, like humans, algorithms are vulnerable to biases (Ferrer et

al., 2021). This can lead to digital discrimination, which means users are treated unfairly, unethically or just differently based on their personal data that is automatically processed by an algorithm (Criado & Such, 2019).

Africa has been a playground for Western researchers (Bai, 2018). While their involvement has produced significant positive outcomes in terms of research impact (van Biljon & Renaud, 2019) it is important to guard against Africa becoming a dumping ground for perceived challenges, imposed problems and barriers which limit the use of AI in the attainment of sustainable development objectives. In trying to make sense of the AI value proposition and the consequences of adoption, researchers can become polarized along disciplinary lines. Researchers in the physical sciences may focus on AI-related solutions and progress which creates a positive AI narrative, while researchers in the humanities focus on the biases and discrimination which create a negative AI narrative. As noted, finding solutions to bias and discrimination in AI requires robust cross-disciplinary collaboration (Ferrer et al., 2021). Therefore, the different stakeholders need to interact to ensure a contextually sensitive, holistic approach; if there is no shared awareness and interdisciplinary knowledge exchange is not managed, then it can lead to siloed thinking, ill-advised actions and policies.

The advancements of AI are set to continue and impact all countries, regardless of their geography, actual or perceived readiness to benefit from the use of AI and other emerging (Butcher et al., 2021). ML applications have proven valuable in addressing some problems associated with less developed contexts and there is the undeniably potential to extend human capabilities. However, to harness the benefits of technology and to blunt its worst disruptions, developing countries will need to take rapid action to ensure they can compete in the economy of the future, specifically by investing in health and education as the building blocks of human capital (World Bank Report, 2019). Accordingly, the potential of ML technology when investigating SSA problems merits attention but there are numerous challenges to consider (Mann & Hubert, 2020). Besides the problems with data availability, computing capacity and relevant skills (Weber & Toyama, 2010), there is the risk of introducing bias, in the data and/or the algorithm. To manage the complexities of adopting ML4D solutions research it is necessary to formulate theorizations that can guide the consideration of goals, prerequisites, stakeholders etc. To address this problem, research agendas (Butcher et al., 2021, Smith & Neupane, 2018) and a ML4D roadmap (De-Arteaga, Herlands, Neill, and Dubrawski, 2018) have been developed to guide the adoption of AI technologies. Given the dynamic, ever-evolving nature of ICT4D a critical overview of SSA literature published on ML4D could be useful to improve our understanding of what has been published and to inform future research in the domain. The rest of the paper is organised as follows: Section 2 will provide an overview of the relevant literature, Section 3 will present the research design, results and findings with the conclusion being discussed in Section 4.

## **2. OVERVIEW OF RELATED LITERATURE**

### **2.1. ICTs and International Development**

Information and communication technology for development (ICT4D) refers to the academic field concerned with the use of ICTs for international development with the focus on the so-called developing countries and particular emphasis on the less materially advantaged members of those societies (Walsham, 2017). Sein et al. (2019) argue that three concepts, namely ICT, D and the catchall term '4' or 'for' lie at the root of the definition of the interdisciplinary ICT4D field. Related to each concept is a specific group of theories, the first group focuses on Development theories i.e., What is development (D)? The second group focuses on theories conceptualizing ICT, i.e., What is ICT in the context of development? The third and final group of theories focus on transformative processes linking ICT to D i.e., How does ICT make D happen?

An in-depth discussion of these groups of theories and the three ICT4D components are beyond the scope of this paper where the focus is on the ML research that has been conducted in the ICT4D domain. Digital information and communication technologies (ICTs) impact international development on the micro and the macro levels. At the macro level, many studies have employed quantitative data analyses to investigate the impact of digital information and communication technologies (ICTs) on international development. For example, Bankole and Mimbi (2017) used previous research in the domain to propose a research direction for macro/micro level impact of ICT on national development on the African continent while Gwagwa et al. (2020) considered policy responses to AI. At the micro (individual) level, impact can be viewed through one or more of the following lenses: (1) economic development as accumulation of financial capital, (2) livelihood development as accumulation of livelihood assets; including not just money but also health, skills, information, etc, and (3) capabilities development as greater freedom on what to be and to do (Heeks, 2014). From the livelihoods and the capability view of the micro level, the core development concept of context comes into play. Sein et al. (2019) argue that any holistic understanding of the use of ICT for development needs to include a contextual understanding of the theoretical premise of ICT, development and the transformative process by which ICT may lead to development. Taking a cultural view of context, Davison and Martinsons (2016) argue that in an increasingly globalised world, the context of locally distinctive values and behaviours will become more important.

As noted, the relationship between ICTs and development is constantly evolving (Heeks, 2020a). Considering poverty alleviation as the main driver in ICT4D 1.0 marginalized the poor in developing countries allowing a supply-driven focus characterising users largely as passive consumers; ICT4D 2.0 focuses on centralizing them thereby creating a demand-driven focus which characterized them as active producers and innovators (Heeks, 2020b). Now ICT4D 3.0, also described as *digital-for-development*, is emerging as a new paradigm between digital technologies and international development. Several factors support the 'digital-for-development' drive. The shift in focus, from the proliferation of communication and information to the extraction of knowledge from the resulting data (Mann & Hubert, 2020). The interest in and the availability of open-source data (Kitchin, 2014), the advent of deep learning, and the proliferation of machine learning tools (Kumar & Sharma, 2021). These developments created the expectation that AI as a tool could be used to solve many of the developing world's problems like the unrealistic expectations once associated with ICT4D.

Considering the issues and contentions in responsible business and digital transformation, Flyverbom et al. (2019) consider the most common ethical issues at the upstream level to be the violation and intrusion of privacy, the consent of the data provider, and the transparency with which big data companies collect the information. Clearly, those issues could apply to developing contexts too. Given the potential risk of bias in the datasets selected and in the algorithms applied (Ferrer et al., 2021), we are at risk of perpetuating historical and contemporary socioeconomic disparities if the research problem, design and interpretation of the findings are not contextually sensitive. The next section will briefly review AI, ML and International Development.

## 2.2. Artificial Intelligence, Machine learning and International Development

Weber & Toyama (2010), when considering on-the-ground projects in development, highlight the following challenges to using artificial intelligence techniques to impact a poor community (a) finding large-scale data in an appropriate format, (b) the low cost of labour and (c) the fact that a lack of intelligence is often not the bottleneck. Mann and Hilbert (2020) identified the tension between global efficiency and local needs as one of the issues at the crossroads of AI/ML and development. For example, automating labour conflicts with local need for jobs. Resolving, or at least, balancing these tensions is pivotal in the quest to harness the potential of AI/ML techniques and practices for designing at the margins. For example, the ability to perform automatic pattern recognition may prove useful when experts are either costly or unavailable in a specific country or

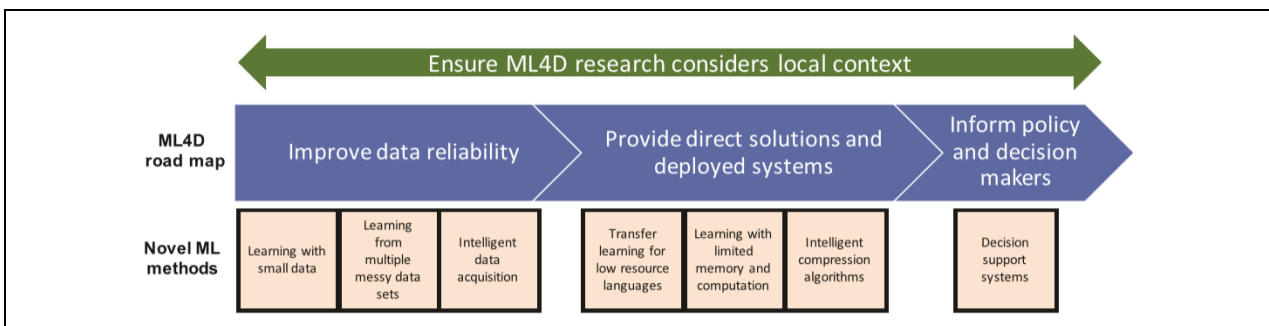
region (Weber & Toyama, 2010). While acknowledging the real challenges mentioned regarding AI4D (with ML4D as sub-field) we argue that there is value in researching the use of AI for development by researchers in developing contexts. Resource-constrained contexts have implications for socio-economic development efforts in the sense that resources are scarce and should not be wasted in pursuit of non-practicable and unsustainable solutions.

To guide future investments in capacity building for responsible AI development and deployment, (Butcher et al., 2021) investigated the AI landscape in SSA, specifically what capacity already exists and what measures stakeholders in the region are taking to ensure that they are AI-ready. They identified the stakeholder groups as Centres of Higher Education and Training, Members of the AI Community and Government and made recommendations on AI-related Academic Activities, Research and Development, Policy Environment, Challenges and Capacity Building Needs and Diversity in AI-related Activities. Researchers applying AI in Developing Contexts have to consider context-specific challenges in terms of data, usability and resources (Weber & Toyama, 2010). Furthermore, Smith & Neupane (2018) suggest that the digital and analogue foundations required for an ethical and equitable application of the technology in many countries in the Global South are largely absent while salient power asymmetries remain. The constraints and challenges mentioned may lead us to question the feasibility of applying AI in Developing Contexts. On the other hand, the proliferation of software tools allows even naive users to develop ML-based models for their respective problem domains (Kumar & Sharma, 2021). Therefore, the more useful question to engage on is: What research has been done on the topic of applying ML to Developing Countries, specifically SSA?

Accordingly, De-Arteaga et al. (2018) proposed the following key properties for ML allocations in the developing world (ML4D).

- Applications and data are geographically constrained to developing countries.
- Problems concern a critical development area for the region of interest.
- Substantially uses ML as an integral element
- Problems being addressed or contextual elements necessitate solutions that differ from existing or plausible solutions in developed regions, and the proposed solution effectively addresses these differences.

The ML4D research roadmap proposed by De-Arteaga et al. (2018) as depicted in Figure 1, suggests three technical stages where ML4D can play an essential role and meaningfully contribute to global development, namely to (1) improve data reliability, (2) provide direct solutions and deployed systems, and (3) to inform policy and decision makers. Each of the three stages is then linked to novel ML methods to be applied in the ML4D context.



**Figure 1: ML4D Roadmap (De-Arteaga, Herlands, Neill, and Dubrawski, 2018)**

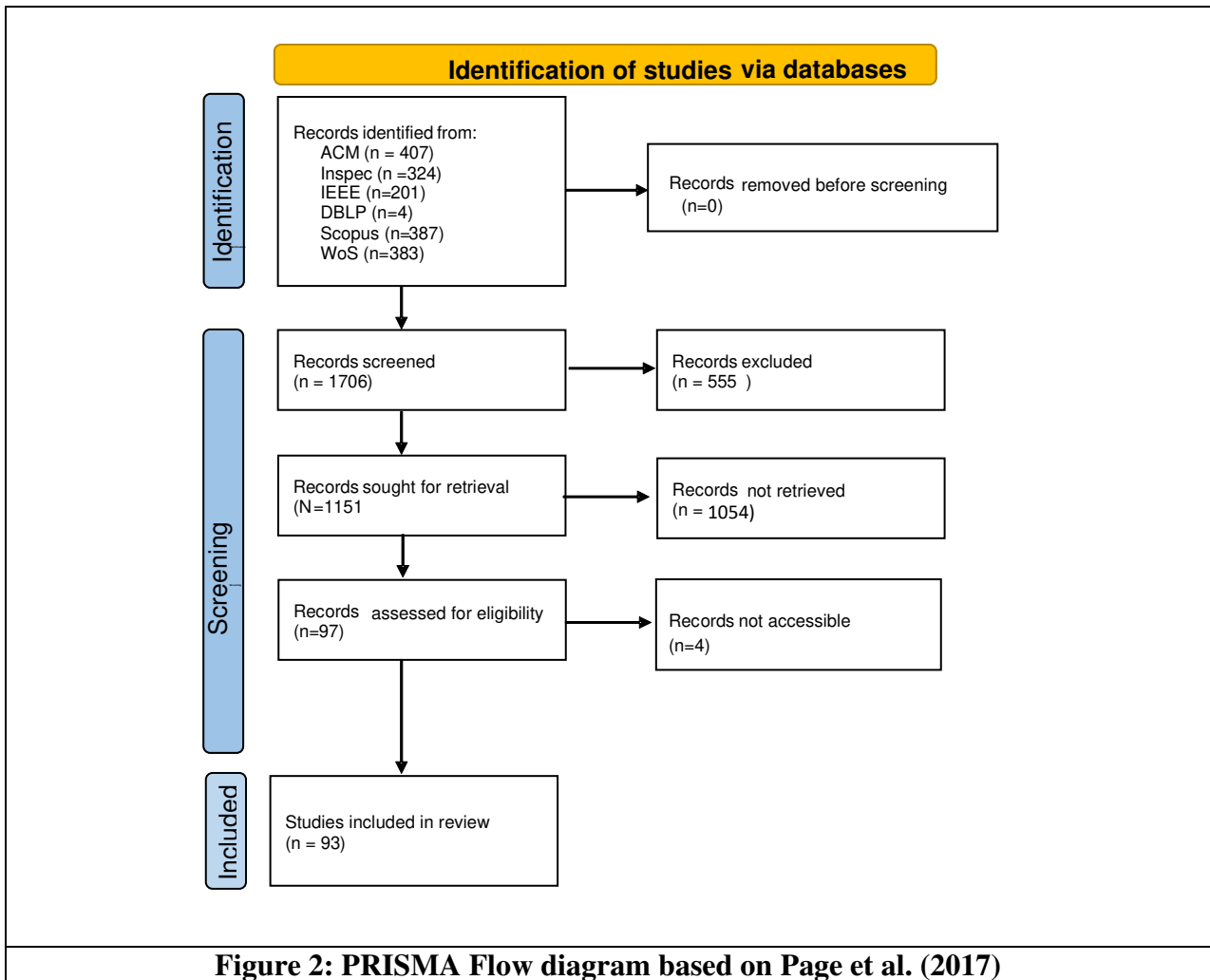
This roadmap was used as a point of departure in analyzing the findings of the literature review as described in Section 3.3. The next section will discuss the research design.

### **3. RESEARCH DESIGN**

The research design involved a critical review (as a type of systematic literature review) on ML4D in SSA based on a selected set of papers from 2010 to 2021. The aim of systematic literature reviews is to go beyond aggregating existing evidence toward constructing lessons from the accumulated literature (Kitchenham and Brereton., 2009) and support the analysis of the literature in a research domain. Grant and Booth (2009) proposed a topology of literature reviews whereby, the most basic type of review (literature review) is useful for building on previous work, avoiding duplication and identifying gaps in the literature. Chatterjee & Davison (2021:228) argue that the gap-spotting approach is built on a fallacious argument since it assumes the fact that a research study has not been done, implies that there is value (a contribution) in doing it. Therefore, we opted for the critical review which has the descriptive purpose of a literature review but goes beyond that, i.e. it aims to demonstrate that the literature has been researched extensively and reviewed critically to include a degree of analysis and conceptual innovation which typically results in a hypothesis or model (Grant & Booth, 2009; Bandara et al., 2015). Bramer et al. (2018) propose a systematic search strategy to develop literature searches starting with clear and focused questions. Considering the purpose mentioned, this research question is formulated as: What research has been done on the topic of applying ML to Sub-Saharan Africa? The review procedure is explained in Section 3.1 and the results in Section 3.2 and Section 3.3 respectively. Ethical clearance for this study was obtained from the College Research and Ethics Committee of the University of South Africa.

#### **3.1. Literature Analysis**

The literature review was carried out to construct a corpus of papers representative of ML4D research conducted in SSA for the period 2010 to 2021. Grant and Booth (2009) characterize the review types in terms of the purpose of the search, appraisal, synthesis and analysis but do not prescribe the relevant procedures. Therefore, we followed the PRISMA (Page et al., 2017) guidelines to be transparent in the reporting of the methods and facilitate replication as depicted in Figure 2. The review was conducted in the week of 6 August 2021, using the search string: "Machine Learning" AND ("Developing Countries" OR "Developing World") for the period 2010 to 2021. Due to the wide and interdisciplinary scope of interest in ML, we included the following databases ACM, Inspec, IEEE, DBLP, Scopus and WoS. The search criteria were set to include only conference and journal papers and 1706 records (publications) were identified as meeting said criteria. These titles were screened to remove items such as patents, magazine articles, abstracts from books, workshops and bulletins published in conference proceedings because the scope and format of these items were not comparable to the conference and journal publications. Removing these 555 items, 1151 records remained for retrieval. Considering these 1151 records, the location where the research was done was captured using the title, abstract or the full paper if necessary. For 337 publications the region was not specified and for 717 publications the research was from other parts of the world (not conducted in SSA). Having removed those, 98 publications remained for analysis on ML research done in SSA. Considering the 98 records of studies conducted in SSA, 4 were inaccessible so the final corpus consisted of 93 papers.

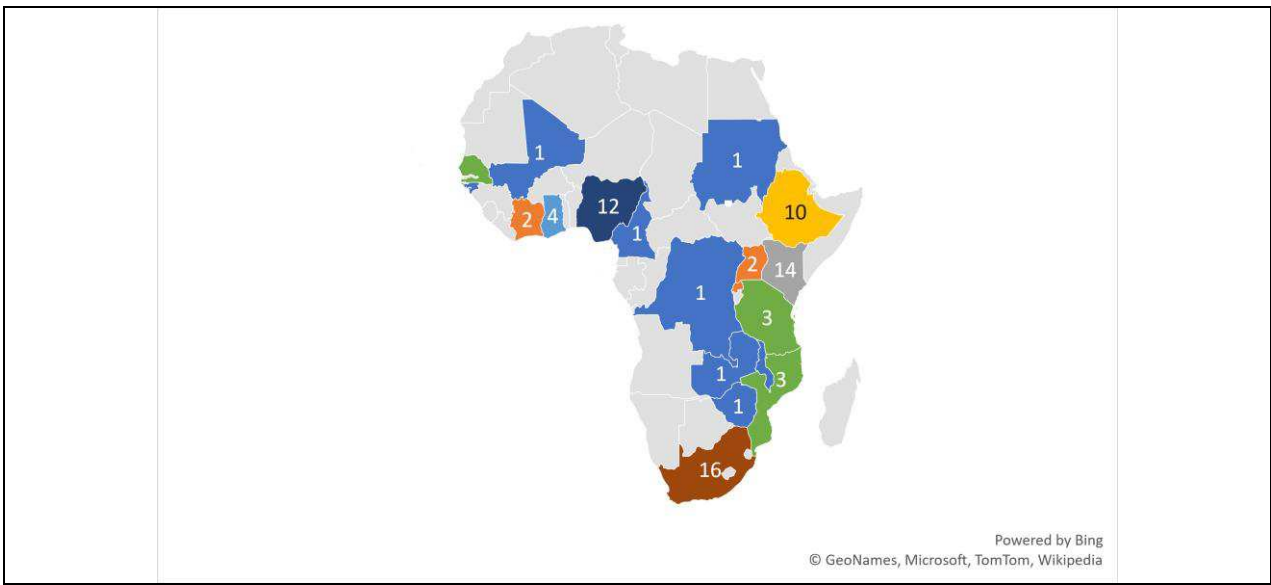


Bandara et al. (2015) suggest the inductive approach for coding when the themes to be reported on are purely derived from the literature analysis and the deductive approach when the themes to be reported on are predetermined to some extent by using coding schemes or theorizations like theoretical lenses, models, frameworks or alignment with specific research questions. Bramer et al. (2018) concurs on the usefulness of predetermined themes and suggests finding articles appropriate for guiding the research and analysis strategy. The most appropriate paper in this regard was the ML4D roadmap (De-Arteaga, Herlands, Neill, and Dubrawski, 2018) as depicted in Figure 1 and hence those themes were applied in guiding the analysis (see Section 3.3).

### 3.2. Publication’s overview

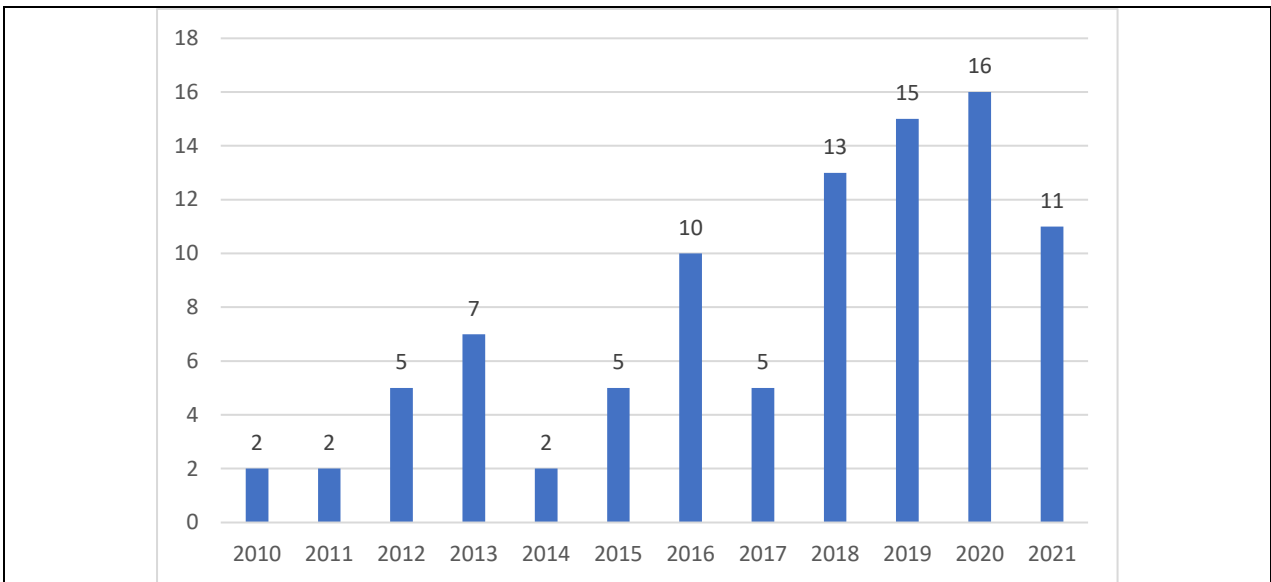
The 93 publications were first analysed to provide an overview in terms of the distribution across years, domains, and countries. Figure 3 depicts the number of publications per country. It can be observed that most studies were done in South Africa (16). Kenya (14) has the same number as the group of cross-border studies (not indicated on the map), followed by Nigeria (12) and Ethiopia (10). This is followed by Ghana (4), Mozambique, Senegal and Tanzania each three with a number of countries contributing two or less studies. It is unknown how many publications appear in other languages but having the English language as an inclusion criterion is possibly a limitation in providing a holistic view of SSA research publications.





**Figure 3: Number of publications per-country**

Figure 4 depicts the papers per year, showing significant growth in the number of papers since 2010, especially since 2017. Note that the year 2021 was captured only until 6 August 2021, so the number of publications in 2021 is likely to exceed all previous years. A simplistic extrapolation would put the final number for 2021 at 19. The publications selected consist of 58 conference publications and 35 journal publications. This resonates with the fact that the field is dynamic, and researchers might need the quicker turnaround times associated with publishing in conference proceedings.

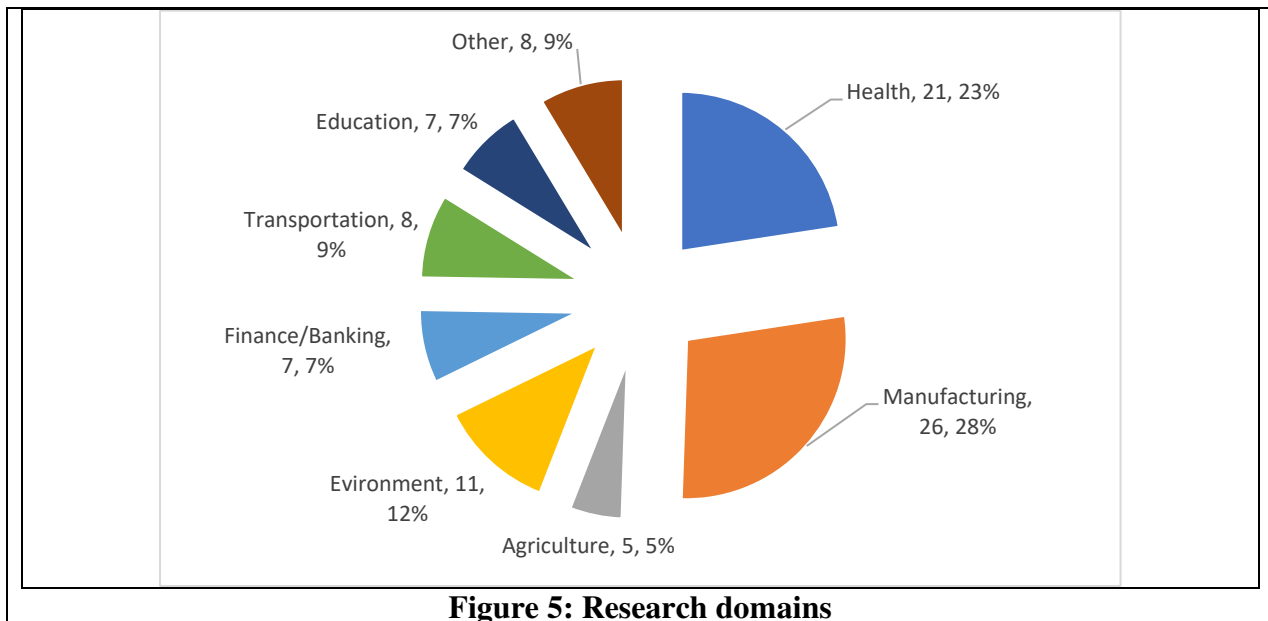


**Figure 4: Paper distribution across years**

The distribution of the publications across domains is presented in Figure 5. The highest concentration of publications was found in Manufacturing (28%) which includes Engineering and several publications where ML techniques were used to optimise systems. This is followed by Health (21%) and the Environment (12%) of the publications with the rest all below 10%. The relatively few papers on Education is noteworthy, considering that Health and Education were highlighted as priority fields in the World Bank report on the changing nature of work (World Bank Report, 2019). The challenges with data availability and quality received from the Education

systems as well as the political and cultural challenges with the classification of students according to their performance (Archer & Prinsloo, 2020) would be some of the factors limiting research in the Education domain. Research on the social aspects would fall into the *Other* category. The papers in that category includes the use of ML techniques to address social challenges like illuminating the working of an informal industry on one of the poorest countries (Björkegren, 2020). The paper on narratives and counternarratives on data sharing in Africa (Abebe et al., 2021) engages with issues arising from power imbalances and identify recurring barriers to data sharing as well as inequities in the distribution of data sharing benefits.

Based on a literature survey (Mehrabi et al., 2021) propose a taxonomy on bias and fairness in Machine Learning which includes 23 types of bias and six types of discrimination. Their paper provides evidence of a dynamic body of research into the social impacts of ML, but apart from a few examples like the paper by Abebe et al. (2021) the latter is not evident from the ML4D SSA literature surveyed in this study. This could be indicative of the complexities involved in classifying interdisciplinary research into distinct categories. However, more research is needed to investigate whether the research on ethical and responsible AI is being published in electronic databases not included in this study or whether this points to an actual gap in the ML4D literature. Assuming that research is published elsewhere this still raises questions about the apparent lack of interdisciplinary research on the ML4D topic.



### 3.3. ML4D roadmap

Another approach is to consider the elements of the ML4D roadmap De-Arteaga et al. (2018) as a point of departure in making sense of the results of the critical review. Notably, the roadmap was suggested for ML4D and not for ML4D research, so it is necessary to consider how, if at all, those elements can inform research. Therefore each of the constructs depicted in Figure 1 (De-Arteaga et al., 2018) are now considered in terms of their relevance for ML4D research.

Considering the *ML4D road map* and *novel ML Methods* components in Figure 1, these constructs could be useful in informing research goals. Therefore, these were used as initial codes in analysing the data with Atlas.ti version 9. For example, the publications related to the code, *limited memory or computational capacity* as depicted in Figure 6. There is evidence of publications related to each of these initial codes as depicted in Figure 7 (the number in brackets indicates the number of sub-categories), additional categories such as *data accessibility* also emerged from the data. Besides confirming the relevance of the suggested categories, the importance of data-related issues when considering ML4D in the SSA context became evident.

The data issues, as depicted in Figure 8 include but are not limited to, data availability, absent data source, unbalanced data, collection challenges, quality issues and accessibility issues. Based on the number of data-related issues the construct *data availability* should be added since that cannot be assumed in SSA. Furthermore, it is a prerequisite for *data reliability* and the rest of the constructs mentioned in Figure 1.

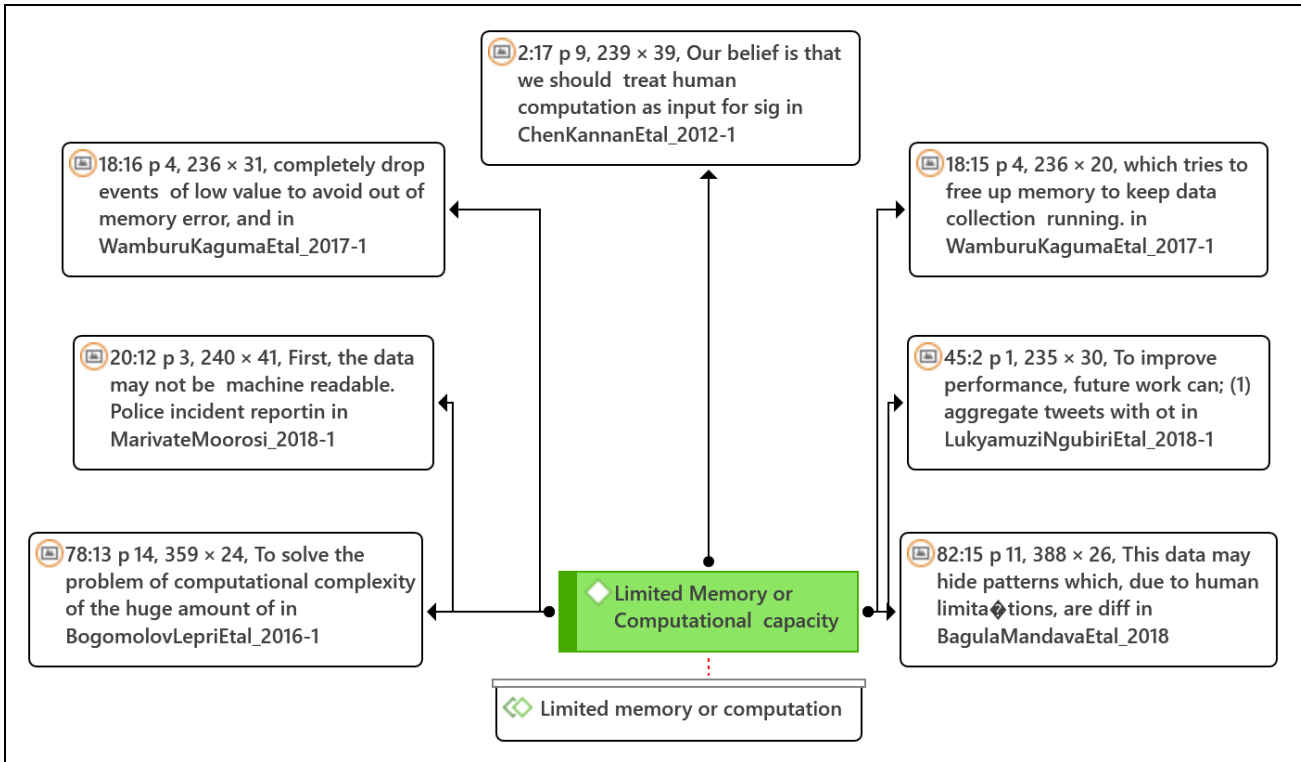


Figure 6: Papers related to the category limited Memory or Computational Capacity

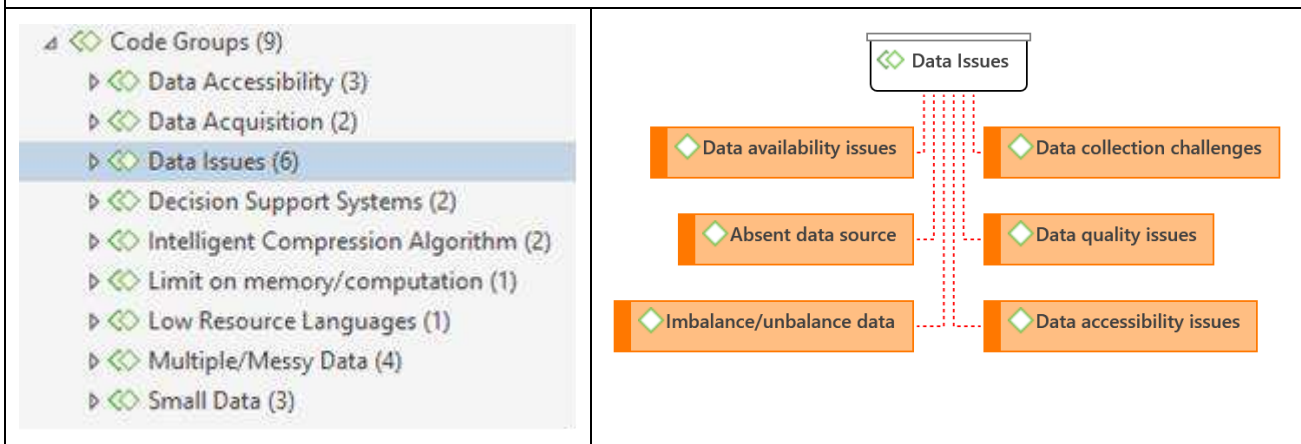


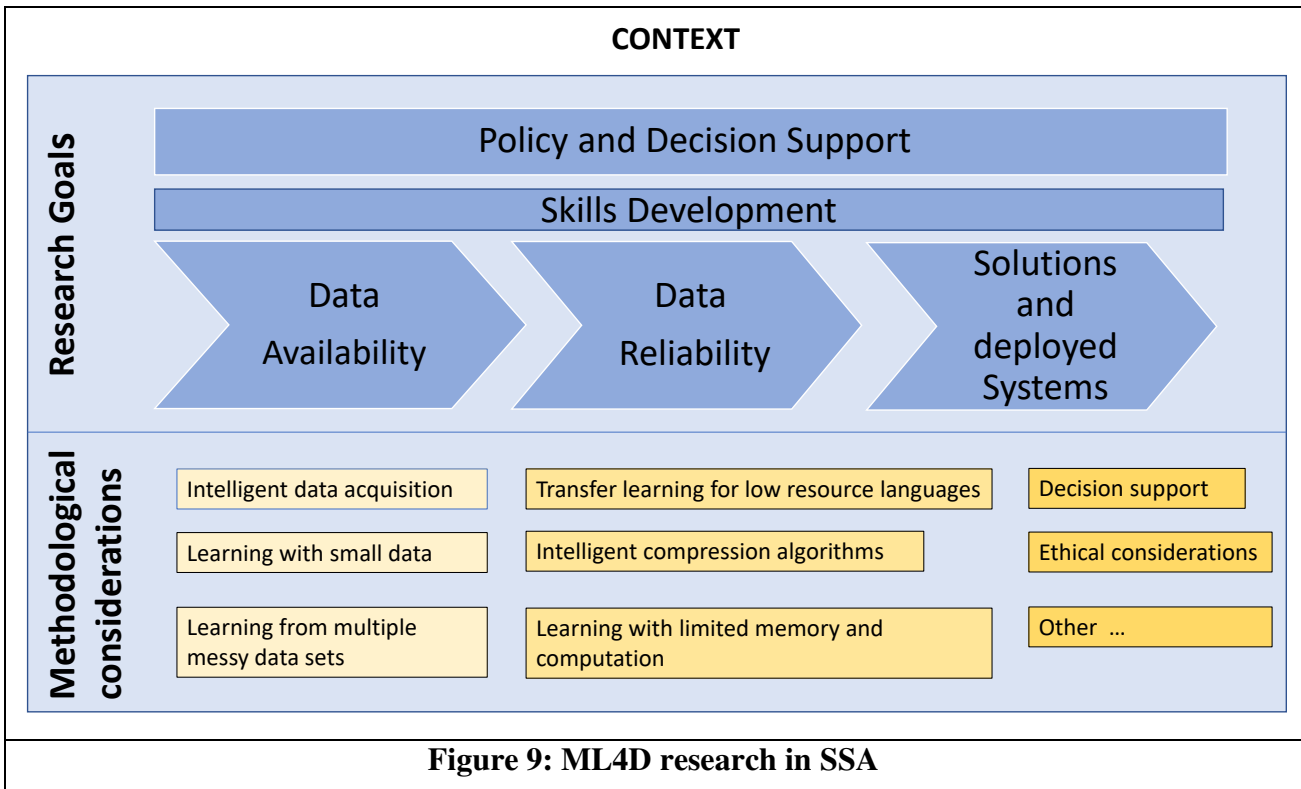
Figure 7: Goals as code groups

Figure 8: Data issue categories

### 3.4. Summary

The insights based on the roadmap presented in Figure 1 and the findings from the literature review are presented in Figure 9. As with all research in developing countries, the context is imperative (Davison and Martinsons, 2016) and this is also evident in the theoretical lenses applied by Sein et al. (2019). Therefore, the context element is presented not as a single construct but rather a pervasive, overarching and encompassing phenomenon. Concerning *Research Goals*, *Data Availability* has been added to the goals of *Data Reliability* and *Solutions and deployed Systems*. The papers relating to Education were focused on predicting academic performance (Mgala & Mbogho, 2015) and a diversity of other projects e.g. a cloud-based infrastructure for

digital content delivery (Mlitwa & Simbarashe, 2019) not specifically on improving ML skills. However, Butcher et al. (2021) highlighted the importance of skills development in the Sub-Saharan regions and therefore *Skills Development* has been added to Figure 9. *Decision Support* is a fundamental goal of ML research and has been retained. *Policy* encompasses all the ethical policies and requirements to mitigate bias and discrimination in the data capturing and the ML procedures applied. The *Methodological Considerations* retained the ML methods from Figure 1 with the addition of Ethical considerations and Other to provide for new methods being added. Not to imply that the set of goals proposed are immutable, but methods are more reliant on technology and hence more susceptible to change.



### 3.5. Limitations

The study has a number of limitations. Keyword searches have several limitations in particular where keywords are related to specific technologies. (Levy & Ellis, 2006). However, given the wide scope of the Artificial Intelligence field using the term ML4D was considered necessary in focusing our efforts. Despite a rigorous approach enforced by the application of the PRISMA guideline, any literature survey can be faulted for having made contestable decisions (both intentionally and unintentionally) about which works to include and which elements to emphasise in the works identified. We also acknowledge the ambiguity and limitations surrounding the term ‘development’ (Sen, 2001; Toyama, 2010) but limit the discussion to aligning our perspective with the holistic perspective provided by Sein et al. (2019). For the purpose of this paper, we consider human development as going beyond socio-economic development toward increased freedom of choice (Sen, 2001). ML is merely a tool whereof the application is determined by the researchers, consequently their theoretical lenses and goals determine what ML methods are relevant. Likewise, De-Arteaga et al. (2018) used specific lenses and granularities in suggesting the methodological considerations in their roadmap thereby making it difficult to evaluate and extend that. Adding a taxonomy of ML methods was considered but linking a taxonomy of methods to a context may be limiting. How to represent the relevant methods as part of informing future ML4D research could be considered towards future refinements of ML4D research roadmaps.

## 4. CONCLUSION

Artificial intelligence, with machine learning as sub-field is like a tidal wave spreading out and affecting all sections of society. Due to data availability, data quality, computing capacity and a lack of technical skills many Sub-Saharan African countries are ill-prepared for managing the possible disruptions or optimizing the opportunities ML brings. This has compelled researchers in different disciplines to research the impact of ML for developing contexts. Our contention is that these efforts can be jeopardised if researchers become polarized into those who seem pro-ML4D since their focus is to develop and promote new ML solutions and those who seem anti-ML4D since their focus on the societal challenges of classifying people and the harm that ML applications can cause through biases in the research procedures, data or the algorithms. Based on existing theorizations (an ML4D roadmap) and a critical literature review of SSA research we share insights towards guiding ML4D research in SSA. These findings highlight the many data-related challenges and the need to consider *data access* together with *data reliability*. Furthermore, the lack of research into the societal affects of ML is a concern. Clearly, this is affected by the selection of the research databases but even so multi-, inter- and transdisciplinary research involving ML might assist in highlighting under-represented yet important societal considerations required to optimise the opportunities while responsibly navigating the context-specific challenges and resource constraints. As noted a contextual understanding of what is feasible and desirable in the SSA context is essential as resources are scarce and should not be wasted in pursuit of impracticable and unsustainable projects.

## 5. ACKNOWLEDGEMENT

This paper is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa (Grant No. 98564).

## 6. REFERENCES AND CITATIONS

- Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S. L., & Sadagopan, S. (2021). Narratives and counternarratives on data sharing in Africa. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 329–341. <https://doi.org/10.1145/3442188.3445897>
- Archer, E., & Prinsloo, P. (2020). Speaking the unspoken in learning analytics: troubling the defaults. *Assessment and Evaluation in Higher Education*, 45(6), 888–900. <https://doi.org/10.1080/02602938.2019.1694863>
- Bai, Y. (2018). Has the Global South become a playground for Western scholars in information and communication technologies for development? Evidence from a three-journal analysis. *Scientometrics*, 116(3), 2139–2153.
- Bandara, W., Furtmueller, E., Gorbacheva, E., Miskon, S., & Beekhuyzen, J. (2015). Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support. *Communications of the Association for Information Systems*, 37, 154–204. <https://doi.org/10.17705/1cais.03708>
- Bankole, F. and Mimbi, L. (2017). ICT Infrastructure and It's Impact on National Development: A Research Direction for Africa. *The African Journal of Information Systems*. 9, 2, 1.
- Björkegren, D. (2020). Measuring informal work with digital traces: Mobile payphone operators in Rwanda. *PervasiveHealth: Pervasive Computing Technologies for Healthcare*. <https://doi.org/10.1145/3392561.3397576>
- Bramer, W. M., de Jonge, G. B., Rethlefsen, M. L., Mast, F., & Kleijnen, J. (2018). A systematic approach to searching: An efficient and complete method to develop literature searches. *Journal of the Medical Library Association*, 106(4), 531–541.

<https://doi.org/10.5195/jmla.2018.283>

- Butcher, N., Wilson-strydom, M., & Baijnath, M. (2021). *Artificial intelligence capacity in sub-Saharan Africa: Compendium report*.
- Chatterjee, S., & Davison, R. M. (2021). The need for compelling problematisation in research: The prevalence of the gap-spotting approach and its limitations. *Information Systems Journal*, 31(2), 227–230. <https://doi.org/10.1111/isj.12316>
- Criado, N., & Such, J. (2019). Digital discrimination. In *Digital Discrimination*. <https://doi.org/10.1093/oso/9780198838494.003.0004>
- Davison, R.M. and Martinsons, M. G. (2016). Context is King! *Journal of Information Technology*, 31(3), 241-249.
- De-Arteaga, M., Herlands, W., Neill, D. B., & Dubrawski, A. (2018). Machine learning for the developing world. *ACM Transactions on Management Information Systems*, 9(2). <https://doi.org/10.1145/3210548>
- Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I. and De Beer, J. 2020. Artificial Intelligence (AI) Deployments in Africa: Benefits, Challenges and Policy Dimensions. *The African Journal of Information and Communication*. 26, 1–28.
- Ferrer, X., Nuenen, T. Van, Such, J. M., Cote, M., & Criado, N. (2021). Bias and Discrimination in AI: A Cross-Disciplinary Perspective. *IEEE Technology and Society Magazine*, 40(2), 72–80. <https://doi.org/10.1109/MTS.2021.3056293>
- Flyverbom, M., Deibert, R., & Matten, D. (2019). The Governance of Digital Technology, Big Data, and the Internet: New Roles and Responsibilities for Business. *Business and Society*, 58(1), 3–19. <https://doi.org/10.1177/0007650317727540>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Heeks, R. (2014). ICTs and poverty eradication: Comparing economic, livelihoods and capabilities models. In *IDPM Development Informatics Working Paper no.58*, University of Manchester, UK.
- Heeks, R. (2020a). ICT4D 3.0? Part 1—The components of an emerging “digital-for-development” paradigm. *Electronic Journal of Information Systems in Developing Countries*, 86(3), 1–15. <https://doi.org/10.1002/isd2.12124>
- Heeks, R. (2020b). ICT4D 3.0? Part 2—The patterns of an emerging “digital-for-development” paradigm. *Electronic Journal of Information Systems in Developing Countries*, 86(3), 1–13. <https://doi.org/10.1002/isd2.12123>
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering - A systematic literature review. *Information and Software Technology*, 51(1), 7–15. <https://doi.org/10.1016/j.infsof.2008.09.009>
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.
- Kumar, P., & Sharma, M. (2021). Data, Machine Learning, and Human Domain Experts: None Is Better than Their Collaboration. *International Journal of Human-Computer Interaction*, 00(00), 1–14. <https://doi.org/10.1080/10447318.2021.2002040>
- Levy, Y., & Ellis, T. J. (2006). A Systems approach to conduct an effective literature review in support of Information Systems research. *Informing Science Journal*, 9.

- Mann, S., & Hubert, M. (2020). AI4D: Artificial Intelligence for Development. *International Journal of Communication*, 14, 4385–4405.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys (CSUR)*, 54(6). <https://doi.org/10.1145/3457607>
- Mgala, M., & Mbogho, A. (2015). Data-Driven Intervention-Level Prediction Modeling for Academic Performance. *ACM International Conference Proceeding Series*, 15. <https://doi.org/10.1145/2737856.2738012>
- Mlitwa, N. B. W., & Simbarashe, M. (2019). A Cloud-based architecture for a regional trans university learning management system collaboration on digital content delivery across Southern Africa. *EDULEARN19 Proceedings*, 1(July), 7819–7827. <https://doi.org/10.21125/edulearn.2019.1903>
- Page, M., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., ... Moher, D. (2017). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. In *BMJ: British Medical Journal*, 372(n71). <http://www.prisma-statement.org/PRISMAStatement/FlowDiagram>
- Sein, M. K., Thapa, D., Hatakka, M., & Sæbø, Ø. (2019). A holistic perspective on the theoretical foundations for ICT4D research. *Information Technology for Development*, 0(4), 7–25. <https://doi.org/10.1080/02681102.2018.1503589>
- Sen, A. (2001). *Development as freedom*. Oxford Paperbacks.
- Smith, M. L., & Neupane, S. (2018). Artificial intelligence and human development Toward a research agenda. In *International Development Research Centre (IDRC/CRDI)*. <https://idl-bnc-idrc.dspacedirect.org/handle/10625/56949>
- Toyama, K. (2010). Human-Computer interaction and global development. *Foundations and Trends in Human-Computer Interaction*, 4(1), 1–79. <https://doi.org/10.1561/11000000021>
- van Biljon, J., & Renaud, K. (2019). Human-Computer Interaction for Development ( HCI4D ): The Southern African Landscape. *IFIP 9.4 International Conference on Social Implications of Computers in Developing Countries*, 1, 253–266. <https://doi.org/10.1007/978-3-030-19115-3>
- Walsham, G. (2017). ICT4D research: reflections on history and future agenda. *Information Technology for Development*, 23(1), 18-41.
- Weber, J. S., & Toyama, K. (2010). Remembering the past for meaningful AI-D. *AAAI Spring Symposium - Technical Report, SS-10-01*(September), 97–102.
- World Bank Report. (2019). *The Changing Nature of Work. World Developemnt Report*. <http://documents.worldbank.org/curated/en/816281518818814423/2019-WDR-Report.pdf>