



HAL
open science

Big Data Readiness Index – Africa in the Age of Analytics

Anke Joubert, Matthias Murawski, Markus Bick

► **To cite this version:**

Anke Joubert, Matthias Murawski, Markus Bick. Big Data Readiness Index – Africa in the Age of Analytics. 18th Conference on e-Business, e-Services and e-Society (I3E), Sep 2019, Trondheim, Norway. pp.101-112, 10.1007/978-3-030-29374-1_9 . hal-02510093

HAL Id: hal-02510093

<https://inria.hal.science/hal-02510093v1>

Submitted on 17 Mar 2020

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

Big Data Readiness Index – Africa in the Age of Analytics

Anke Joubert¹, Matthias Murawski¹ and Markus Bick¹

¹ ESCP Europe Business School Berlin, Berlin, Germany
{mmurawski, mbick}@escpeurope.eu
anke.joubert@edu.escpeurope.eu

Abstract. Big data promises to drive economic growth and development, but if not applied across borders it can lead to a greater digital divide. There are, however, no indexes measuring big data readiness on country level. Most existing indexes cover topics such as digitalization and focus on developed economies with a clear underrepresentation of Africa. Thus, the underlying question to answer is: what are the required components for an index measuring big data readiness in Africa? The design science approach is used to design the Big Data Readiness Index (BDRI). The BDRI is developed in line with the academic definition of big data under the five V's: volume, variety, velocity, veracity and value, and consists of 75 indicators, grouped into drivers that build up the five components. We apply this index to all African countries which yields, amongst others, country rankings that show relatively high BDRI levels for coastal countries such as South Africa, Kenya and Namibia as well as for islands such as Mauritius. Limitations and further development of the BDRI are discussed.

Keywords: Africa, Big Data, Design Science Research, Readiness Index

1 Introduction

Big data analytics has gained substantial interest among academics and business practitioners in the digital era and promises to deliver improved operational efficiency, drive new revenue streams, ensure competitive advantage and facilitate innovation, amongst other things [1, 2]. Big data is not only of interest to academics, but also to governments, businesses and the general public [3]. Fosso-Wamba et al. (2015) state that the topic 'big data' was searched on Google about 252 000 times in November 2011, whereas the peak is at 3.13 billion in March 2017 [4]. This shows not only the increasing interest in the subject of big data, but also indicates that the topic is not losing momentum.

The concept of data-driven decision making and automation through big data can be applied across various sectors including healthcare, security, services, manufacturing, retail and technology. In order to gain value, the concept of big data must be clearly defined and understood, both in terms of its benefits, application possibilities, as well as the shortcomings and risks involved - these include for example privacy concerns, human resource scarcity, infrastructure shortages and lack of institutional governance [1].

As countries differ substantially across these fields, we believe it would be interesting to compare the structural, technological and institutional readiness for countries to

adapt big data technologies in order to harvest the economic benefits. Most authors of published big data related literature up until 2015 are from China, followed by the USA, Australia, the UK and Korea [5]. Africa produces less than 1% of the world's research even though 12.5% of the world's population is from Africa [6]. This unequal geographic coverage of literature raises a vital question as to whether the lack of interest in African countries is due to a global big data divide or whether it is due to a lack of essential knowledge and proficiency to undertake big data related research within African countries.

Literature focusing on firm level e-business adaption shows that technology readiness has the strongest impact on facilitating assimilation of e-business in developing countries [7]. Lawrence & Tar (2010) [8] discuss issues hindering e-commerce in developing countries, concluding that the absence of basic infrastructure, socio-economic factors and inadequate government ICT strategies slows down adoption and growth. Furthermore, e-readiness does not have an established theoretical basis with the existing e-readiness measurement tools largely focusing on ICT, business and policy, underplaying the information access factor [9]. Thus, the lack of country level research in this field is established and leading to our research focus.

Furthermore, the Digital Transformation and Sustainability (DTS) model proposed by Pappas et al. (2018) [10] explains how different data actors and big data analytical capabilities interact to create sustainable societal value through critical factors such as a data-driven organisational culture, technology investment and technical and managerial skills. They mention the need for further research to focus on data capacities and availability in the digital society to take into account the capacity and availability of big data, as well as differences between countries, continents, and cultures [10].

In this paper, we will attempt to develop an index that can benchmark Big Data Readiness in the age of analytics. The focus area for data application will be the African continent. Establishing such an index will not only allow comparison between countries but will also set a benchmark of the indicators and topics which should be covered by an index measuring big data readiness. A new index covering Africa will help to fill the current gap in big data research for this region as well as the gap left by the existing indexes' limited African coverage. Based on this, we formulate the following research question:

What are required components for a big data readiness index (BDRI) and how could a comprehensive BDRI be designed?

As previously mentioned, the developed index will be applied to African countries with publicly available data. Before entering this empirical part, we start with a literature review which will yield a brief overview of big data. In the next step, Hevner's design science research approach [11] (outlined in the methodology section) will be applied as a baseline for constructing the index, including descriptions of the data collection, cleaning, weighting and normalization process. The section that follows includes an analysis of the BDRI by using data about African countries. The final section concludes by summarizing key findings and main contributions.

2 Theoretical Background

2.1 Introduction to big data

Big data analytics promises to deliver improved operational efficiency, drive new revenue streams, ensure competitive advantage and facilitate innovation, amongst other things [2, 12]. Fosso-Wamba et al. (2015) [4] summarize some descriptions of the impact of big data in previous literature as the *next big thing in innovation* [13]; *the fourth paradigm of science* [14]; *the next frontier for innovation, competition, and productivity* [15]; *the next management revolution* [16]; and that big data is *bringing a revolution in science and technology* [2].

There is no globally accepted definition for the term ‘big data’. The complexity of defining this term was aggravated by a shared origin between academia, industry, the media and widespread public interest. The fact that various stakeholders provide diverse and even contradictory definitions, leads to the emergence of literature that attempts to establish a common definition.

The first concepts predominantly associated with big data are far from new and relate to data storage and data analysis. An early definition describes big data using the three V’s: volume, velocity, and variety. This approach has been reiterated in various studies [16, 17]. Volume refers to the size of the data, whereas velocity refers to the speed of data generation and the frequency of data delivery. Variety takes the data types into consideration and highlights the fact that big data is generated from a large number of sources and formats that include structured and unstructured data [17].

A fourth V for value was subsequently suggested. Value implies that extracting economic benefits from the data is a key attribute of big data analytics [5, 18, 19]. A fifth V for veracity followed. The fifth V is sometimes referred to as verification in literature [20]. Veracity attempts to stress the importance of data quality and that trusting the various data sources is integral. This allows analysts and users to be sure that security measures and techniques are in place to assure trustworthy data analysis [4]. The use of inaccurate data will lead to biased and inconsistent results and recommendations. Both value and veracity are considered as important in this study, as without added value and trust towards the quality of the data used, the benefits of big data analytics deteriorate.

2.2 Big data in developing countries

If not applied across borders, increased data generation and the availability of analytical tools, can lead to an even greater digital divide – where countries who successfully implement this will advance leaving the rest behind. The digital divide originated as a concept to describe the gap between countries that have and do not have access to computers and the internet. The digital divide does not only refer to physical access but also to skills and usage of technology. Van Dijk (2006) writes that “in terms of physical access the divide seems to be closing in the most developed countries; concerning digital skills and the use of applications the divide persists or widens” [21, p. 221].

A framework for ICT development strategies that can enable countries to participate competitively in the emerging digital economy shows an acute global shortage of high skilled staff needed to steer the emerging digital economy in both developed and developing countries [9]. Hilbert (2016) mentions that privacy concerns, human resource scarcity, infrastructure shortages, limited economic resources and institutions are areas that can be barriers to developing countries, keeping them from successful digitalization [22]. Although additional challenges exist that hinder developing countries to easily implement big data analytics, the positive impacts on developing countries' economies can be even greater. Thus, big data in itself can help alleviate the initial shortcomings that are barriers to the implementation of big data. For example, due to structural inconsistencies, statistical authorities in Ghana took 17 years to adopt the UN national accounting system. Afterwards, Ghana's GDP was 62% higher than previously estimated [23]. Some of the main opportunities big data can offer to developing countries are early warnings in times of crisis, detecting digital media and other anomalies, real-time awareness of data changes, a more honest representation of reality and real-time feedback. This can lead to early interventions for policies or programs that deliver different results than expected.

Bifet (2013) mentions that Global Pulse, a United Nations initiative launched in 2009, is using big data for the benefit of developing countries. This involves researching innovative techniques for real-time data analysis, detecting early emerging vulnerabilities and sharing open source real-time data among those integrated in the global network of Pulse Labs [24].

Some other interesting examples of how big data analytics is being applied to assist development include using machined roof counting to measure poverty in Uganda, using mobile phone data and airtime credit purchases to estimate food security in East Africa and mining citizen feedback data in order to gain input for government decision making in Indonesia [25]. These are only some of the many ongoing projects where big data analytics can be implemented for positive and productive outcomes in developing countries.

3 Research design

3.1 Index development procedure

Developing an index will be approached in the same manner as designing an artefact in the context of design science research [11]. This approach is relevant as the BDRI is applicable to a technology related field. The index will be designed and evaluated accordingly. The seven steps of the design science approach will be applied to designing an index in the following manner:

Steps	Description	Application
1. Design as an artefact	<i>Produce a viable artefact in the form of a model, method, or instantiation</i>	In this research constructing an index is the artefact produced by the design science approach
2. Problem relevance	<i>Develop technology-based solutions to relevant business problems</i>	The BDRI fits into a technology related field, with very relevant implications for technological development and policy recommendations
3. Design evaluation	<i>Demonstrate the utility, quality, and efficacy via well-executed evaluation</i>	The BDRI is designed and tested with the help of evaluating four existing indexes in the digitalisation field and application is shown on a scenario basis
4. Research contributions	<i>Provide clear and verifiable contributions</i>	The BDRI overcomes geographical and conceptual gaps in research. The components and underlying drivers set out methodology of what to include when measuring big data.
5. Research rigor	<i>Rely on rigorous methods in construction and evaluation</i>	The application of the index shows research rigor. After careful construction the BDRI is compared to other indexes to evaluate ranking reliability.
6. Design as a search process	<i>Search requires utilizing the available means to reach desired ends</i>	Data collection followed a two phased methodological approach in order to overcome the shortcomings and limitations of missing data in Africa
7. Communication of research	<i>Present results effectively to technology and management audiences</i>	The BDRI consists of five components with three drivers each and can easily be understood by different audiences. The development and application is well documented

Fig. 1. Application of the design science approach

As described in the previous section, big data has been researched intensively and defined by many researchers. The definition as followed by this paper includes the five V's: volume, variety, velocity, veracity and value. Thus, the five V's also form the five components of the BDRI.

Each component is formed up by three drivers. These drivers are inspired by an analysis of existing digital related indexes such as the Enabling Digitalization Index (EDI), the Digital Evolution Index (DEI), the IMD World Digital Competitiveness (WDC) and the Networked Readiness Index (NRI) [26–30]. Certain components from each index could partially cover certain aspects of the five V's. These aspects were evaluated on a basis of relevance by referring to the definition as outlined in the literature review. Complementary drivers based on available data were added to give a global overview of big data. A total of 170 variables were collected and after analysis, taking missing values into account, and referring to the base definition, 75 variables were selected and aggregated in groups of five to build the fifteen drivers that form the basis of the final five components of the BDRI which is visualized in Figure 2.

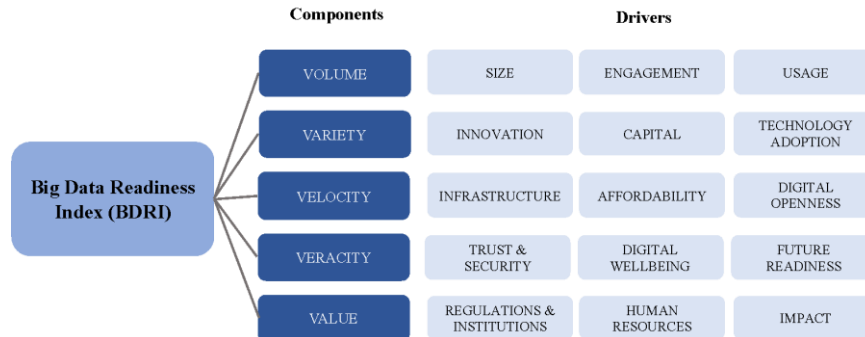


Fig. 2. Components of the BDRI

3.2 Data collection

The data collection process followed a two-phased methodological approach. In the first phase, a comprehensive overview of the reviewed indexes' data sources was completed. This allowed an assessment of relevant available data to include in a BDRI, as well as ways to set up an index, measure outcomes, assign weights, and normalize variables. It also indicated which first-hand sources cover all the African countries. Some of the sources [28, 29] could be reused to build up and define new components of the BDRI, but due to the limited coverage of Africa and the fact that big data requires other components not covered in the reviewed digitalization indexes, this was not exhaustive. This phase gave a broad overview that helped identify the gaps in data availability to construct the BDRI.

The second phase considered alternative sources to measure Africa's big data readiness. Due to data scarcity for some African countries, this phase had its obstacles. Dealing with data recency and missing values was a major part of the data cleaning process. This phase was successful in finding additional open data sources to help build the BDRI and also helped to verify and fill gaps for data collected in the first phase.

3.3 Evaluation of the index

A thorough comparison of the created BDRI with the evaluated indexes (EDI, DEI, IMD WDC, and the NRI) was performed to function as a design evaluation. Even though the compared indexes have a less extensive coverage of Africa, these indexes provided the possibility to benchmark the countries that had been covered in similar topics and compare these rankings to the rankings of the BDRI. The evaluation concluded that the BDRI output is aligned with expectation and that the differences in rankings could be accounted for by the fact that the BDRI takes different features that relate to big data into account.

Hevner et al. (2004) state that scenarios can be applied as an evaluation technique for innovative artefacts. This is particularly relevant for new and complex artefacts which cannot be evaluated in one step [11]. Three scenarios that would be applicable

use cases for the BDRI include governmental policy development, market entry analysis and institutional usage. The diverse scenarios illustrate research rigor through the wide range of usage possibilities and display how the development of the BDRI fits into the information system research framework.

4 Findings and discussion

4.1 Country rankings

We calculated the BDRI for all African countries. Ranking the countries by score allowed an analysis of the top performers. Figure 3 shows the top ten performers.



Fig. 3. BDRI - Top 10 African countries

Rwanda (position 3) is the only country in the top 10 that is not a coastal country. Furthermore, three islands, Mauritius (position 1), the Seychelles (position 4) and Cape Verde (position 10) belong to the top performing BDRI countries in Africa. South Africa, an influential economy in Africa, ranks in second position. Figure 4 shows the evolution of the BDRI.

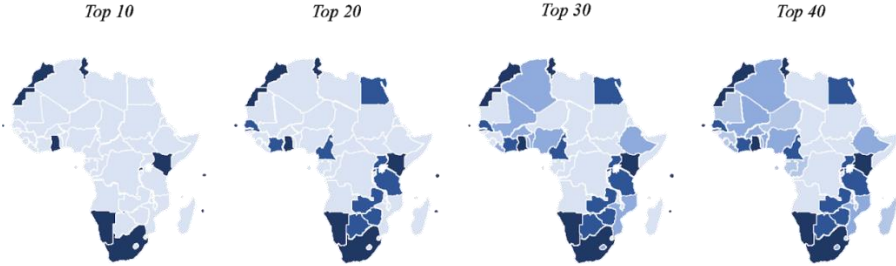


Fig. 4. BDRI - Top 10, 20, 30, and 40

Southern Africa is the top performing region, with Eastern Africa following in the top 20. The top 30 grows to include Nigeria and the neighboring Benin and Togo and coastal Liberia as well as Burkina Faso, Mali and Algeria in West Africa. Nigeria, one of Africa's largest economies, ranks at position 21. The top 40 fills the neighboring gaps in West Africa, with Guinea-Bissau as the only exception.

Looking at the location of the top performers, coastal countries perform well and neighbors can influence each other in terms of technology adoption.

Evaluating overall performance by component shows that volume produced the most unequal performance. This indicates a digital divide within Africa in terms of size, engagement and usage. On the opposite end, countries show less diverse performance in variety and veracity, yet this narrow spread has a low base. On average, velocity and value show a higher overall performance. Noteworthy positive outliers include Mauritius (usage), the Seychelles (usage and infrastructure) and South Africa (size, innovation and infrastructure). On the opposite end recurring negative outliers are Somalia, South Sudan and the Central African Republic.

4.2 Hierarchical clustering analysis

Hierarchical cluster analysis is an algorithm that groups similar objects together in clusters to get a set of clusters that are distinct from one another [31]. In order to perform this clustering, the raw data needs to be transformed into a distance matrix. A dendrogram was created, where the vertical axis represents the distance or dissimilarity between clusters, whereas the horizontal axis represents clusters in terms of size. Selecting a distance measure of one resulted in three clusters of similar size. The first cluster is expected to be more different to the second and third cluster that branches from the same clade of the dendrogram.

The countries are assigned to one of the three clusters in terms of similarity. We defined these three distinct groups as the countries *forging ahead* and thus outperforming the rest of Africa, the countries *gaining momentum* and thus catching up to big data implementation trends and the ones that are *lagging behind*.

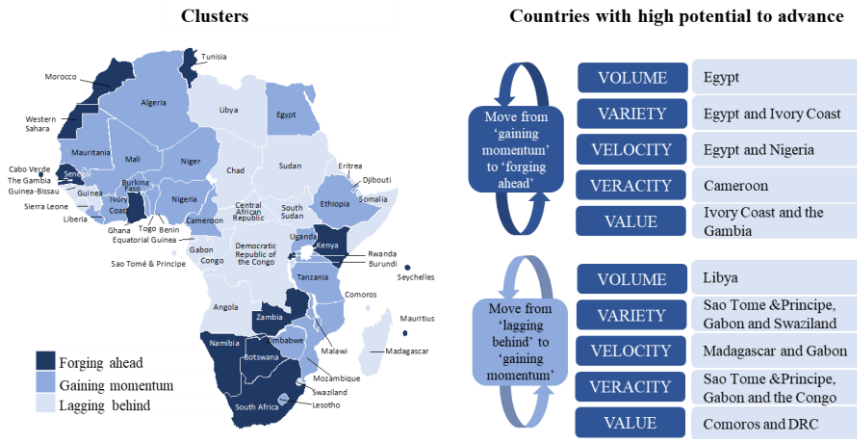


Fig. 5. Cluster analysis results

The map (Figure 5) shows that the *lagging behind* cluster that requires the most intervention is the central geographical strip from Angola in the south to Libya in the north, as well as Sudan, Eritrea and Somalia on the Eastern coast. A small cluster in West Africa, consisting of Guinea-Bissau, Guinea and Sierra Leone, is also in need of large interventions to become technologically ready to gain value from big data.

4.3 Limitations and next steps

The major limitations of developing the BDRI for Africa was induced by missing data, aggregation methodology and the limited coverage of available indicators.

Data scarcity. One of the major constraints to building an index for big data readiness in Africa is data scarcity. The limited data available makes analysis more difficult.

Aggregation methodology. A major challenge was how to overcome aggregation where some indicators of the drivers were missing. There is little research on how to deal with aggregation as some indexes, such as the Digital Evolution Index (DEI) chose to cover countries based on data availability [26].

Limited coverage. Some data that is available to include in indexes covering other countries was not available for Africa. After filling missing values using projections from data available in the past ten years, the following indicators still had a high number of missing values, making them unusable in the BDRI: Public private partnership investment in ICT consisted of 91% missing data and was thus only being available for five countries. Furthermore, software piracy rate has 86% missing values; patent applications 85%; commercial value of unlicensed software 74%; labor force with advanced education 72%; firms expected to give gifts to get a phone connection 63%; telephone faults cleared by next working day 60%; researchers in R&D 57% and average duration of power outages 52%.

Taking these three limitations into account, we include the following recommendations to help overcome these limitations, to improve the developed index and to expand its use in future research.

Look into non-traditional data sources. Future studies can look beyond the traditional data sources. Data goes beyond tables of numbers and includes society's digital footprints created by cell phones, internet searches, personal transactions and other sources [32]. Pictures, sound and social media can also be used as data sources and are creative ways to capture the relevant data – especially on a continent where data capturing is limited. Together with alternative data sources, streamlined open data would be beneficial. Support and investment in platforms such as Open Africa could lead to increased published databases and higher coverage [33]. Africa has a population of 1.2 billion people of whom around 60% are young – thus Africa has a growing number of future digital natives who will be generating increasing amounts of data [32]. This indicates a great opportunity for non-public organizations to get involved in data collection, curation, labelling and help to increase open data sources. There are many other opportunities for data collection in Africa. Data collection should not only be the responsibility of the government but could also be an opportunity for start-ups. Kenyan and Namibian mobile-phone operators have made data records available to researchers to help combat malaria, which was used to compare caller movements with malaria outbreak data to predict where the disease might spread [34]. By looking into alternative sources of data, calling for more open data and by inviting the private sector to get involved, there is hope to ultimately eliminate the missing data issues on the African continent.

Expand BDRI coverage. The BDRI could be expanded to cover countries outside Africa. We recommend covering other developing regions first, as the BDRI methodology is designed for measuring big data readiness specifically in developing countries.

Comparability. We recommend that the index is measured annually to create a basis for comparison. This will assure that country specific progress can be tracked, and the effectiveness of policy interventions can be compared over time. The inclusion of an in-depth country overview is also suggested.

Index evaluation. At the current stage the BDRI is evaluated on a scenario basis. Users can be asked to use the index for their purposes and give feedback that could lead to redesign.

5 Conclusion

Big data has gained substantial interest among academics and business practitioners in the digital era, and with 2.41 billion searches for this term in July 2018 [35], this topic is not losing momentum. A new index, the Big Data Readiness Index (BDRI) that covers Africa extensively was developed in this paper. Data availability and limited research were key challenges, but the emergence of mobile technology which in Africa is double as high as global rates ensures higher big data potential [36]. The seven steps of the design science approach [11] were applied throughout

design, development and evaluation of the BDRI. Research rigor is ensured through careful construction, comparison and evaluation of the ranking reliability. Finally, the BDRI is set up in an understandable format based on the five V's.

Looking at the rankings shows that coastal countries and islands perform best, Rwanda being the only non-coastal country in the top 10. Neighboring countries influence each other through clear diffusion patterns. Outlier detection shows noteworthy positive outliers within components such as Mauritius and the Seychelles in the value component and South Africa in some components from volume, variety and velocity. On the opposite end recurring negative outliers are Somalia, South Sudan and the Central African Republic. The hierarchical clustering analysis reinforces these results by classifying the central geographical strip as *lagging behind*. The larger distance between the *forging ahead* group and the other two clusters show that the top performers are far ahead in terms of big data readiness.

The development of the BDRI adds value in various ways. Firstly, by linking big data theory to propose a practical way to measure big data, we contribute to the development of, amongst others, 'big data analytics capability' [37]. Secondly, the BDRI covers Africa extensively, overcoming the geographical gap mentioned by scholars [5] and seen in underrepresentation of Africa in most existing digitalization indexes. Thirdly, no prior research could be found that explicitly focuses on measuring big data as all the reviewed indexes focused more on digitalization and ICT with the focus on firm level e-business.

African countries might have similar weaknesses and strengths and analysis will help countries to systematically review available policy options to foster data-driven opportunities and maximize big data driven growth and development. In this way, Africa will be able to overcome developmental challenges and play an increasing role in the age of analytics.

References

1. Grover, V., Chiang, R.H.L., Liang, T.-P., Zhang, D.: Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems* 35, 388–423 (2018)
2. Keller, S.A., Koonin, S.E., Shipp, S.: Big data and city living - what can it do for us? *Significance* 9, 4–7 (2012)
3. Agarwal, R., Dhar, V.: Editorial —Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. *Information Systems Research* 25, 443–448 (2014)
4. Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., Gnanzou, D.: How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics* 165, 234–246 (2015)
5. Sivarajah, U., Kamal, M.M., Irani, Z., Weerakkody, V.: Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research* 70, 263–286 (2017)
6. Duermeijer, C., Amir, M. and Schoombee, L.: Africa generates less than 1% of the world's research; data analytics can change that. An in-depth analysis of the

- continent's research reveals promising developments – and strategies for continued improvement, <https://www.elsevier.com/connect/africa-generates-less-than-1-of-the-worlds-research-data-analytics-can-change-that>
7. Zhu, K., Kraemer, K.L., Xu, S.: The Process of Innovation Assimilation by Firms in Different Countries: A Technology Diffusion Perspective on E-Business. *Management Science* 52, 1557–1576 (2006)
 8. Lawrence, J.E., Tar, U.A.: Barriers to ecommerce in developing countries. *Information, Society and Justice* 3, 23–35 (2010)
 9. Mutula, S.M., van Brakel, P.: E-readiness of SMEs in the ICT sector in Botswana with respect to information access. *The Electronic Library* 24, 402–417 (2006)
 10. Pappas, I.O., Mikalef, P., Giannakos, M.N., Krogstie, J., Lekakos, G.: Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. *Inf Syst E-Bus Manage* 16, 479–491 (2018)
 11. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. *MIS Quarterly* 28, 75–105 (2004)
 12. Chen, H., Chiang, R.H.L., Storey, V.C.: Business intelligence and analytics: from big data to big impact. *MIS Quarterly* 36, 1165–1188 (2012)
 13. Gobble, M.M.: Big Data: The Next Big Thing in Innovation. *Research-Technology Management* 56, 64–67 (2013)
 14. Strawn, G.O.: Scientific Research: How Many Paradigms?, <https://er.educause.edu/-/media/files/article-downloads/erm1231.pdf>
 15. Manyika, J, Chui, M, Brown, B, Bughin, Dobbs, R and Roxburgh, et al.: Big data: The next frontier for innovation, competition, and productivity, <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>
 16. McAfee, A., Brynjolfsson, E.: Big data: the management revolution. *Harvard Business Review* 90, 61–67 (2012)
 17. Russom and P: TDWI Best Practices Report | Big Data Analytics | Transforming Data with Intelligence, <https://tdwi.org/research/2011/09/best-practices-report-q4-big-data-analytics.aspx?tc=page0&tc=assetpg&tc=page0&m=1>
 18. Murawski, M., Bick, M.: Demanded and Imparted Big Data Competences: Towards an Integrative Analysis. *Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017*, 1375–1390 (2017)
 19. Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M., Feldberg, F.: Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems* 26, 191–209 (2017)
 20. Opresnik, D., Taisch, M.: The value of Big Data in servitization. *International Journal of Production Economics* 165, 174–184 (2015)
 21. van Dijk, J.A.G.M.: Digital divide research, achievements and shortcomings. *POETICS* 34, 221–235 (2006)
 22. Hilbert, M.: Big Data for Development: A Review of Promises and Challenges. *Dev Policy Rev* 34, 135–174 (2016)
 23. Devarajan, S.: Africa's statistical tragedy, <https://blogs.worldbank.org/african/africa-s-statistical-tragedy>
 24. Bifet, A.: Mining Big Data in Real Time. *Informatica* 37 (2013)

25. Decuyper, A., Rutherford, A., Wadhwa, A., Bauer, J.-M., Krings, G., Gutierrez, T., Blondel, V.D., Luengo-Oroz, M.A.: Estimating Food Consumption and Poverty Indices with Mobile Phone Data (2014)
26. Chakravorti, B., Chaturvedu, R. S., Troein, C., Pagan, C., Filipovic, C. & Beck, M.: Digital Planet: How competitiveness and trust in digital economies vary across the world,
https://sites.tufts.edu/digitalplanet/files/2017/05/Digital_Planet_2017_FINAL.pdf
27. Euler Hermes: Enabling Digitalization Index 2018: measuring digitagility, <http://www.eulerhermes.com/mediacenter/news/Pages/Enabling-Digitalization-Index-2018-Measuring-digitagility.aspx>
28. International Telecommunications Union: ICT Development Index, <https://www.itu.int/net4/ITU-D/idi/2017/#idi2017economycard-tab&AFG>
29. World Economic Forum: The Global Information Technology Report: Innovating in the digital economy.,
http://www3.weforum.org/docs/GITR2016/GITR_2016_full%20report_final.pdf
30. IMD World Competitiveness Centre: World Digital Competitiveness Rankings 2017, <https://www.imd.org/wcc/world-competitiveness-center-rankings/world-digital-competitiveness-rankings-2018/>
31. Bock, T.: What is Hierarchical Clustering?, <https://www.displayr.com/what-is-hierarchical-clustering/>
32. Sengeh, D.: Missing Data: Achieving the Value of Big Data Analytics in Africa, <https://www.linkedin.com/pulse/missing-data-achieving-value-big-analytics-africa-david-sengeh>
33. Open Africa.: openAFRICA,
https://africaopendata.org/dataset?q=&sort=score+desc%2C+metadata_modified+desc
34. The Economist: Off the map,
<https://www.economist.com/international/2014/11/13/off-the-map>
35. Google Trends, <https://trends.google.com/trends/?geo=US>
36. Matinde, V.: Africa's missed opportunity in the multi-billion dollar mobile phone industry, <https://www.idgconnect.com/idgconnect/opinion/1029329/africas-missed-opportunity-multi-billion-dollar-mobile-phone-industry>
37. Gupta, M., George, J.F.: Toward the development of a big data analytics capability. *Information & Management* 53, 1049–1064 (2016)