Opportunities and Challenges of Digitizing the Education Sector in Developing Countries

Can Learning Apps Be Used to Inform Policymaking in The Gambia?

Jessie Yue Guan

Master's in Programming & System Architecture
60 Credits

Department of Informatics
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

Spring 2022

Opportunities and Challenges of Digitizing the Education Sector in Developing Countries

Can Learning Apps Be Used to Inform Policymaking in The Gambia?

Jessie Yue Guan

Opportunities and Challenges of Digitizing the Education Sector in Developing Countries

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

Acknowledgements

I would like to first and foremost thank my supervisor Terje Aksel Sanner for his knowledge of, insight into, and reflections around education management information systems. I would also like to thank Marit Linnebo Olderheim for pushing me out of my comfort zone and making things happen.

I would like to express my deepest admiration of Kim Frost’s friendliness and cheerfulness, Pamod Amarakoon’s wisdom and tranquility, and Sophia Kousiakis’ diligence and hospitality. Many thanks to Lamin Conteh for setting up interviews and tours and for driving us to them personally, to Lucas Barrow-Townsend for showing us around and being so passionate about his work, and to Bjørnar Valbø for keeping me company throughout the trip.

I am very grateful for my wonderful boyfriend for always being there for me, showing me so much love, and bringing so much joy into my life. I am also grateful for my dear classmate and study buddy for keeping me company through long days at the library. This thesis would not have been possible without them.

Finally, I am very thankful that the work of this thesis was mostly done after the pandemic restrictions had been lifted. Getting the extraordinary opportunity to take a field trip to The Gambia has been amazing, memorable, and humbling. Being able to spend quality time with my partner, friends and family outside in the fresh air under the bright sun on days off has filled me with a lot of energy and motivation.

Abstract

This thesis aims to look at the opportunities and challenges of digitizing the education sector in developing countries with the integration of data from Leap Learning apps into DHIS2 dashboards for The Gambia as an example of this. First, a literature review of representation theory, technology-enhanced learning, educational data mining, and learning analytics will be performed with a special focus on criticisms and applications in low-resource contexts. Then, an introduction to DHIS2, Leap Learning, and The Gambia will be provided. Next, thematic analysis will be used on documents, workshops and interviews in order to answer the research questions and a detailed summary of the results will be described. Finally, these results will be discussed in light of the topics from the literature review.

Table of Contents

List of Figures

List of Tables

List of Acronyms

3IR Third Industrial Revolution

4IR Fourth Industrial Revolution

BDRI Big Data Readiness Index

CMM Cluster Monitor Manual

CR Completion Rate

CREDD Curriculum Research, Evaluation and Development Directorate

DHIS2 District Health Information Software 2

DM Data Mining

DPG Digital Public Good

EBD Educational Big Data

ECD Early Childhood Education

EDM Educational Data Mining

EGMA Early Grade Mathematics Assessment

EGRA Early Grade Reading Assessment

EMIS Education Management Information System

ESSP Education Sector Strategic Plan

ESY16 Education Statistics Yearbook 2016

ESY21 Education Statistics Yearbook 2021

GABECE Gambia Basic Education Certificate Examinations

GER Gross Enrollment Rate

GIR Gross Intake Rate

HEI Higher Education Institutions

HISP Health Information Systems Programme

HMIS Health Management Information System

IS Information System

KDD Knowledge Discovery in Databases

LA Learning Analytics

LBE Lower Basic Education

LMIC Low and Middle Income Countries

LMS Learning Management System

MOOC Massive Open Online Course

MoBSE Ministry of Basic and Secondary Education

MoHERST Ministry of Higher Education, Research, Science and Technology

NAT National Assessment Test

NER Net Enrollment Rate

NST National Standardized Test

PTA Parent Teacher Association

RT Representation Theory

SDG Sustainable Development Goal

SIS Student Information System

SMC School Management Committee

SMM School Management Manual

SMT Senior Management Team

SSE Senior Secondary Education

STED Science and Technology Education Directorate

TEL Technology-Enhanced Learning

TT Teacher Trainee

TaD Teachers as Designers

UBE Upper Basic Education

WASSCE West-African Secondary School Certificate Examinations

Chapter 1

Introduction

1.1 Motivation

> "Architects want to build better buildings, medics want to help people live longer and healthier lives, engineers want to build more effective technological systems to improve efficiency and artists want to stimulate our subtler senses with their work. [Information systems] scholars and practitioners should be concerned with how to use [information and communication technologies] to help make a better world, where everybody has the opportunity and capability to use technologies to make better lives for themselves, their communities and the world in general." (Walsham, 2012, p. 89).

1.1.1 Digital Transformation

We are now at the dawn of the Second Machine Age and the Fourth Industrial Revolution. The first-mentioned was introduced by Erik Brynjolfsson and Andrew McAfee (Spencer, 2017). The birth of industry and machine manufacturing is referred to as the First Machine Age and the recent great strides in robotics and artificial intelligence is called the Second Machine Age. The negative consequences of this is that "[these] advances (…) are going to sweep away many existing jobs, leading to potentially higher unemployment and greater inequality" (p. 142-143). The positive effects of this is that "such advances bring forth the possibility of a bountiful future of less toil, more creative work, and greater human freedom" (p. 143). However, Brynjolfsson and McAfee have been criticized for implicitly taking on a capitalist view on advancement. Since most new technologies are developed under free market economies, they will inevitably support profit generation and favor managers over workers. The last-mentioned was conceived by Klaus Schwab (Xu et al., 2018). The invention of the steam engine and the transition from agriculture to manufacturing is known as the First Industrial Revolution. The innovation of the internal combustion engine and the introduction of electricity and oil marks the Second Industrial Revolution. The launch of the mainframe computer and the implementation of information systems to automatize tasks and maximize productivity is viewed upon as the Third Industrial Revolution (3IR). The evolution of the

Internet and the blurring of the digital, physical and biological signals the Fourth Industrial revolution (4IR). This includes new technologies like mobile devices, smart sensors, cloud computing, blockchain, augmented/virtual reality, data analytics, artificial intelligence, and state-of-the-art robotics that have the potential to transform and disrupt both corporations and institutions within all sectors. For each revolution, there are massive increases in productivity, shifts from manual labor to idea generation, fusion of different technologies or disciplines and accelerating momentum of innovation. However, they also come with displacement of jobs and markets, increased segregation and inequality, complex ethical and privacy issues, and highly alarming security threats.

1.1.2 Information Systems

An in-depth description and discussion of all these changes would be a great undertaking and outside the scope of this thesis. However, what all of these technological innovations have in common is that they are information systems (IS) which have wide-ranging after-effects and therefore must be understood from a sociotechnical and organizational perspective. ISs as an academic field was established during the 3IR when computers were few, large, and expensive (Sørensen, 2016). At this point, they were mainly used by enterprises to automate business processes, and because of this, early ISs research has focused heavily on management (Watson, 2014). As a result, it became a highly interdisciplinary field that borrowed theories from economy, sociology, psychology amongst others, but with no theoretical foundation of its own. “The field has often wrestled with justifying its existence and establishing itself as a legitimate academic endeavor” (p. 516). Since then, technological advancements have accelerated under the 4IR and computers have become many, small, and affordable (Sørensen, 2016). Digitalization of all sectors has led to the development of large infrastructures and platforms. Distribution of computers have led to an explosion of ubiquitous mobile devices. Scaling of computers have led to the realization of artificial intelligence and robotics. ISs are now coming in numerous shapes, applied to various settings, and used for different goals. This has exacerbated the “identity crisis” of ISs as a field of research (Walsham, 2012). A possible point of convergence for all ISs could be representation theory (RT), “one of the few native theories of [ISs] that seeks to describe the fundamentals of the phenomenon we study” (Fossum & Kempton, 2020, p. 2). It asserts that:

> [The] fundamental purpose of an [IS] is to provide users with representations of real-world phenomena. There are things out there - factories, airplanes, people, trees, ideas - which we want represented in here, so that we can understand, work with, and coordinate around these things, without observing or manipulating them directly. Digital representations enable workspaces where aspects of the world previously separated because of time, space, or human inability to handle scope, can now through modelling, data capture, and computation be combined and represented. (p. 1)

The *capture* of big data, the *modelling* of data mining techniques, and the *computation* of data analytics are some of the ways in which digital representations are made. In order to narrow down the focus on the text, this thesis will only look at these technologies which have been taking the research community by storm in the last decade (Joubert et al., 2021). Some believe that the mainframe heritage of ISs is a detriment to its academic agility, which makes it challenging for the field to address big data and artificial intelligence (Sørensen, 2016). Others argue that the organizational focus of ISs, its emphasis on "value creation through data", and its cross-disciplinary history all work to its benefit in addressing big data and data analytics (Agarwal & Dhar, 2014, p. 445).

1.1.3 Education

"The concepts of data-driven decision-making and automation through big data can be applied across various sectors, including security, services, manufacturing, retail, and healthcare" (Joubert et al., 2021, p. 1). Nonetheless, it can be argued that education is perhaps one of its most important areas of research and application for several reasons. On the one hand, it can give rise to complex ethical challenges. First, the focus on resource optimization can lead to the commercialization and commoditization of education which again can give rise to skewed power relations, surveillance, exploitation similar to that of a Panopticon (Drachsler & Greller, 2016; Slade & Prinsloo, 2013). Second, most students are not of legal age and thus either unable to fully comprehend the possible consequences of data collection or independently make decisions regarding their privacy without approval from a guardian (Drachsler & Greller, 2016). Third, students from discriminated groups in vulnerable developmental stages run the risk of being trapped in cycles of self-fulfilling prophesies if an algorithm is biased or its results are misused or leaked (Slade & Prinsloo, 2013; Wang, 2016). This is why the application of big data, data mining, and data analytics in education has been under considerably scrutiny and lagging behind that of other sectors (Drachsler & Greller, 2016; Siemens, 2013). While "several authors refer to the obligation that institutions have to act on knowledge gained through analytics[,] it is fair to say that there may also be instances where institutions [should] not to act on data" (Slade & Prinsloo, 2013, p. 1513). On the other hand, it can also bring about promising learning opportunities. First, the domain of *technology-enhanced learning (TEL)* can lead to efficiency by making existing processes more cost-effective or time-effective, enhancement by improving existing processes and their outcomes, and transformation by introducing new processes of radical positive change (Kirkwood & Price, 2013). Second, "the increased availability of [*educational big data (EBD)* and *educational data mining (EDM)*] has led to new frontiers in how we monitor, understand, and evaluate processes in educational contexts and has informed decision making and efforts to improve educational effectiveness" (Fischer et al., 2020, p. 131). Third, the field of *learning analytics (LA)* has the potential to "provide personalized feedback and learning experience, (…) identify biases affecting student success, [and] promote the development of 21st century skills" (Gašević, 2018, p. 5). The connections between big data, data mining, and data analytics on one side and EBD, EDM, and learning analytics on the other side are obvious and they are all part of the 4IR. While TEL tends to end up reproducing existing paper-based practices and can be viewed upon as part of the 3IR, it also encompasses all of the other terms and can thus also be considered a part of the 4IR (Kirkwood & Price, 2014). Nevertheless, in order to completely apprehend the disruption brought about by big data and its analysis, it is important to also be familiarized with traditional manual approaches and the mechanic automatization of them.

1.1.4 Developing Countries

Another area (both metaphorically and literally) which has been lagging behind when it comes to big data, data mining, and data analytics is the Global South. "Most countries in Latin American, Middle East, Africa, Southeast Asia, and some parts of southern Europe are designated as 'developing countries' because of their lower rank in the United Nations Development Program Human Development Index" (Gulati, 2008). Although they are often grouped together, "these countries differ in their political circumstances, the history of their educational development, culture, language, religion, gender issues, population size, resources and the contemporary influx of technology" (p. 2). In fact, some of these countries are historically considered the "cradles of education" so to speak, as the first school was built in

Egypt and the first university in Babylonia. However, because of a long history of complex socioeconomic developments, perhaps most notably Western colonialism and imperialism, the gap to the Global North is only increasing. One of these nations is The Gambia, a very small country at the West Coast of mainland Africa. It surrounds a river by the same name which was one of the first trade routes to inland Africa that Europeans battled over and sought to control as slavery became more profitable and lucrative (Britannica, 2020-b; UNESCO, n.d.). After WWII and the 3IR, the income gap between the richest and the poorest people in the world has increased from 13:1 to 74:1 (Gulati, 2008). In 1960, only one fourth of children went to school in Sub-Saharan Africa, and approximately only one half of children in Asia and Latin America. After decolonization, many developing countries started to invest in televised primary schooling aided by different donor organizations. During the 70s, they focused their efforts on improving secondary education with the help of the World Bank, but resources were mostly given to the elite few. In the 80s, oil crises, food shortages, falling prices and increased debt hit developing nations the hardest whereas developed nations got by relatively unscathed. Since the 1990s, the national policies of many developing countries have started to emphasize the development of computer skills in education which is believed to be "necessary (…) to compete in a modern, IT-driven, global economy" (p. 8). There has been very little research on the use of TEL, EDM, and LA in low- and middle-income countries. This must be changed if we are to ensure that the introduction of data mining and analytics, as well as the digitization and transformation of education in a broader sense, result in the least possible harm. Although the true motivations of the Second Machine Age and the 4IR can be questioned, there is no denying that the arrival of them is inevitable and unavoidable (Prinsloo & Kaliisa, 2021). If they are really going to lead to "greater human freedom" and "accelerating innovation", there has to be equal opportunities for everyone and a full realization of all human potential.

1.1.5 TEL Tool Requirements

EDM and LA tend to use TEL tools like learning management systems (LMS), massive online courses (MOOC), and intelligent tutoring systems (ITS) amongst others (Romero & Ventura, 2013; Pardos, 2017). Additionally, they may also employ administrative or paper-based educational data such as student information systems (SIS) and national standardized tests (NST). Furthermore, they can also utilize any data from anywhere that is relevant to learning including social media, online forums, public chatrooms, web browsers and more

(Wang, 2016). Of special interest is the “entry of open source projects in mobile computing [which] has led to low cost smartphones (…). Students have started using smart phones to access learning content” (Sin & Muthu, 2015, p. 1035). In fact, “some schools have already adopted mobile learning curricula. Students use smartphones and tablets for language learning and science inquiry” (Wang, 2016, p. 382). This has been especially relevant recently; “as institutions responded to the disruption caused by COVID-19, the digitisation and datafication of teaching and learning has intensified” (Prinsloo & Kaliisa, 2021, p. 2). At the beginning of the pandemic, UNESCO launched the Global Education Coalition for all member states of UN with a special focus on 39 African countries (UNESCO, 2021-b). In celebration of Africa Day, an online forum event was held in which education stakeholders such as ministries of education, higher education institutions, non-profit organizations, educational technology companies, and telecom operators were invited to share their experiences. It highlighted some important considerations, required specifications, and desired features of TEL tools in an African setting. Within the continent, the education sector was hit hard with schools being closed for up to 10 months and “governments responded quickly by launching educational broadcasts via television, radio or the internet” (p. 1). However, only 34% of households have internet access and 11% of learners have a computer they can use at home. In addition, it requires adaptation of curriculum content, training of teachers in digital and pedagogical skills, and children having basic digital literacy and computer skills. Moreover, a lack of virtual storage space and funds to provide digital access for all students leads to a divide between urban and rural areas. Because of this, having hybrid as opposed to purely physical or digital learning is essential – this emphasizes the significance of distributing paper-based learning material and having face-to-face classes whenever possible. Mobile learning devices with efficient energy use, optimal data storage management, and offline functionality in the form of preloaded content are also invaluable. From 2015 to 2019, investments in Africa increased from 300 million to 2 billion dollars. However, this figure is still only equivalent to those of a single European country. Unfortunately, investors tend to be drawn to countries and regions where telecommunication networks and infrastructures are already in place. This stresses the importance of collaboration between the public sector, the private sector, and international donors. They need to share and learn from each other in order to avoid “reinventing the wheel”. Data-driven and informed decision-making enabled by robust education management information systems (EMIS) can play an important role in policymaking for education both before, during, and after COVID-19. Taking all of this into account, Leap Learning is a blended flipped classroom learning program that can also be used for distance education through mobile learning apps and will therefore be the focus of this thesis (Leap Learning, n.d.-a). It is one of the few TEL tools that provide the same learning content for both classroom settings and at-home learning that has been implemented in multiple developing countries around the world.

1.1.6 EMIS Requirements

Leap Learning is only one of a large number of TEL tools that are currently being used in Africa. In the Global Education Coalition for COVID-19 alone, there are more than 66 active educational technology projects and an additional 38 more under discussion (UNESCO, 2021-b). For instance, Technovation is a mobile learning program that empowers girls in STEM by teaching them programming, leadership, and entrepreneur skills to solve problems in their local communities. Pix is a non-profit adaptive and personalized learning platform that assesses, develops, and certifies digital skills for overcoming real-world challenges using gamification. Curious Learning "creates and curates open-source localized early learning content apps" in order to provide education resources and assessment tools within literacy and numeracy and distribute them on mobile devices (p. 8). LabXchange provides virtual labs and lessons targeted towards learners above the age of 12 that lack access to them in which the content is aligned with national curricula and localized into different languages. Finally, Blackboard is a LMS provided by an international educational technology company with 22 years of experience working with 38 African ministries of education. It provides data-driven personalized learning experiences, virtual interactive classrooms, data reporting systems, and digital education certifications that can be used for both primary, secondary, and higher education. It is important to remember that this is only a random sample and a very tiny fraction of the available TEL tools throughout the continent. Similar to within the health sector, a focus on technical quick fixes instead of long term strategies can lead to fragmented silo systems as opposed to integrated standardized architectures (Sanner, 2021-b). A solution for this can be harmonizing philanthropic projects with existing governmental programs and this can be accomplished by the use of an EMIS. It collects data from different sources and can identify gaps and overlaps in current efforts which can be used to ensure that human, economic, and digital resources are being used efficiently. This also enables the creation of new indicators and cross-comparison of regions, districts, schools, grades, classes, students, teachers and so on. EBD can be gathered from population censuses, SISs, and NSTs and then be processed with EDM and LA techniques and ultimately be visualized as maps, charts, and tables in dashboards. These can then be used to monitor learning and performance as well as to create and evaluate educational policies (West, 2012). With all of this in mind, the Health Information Systems Program (HISP) has recently taken their 28 years of experience within the health sector and adapted their District Health Information Software (DHIS2) for education (HISP, n.d.-a; DHIS2, n.d.-a). They collaborate with a wide range of large internationally recognized humanitarian organizations as well as numerous ministries of health and education in almost half the world.

1.2 Goal

In light of all this, this thesis will attempt to answer the following five research questions.

- RQ1: What are the opportunities and challenges or digitizing the education sector in developing countries?
  - RQ1a: What are the existing strengths and weaknesses of the Gambian education system and how can technology be used to leverage it?
- RQ2: How can data from educational apps be used to inform school management and planning in a low resource context?
  - RQ2a: How can DHIS2 be used for educational data collection in The Gambia?
  - RQ2b: How can data from Leap Learning be used to inform educational policymaking by the Gambian Ministry of Education?

1.3 Approach

The epistemology of this thesis aligns with that of critical research, and the theoretical perspective is accordingly critical theory. Critical theory "combines the different, but interlinked purposes of theorisation and transformation of a status quo characterised by socially oppressive conditions" (Masiero, 2021-a, p. 26). Critical research "assume that social reality is historically constituted and that it is produced and reproduced by people" (Masiero, 2021-b, p. 10). As shown by section 1.1.4 and 1.1.5, being familiar with the history of Africa is of the essence in order to fully understand the current state of technology adoption within the continent. The qualitative methodology of case study is used on The Gambia with document analysis, conference workshops, and interviews as methods. Thematic analysis is used to process and make sense of the data.

In order to answer the research questions, this thesis will begin with a literature search in order to get an overview of RT, TEL, EDM and LA. The literature review occurred in three stages. It started with searches on Google Scholar with the terms “representation theory”, “technology-enhanced learning”, “big data in education”, “educational data mining”, and “learning analytics”. Based on the paper title, all relevant results on the first three search pages were included. After that, the articles were scanned for general and recurring themes and a subset of about 10-15 articles in each topic were read. Lastly, only about the 5-10 most representative, comprehensive, or rigorous articles in each subject made the final cut. Each topic has a section and multiple sub-sections. Some sub-sections consist of an in-depth exploration of a paper, others detail a specific aspect of the topic from multiple perspectives, whereas others still provide explanations of numerous similar concepts. Special care was taken to find at least one criticism and one application in developing countries which take up the second last and last sub-sections of the review respectively.

After that, a brief overview of the context will be provided. The first two sections offer a deep dive into the complex sociotechnical and technical aspects of the enormous ecosystem that is DHIS2. This includes a summary of its history, use cases, conceptual definitions and features. The third section will take a quick look at Leap Learning. The last two sections provide an outline of The Gambia including its demography, geography, economy, government and history, as well as the Gambian education system including its structure, terms and schedules, curriculum, assessment and management. These sections were written after a systematized investigation of numerous pages on various websites that were deemed credible after which the contents were rearranged and organized by topic.

Next, the methods used will be described in more detail followed shortly by the results of the document analysis, workshops, and interviews respectively. The former was performed on resources that were found on the ministry of education website in The Gambia. The second-mentioned was based on a conference of DHIS2 for education. The latter was held with various education directorates and Leap Learning. Thematic analysis was used on all of the forementioned. The document analysis yielded the themes “education access: management, schools, enrollment”, “equitable access: regional difference, gender differences, special needs children”, and “available resources: capacity, teachers, textbooks, facilities”. The themes that emerged in the workshops were “organizational aspects: strategy, systems, data” and “technical aspects: features, implementation”. The interviews resulted in the themes “learning apps: Leap Learning, Play to Learn”, “management: cooperation, training, challenges”, and “considerations: evaluation, localization”.

Finally, a discussion and conclusion on the topics from the introduction and literature review will be provided. The discussion looks specifically at Leap Learning data in DHIS2 for education in light of RT, TEL, EDM, and LA, whereas the conclusion focuses on its implementation within the Gambian education system.

Chapter 2

Literature Review

2.1 Representation Theory

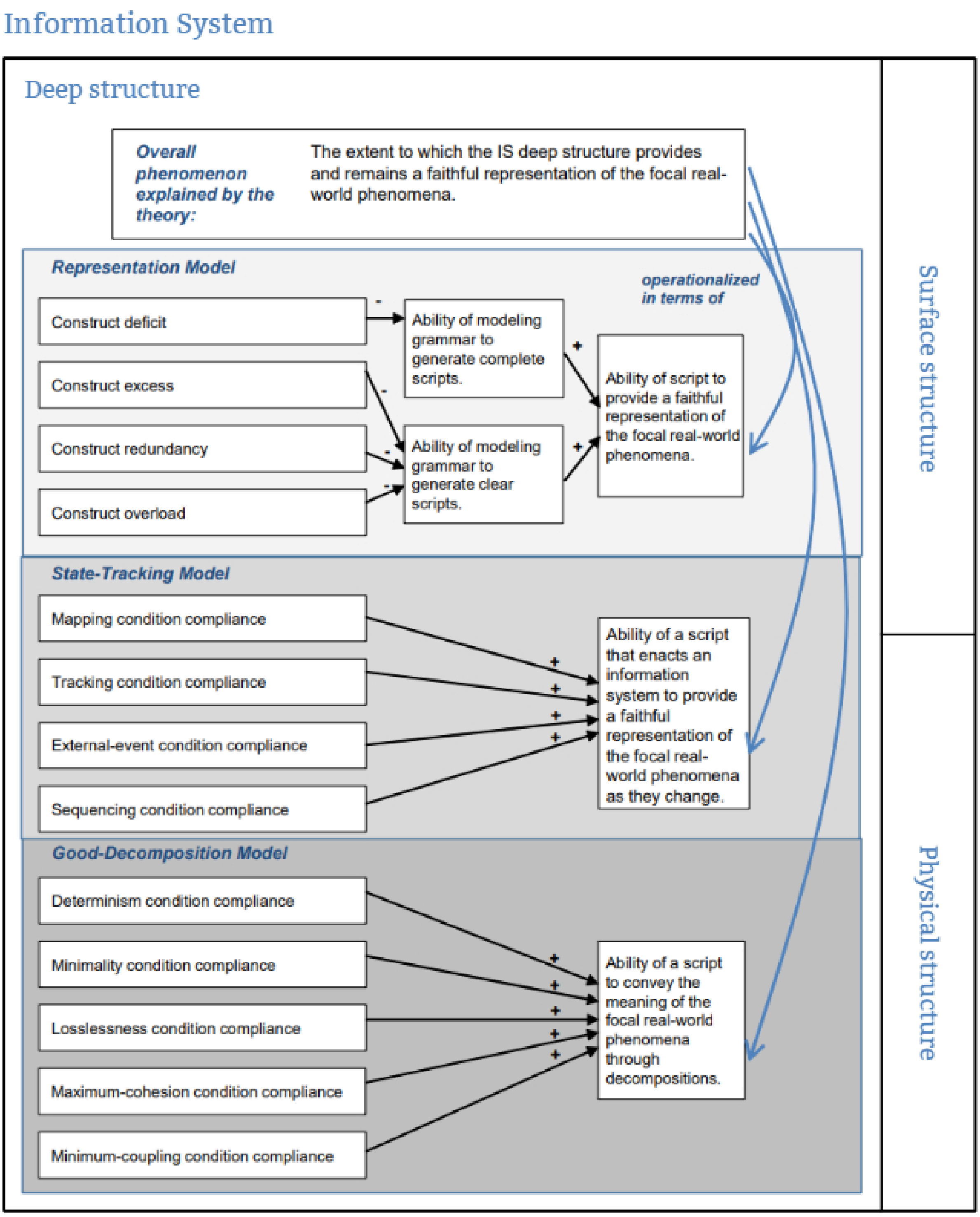
"The core purpose of all [information] systems is to help people understand the states of some real-world systems that are relevant to them" (Burton-Jones & Grange, 2013, p. 636). In order to do this, these systems have to be faithful representations of their physical counterparts. For example, an electronic health record system can be a representation of medical records, and a learning content management system can be a representation of school textbooks. Within IS research, Wand and Weber's RT focuses on this inherent characteristic of ISs (Fossum & Kempton, 2020). More specifically, "the presence or absence of certain properties in [IS] grammars or scripts enhances or undermines (…) a faithful representation" (Burton-Jones et al., 2017, p. 1309). Wand and Weber claimed that an internal view of the IS itself can be decoupled from an external view of its development, implementation, and management (Burton-Jones & Grange, 2013). However, this assumption has later been criticized and disputed, as it is argued that an IS cannot be viewed independently from its organizational and social context (Fossum & Kempton, 2020). This section will provide a brief introduction to RT and how it can be used in the creation and evaluation of information systems.

2.1.1 Overview of RT Structures and Models

RT holds that ISs consists of a surface structure, a deep structure, and a physical structure (Fossum & Kempton, 2020). The former refers to the user interfaces of the system, the second-mentioned refers to the representations of the system, and the latter refers to the hardware of the system. The deep structure can be further divided into three parts. The representational model "describes how well the structure conforms to [the] given ontology". The state-tracking model refers to how well it "track[s] events and state changes in the real-world phenomena". The good-decomposition model "specifies the conditions for how the representation should be decomposed" (p. 2-3). There are four ways in which the representational model can be weakened (Burton-Jones et al., 2017). Firstly, when an

Figure 2.1

Overview of RT Structures and Models



Note. Adapted from "Assessing Representation Theory with a Framework for Pursuing Success and Failure" by A. Burton-Jones, Recker, J., Indulska, M., Green, P., and Weber, P., 2017, *MIS Quarterly*, *41*(4), p. 1307-1334 (https://doi.org/10.25300/MISQ/2017/41.4.13).

ontological construct is not covered by a grammatical construct (construct deficit). Secondly, when a grammatical construct does not cover an ontological construct (construct excess). Thirdly, when multiple grammatical constructs map to one ontological construct (construct redundancy). Finally, when one grammatical construct maps to multiple ontological constructs (construct overload). In addition, there are four ways in which the state-tracking model can be strengthened. Firstly, "a state of the [IS] must map to only one state of its focal real-world phenomena" (mapping condition; p. 1310). Secondly, "an [IS] must faithfully track state changes that occur to its focal real-world phenomena as a result of (…) internal event[s]" (tracking condition). Thirdly, "an [IS] must provide a faithful representation of any change in the state of its focal real-world phenomena that arises due to (…) external event[s]" (external-event condition). Finally, "an [IS] must faithfully record the sequence in which external events occur in its focal real-world phenomena" (sequencing condition; p. 1312). Furthermore, there are five ways in which the good-decomposition model can be strengthened. Firstly, all internal events have to be predictable and only external events can be unpredictable (determinism condition). Secondly, there should be no redundant state variables that are never used (minimality condition). Thirdly, all hereditary (i.e., individual) and emergent (i.e., collective) properties of the focal real-world phenomena must be preserved in the decomposition (losslessness condition). Fourthly, it should not be possible to break a subsystem into two further independent subsystems (maximum-cohesion condition). Finally, as many events as possible should be internal as opposed to external (minimum-coupling condition). A hierarchical diagram of RT structures and models is shown in Figure 2.1.

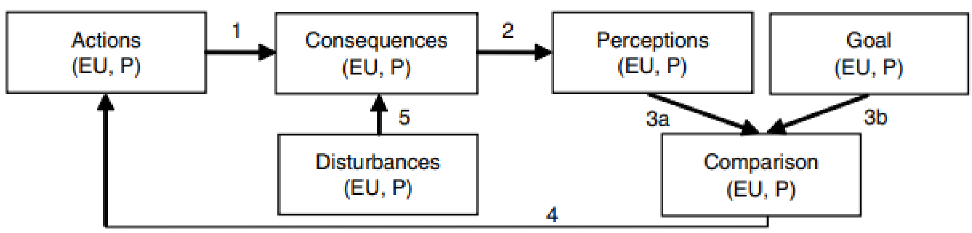
2.1.2 Expanding RT with Interpretive Structure

Fossum and Kempton (2020) suggest an expansion of RT after a short literature review of the topic and an observational study of an emergency care unit (ECU) in a Norwegian hospital. Various researchers have proposed that ISs – and by extension, their representations – have multiple dimensions. Let us say that a nurse reads the blood pressure of a patient as part of a routine physical examination. The monitoring equipment, the fraction format, the systolic and diastolic pressure, and the health assessment would comprise the physical, syntactic, semantic, and pragmatic dimensions of the representation respectively. Fossum and Kempton suggest the addition of three further dimensions based on previous work in the field. The *correctness* dimension reflects the causality between blood being pumped and the readings of the test. The *subjectivity* dimension reflects the evaluation that the blood pressure is higher than average. The *situation* dimension reflects the implication of a diet high in cholesterol. They collectively refer to these dimensions as the "interpretive structure of an IS and its representations" (p. 5). In other words, how a representation is interpreted is equally important as how the representation is formed.

In order to test their hypothesis, Fossum and Kempton followed 15 complete patient trajectories and interviewed 25 clinicians at the ECU. They found that the same representations of a patient's health condition could be interpreted and used in different ways depending on their pragmatic, situational, and subjective dimensions. The triage nurse who is responsible for prioritizing patients in the ECU would be more selective in which data to document and communicate. The polyclinic nurse who is responsible for continuous care in the polyclinic would be more inclusive in which data to document and communicate. Further, the correctness (both actual and perceived) dimension would determine if additional representations would be used, as in the case of the triage nurse using multiple thermometers to take the patient's temperature. Fossum and Kempton argue that good representations, decompositions, and state trackers emerge "by, in, and for the practice". In other words, the deep structure of ISs "will never be fully complete nor (…) fully designed up front". Because of this, "the IS needs (…) to let users compose and re-compose representations" (p. 11). In other words, there is also a need for the inclusion of a temporal dimension as representations and their interpretations can change over time.

Figure 2.2

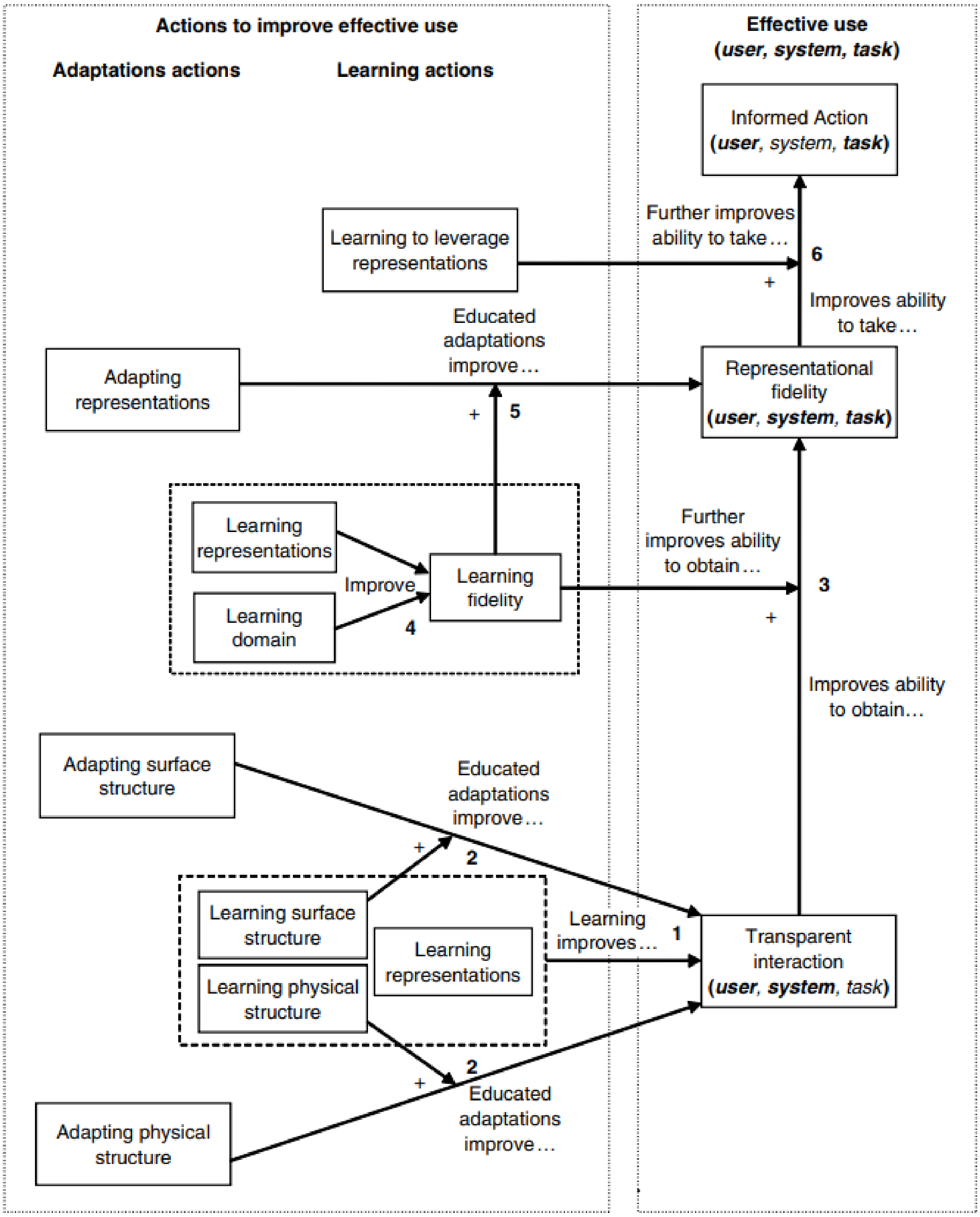
Conceptual Framework of Effective Use of ISs



Note. From "From Use to Effective Use: A Representation Theory Perspective" by A. Burton-Jones and C. Grange, 2013, *Information Systems Research*, *24*(3), p. 632-658 (https://doi.org/10.1287/isre.1120.0444).

Figure 2.3

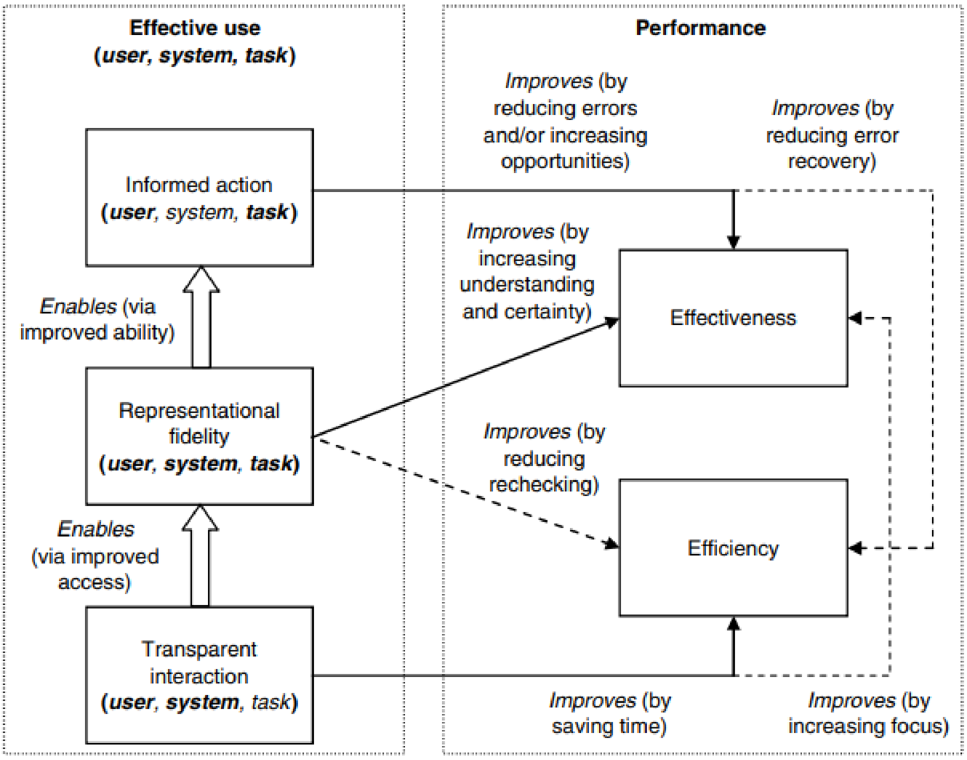
The Nature of Effective Use of ISs



Note. From "From Use to Effective Use: A Representation Theory Perspective" by A. Burton-Jones and C. Grange, 2013, *Information Systems Research*, *24*(3), p. 632-658 (https://doi.org/10.1287/isre.1120.0444).

Figure 2.4

The Drivers of Effective Use of ISs



Note. From "From Use to Effective Use: A Representation Theory Perspective" by A. Burton-Jones and C. Grange, 2013, *Information Systems Research*, *24*(3), p. 632-658 (https://doi.org/10.1287/isre.1120.0444).

2.1.1 Using RT to Define Effective Use of IS

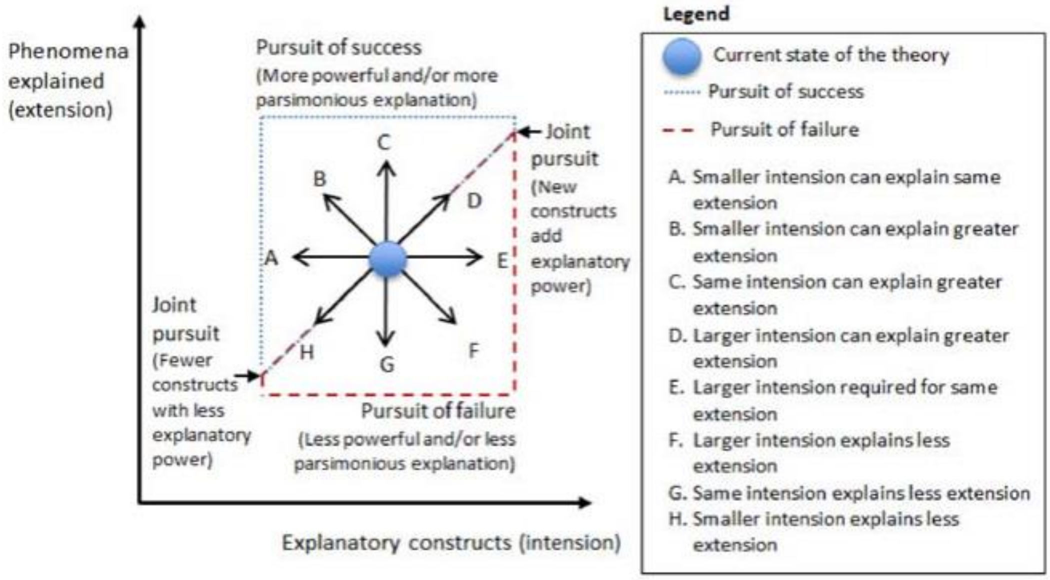
Burton-Jones and Grange (2013) looks at how RT can inform and model the effective use of ISs. Based on a thorough literature review, they propose a framework and two models of the effective use of ISs. They differentiate between effective use (EU) as the means and performance (P) as the end of achieving a goal. As shown in Figure 2.2, users can take actions based on both EU and P which have consequences for both EU and P with EU improving P. They also distinguish between performance in terms of effectiveness as in *whether* a goal is achieved or not, and efficiency as in the *rate* at which the goal is achieved. Moreover, as presented in Figure 2.3, there are three dimensions of effective use that approximately reflect the surface structure, the deep structure, and the physical structure of RT. The first is *transparent interaction*, which is "the extent to which a user is accessing the system's representations unimpeded by its surface and physical structures". It primarily improves efficiency by saving time, and secondarily improves effectiveness by increasing focus. The second is *representation fidelity*, which refers to “the extent to which a user is obtaining representations from the system that faithfully reflect the domain being represented”. It primarily promotes effectiveness by increasing certainty, and secondarily promotes efficiency by reducing double checking. The third dimension is *informed actions*, “the extent to which a user acts upon the faithful representations he or she obtains from the system to improve his or her state” (p. 642). It primarily supports effectiveness by increasing accuracy, and secondarily supports efficiency by reducing error recovery. Transparent interaction enables representation fidelity which in turn enables informed actions. Burton-Jones and Grange also contrast adaptation actions with learning actions, as displayed in Figure 2.4. The former refers to the user’s adjustments of the surface structure, deep structure, or physical structure of the IS. The latter refers to the user’s grasp of the same structures in addition to the domain, representations, and fidelity of the IS. Learning the surface and physical structures of the IS improve adaptation of them, which together with learning its deep structure improves transparent interaction. Learning the representations and domain results in learning the fidelity of the IS, which enables adaptation of the representations, and in turn promotes representational fidelity. Finally, learning to leverage the representations of the IS supports informed action.

2.1.2 Assessing RT Successes and Failures

In a comprehensive paper, Burton-Jones and colleagues (2017) provide an outline of RT constructs, evaluate it by looking at its merits or lack thereof, and propose directions for future research. They use Gray and Cooper’s framework to pursue the failures of the theory and Kaplan’s framework to pursue the successes of it. The first-mentioned involves identifying questionable assumptions, delineating the theory’s boundaries, seeking out competing theories, and trying to explain contradictory findings. The last-mentioned consists of improving the theory’s intension by refining its constructs (reduction), by improving its extension by explaining new phenomena (addition), or by improving explanations of existing phenomena (precision), or all three. Burton-Jones and colleagues develop their own framework by elegantly combining these into a coordinate system, as evident in Figure 2.5. The diagonal line H-D depicts where the two different frameworks meet. A move towards H

Figure 2.5

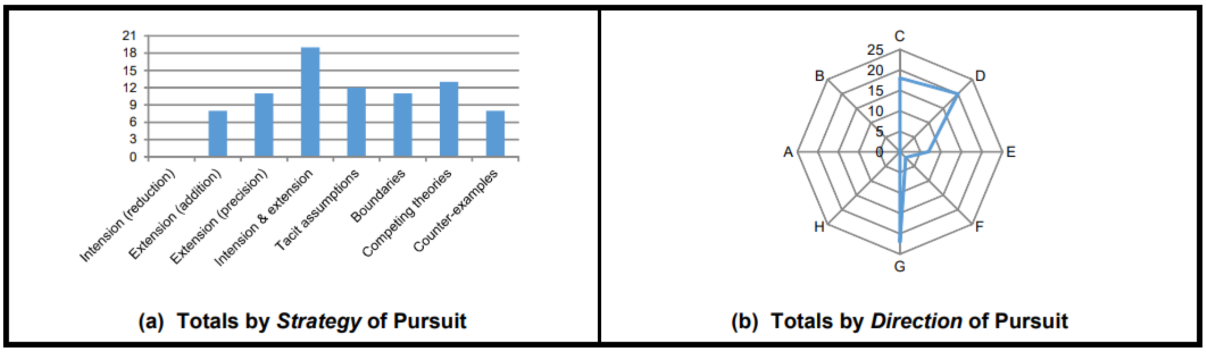
Framework for Assessing RT



Note. Adapted from "Assessing Representation Theory with a Framework for Pursuing Success and Failure" by A. Burton-Jones, Recker, J., Indulska, M., Green, P., and Weber, P., 2017, *MIS Quarterly*, *41*(4), p. 1307-1334 (https://doi.org/10.25300/MISQ/2017/41.4.13).

Figure 2.6

Results of RT Assessment



Note. Adapted from "Assessing Representation Theory with a Framework for Pursuing Success and Failure" by A. Burton-Jones, Recker, J., Indulska, M., Green, P., and Weber, P., 2017, *MIS Quarterly*, *41*(4), p. 1307-1334 (https://doi.org/10.25300/MISQ/2017/41.4.13).

follows Gray and Cooper's ideals of having a limited theory explaining few phenomena, whereas a move towards D follows Kaplan's ideals of having an expanded theory explaining many phenomena. Additionally, "a move to B reflects the ideal of simplifying a theory and gaining explanatory power", whereas "a move to F (…) show that a theory cannot explain what it purports to explain even with additional constructs" (p. 1315). Finally, any moves towards A, C, G and E fall in-between these.

In a literature review, Burton-Jones and colleagues used their framework on 69 articles that were found to be relevant after multiple rounds of independent coding the strategy and direction of pursuit. Whereas the simultaneous use of intension & extension, and the use of addition or precision were all popular strategies, no articles used reduction. As displayed in Figure 2.6, use of the other strategies was more evenly distributed. Correspondingly, the pursuit of most papers were in direction G, D, and C respectively which all emphasize changes in extension of the theory. Some articles that pursued the failures of RT did not offer an alternative explanations, and most of these did not provide empirical evidence backing their claims. The majority of the papers examine the representational model, some look at RT in general, a few analyze the good-decomposition model, but none study the state-tracking model, which could be an area for future research. Other immediate topics of further investigation based on gaps in previous research would first and foremost include reduction, followed by addition, boundaries, and counterexamples. Several scholars have argued that the criteria of construct redundancy is superfluous, and this could indicate a possibility for reduction. An opportunity for addition could be auditors using state-tracking conditions of process mining logs in financial information systems to assess risk of misstatements of accounts. In the case of uncovering boundaries, the theory could be of limited use if different stakeholder groups are unable to reach a consensus on the meaning of their focal real-world phenomena. Finally, in terms of counterexamples, multiple researchers have found that construct overload, excess, and deficit do *not* have detrimental effects on the representational model which should be explored in future studies.

2.2 Technology-Enhanced Learning

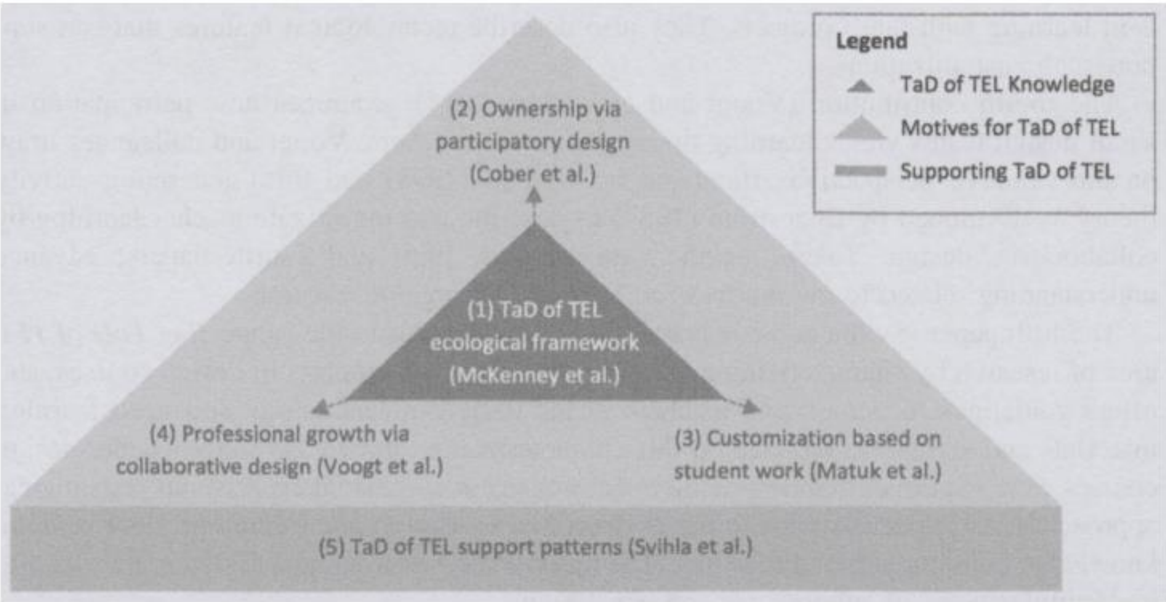
Technology has been used in education for a long time – from rudimentary abacuses and calculators with graphing features to computers capable of sophisticated statistical analyses. Various terms have been adopted to describe different aspects of the relationship between education and technology (Bayne, 2015). These includes “educational technology”, “computer-based learning”, “online education”, “learning technology”, “e-learning”, “instructional technology”. Recently, the term “technology-enhanced learning” has become increasingly popular, which can simply be defined as “enhancing learning and teaching through the use of technology” (Higher Education Funding Council for England, 2009 as cited in Kirkwood & Price, 2014, p. 8). TEL is a somewhat vague term that has a wide range of diverse applications. Since listing all of these would be outside the scope of this thesis, the review will only focus on the strengths and weaknesses of TEL in and of itself.

2.2.1 Design-Based Research in TEL

The successful implementation of any TEL tool begins with design. There are several proponents for the use of design-based research and participatory action research on learning technologies (Mor & Winters, 2007; Wang & Hannafin, 2005). It is argued that TEL environments (TELE) are often based on “contradictory theoretical (…) foundations” and that there is a “lack of (…) frameworks applicable to practice” (Wang & Hannafin, 2005, p. 12-13). Design approaches can resolve both these problems by closing the gap between theory and practice and by encouraging continuous synergy between them through iterative processes. It shows a lot of promise as it contributes to knowledge building, while at the same time, it moves the research process out of artificial laboratory settings and into organic classroom environments. Moreover, design experiments employ a wide range of both qualitative and quantitative methods such as surveys, interviews, observations, case studies, document analysis, inquiry methods, expert reviews, and more. Thus, it can be a source of both general meta-design knowledge and specific context-based data. In addition, design research takes advantage of the fact that “many experienced practitioners in education have tried and tested methods for solving recurring problems or addressing common needs” (Mor & Winters, 2007). The design researcher can mediate between different stakeholder groups like administrators who want measurable results and teachers who are more concerned with didactic processes. Furthermore, design-based research can offer valuable inputs “on the engineering aspects of designing, developing, and evaluating good technology for (…) instruction” (p. 67). Designs can be tested almost immediately using formative and summative evaluations, which can also be used to assess the effectiveness of reform policies in everyday settings. However, design-based research is a resource intensive activity that is often cut short due to time constraints. Large amounts of data are generated, although most of it is deleted and never used. It can be difficult to convince policymakers to invest in it when it may not be compatible with their ideas of scientific research, it may not lead to generalizable results, or it might be viewed upon as a distraction or an intrusion. Therefore, it is important that design researchers demonstrate rigor and thoroughness, determine applicability to other contexts, and minimize any Hawthorne effects to the best of their abilities.

Figure 2.7

Conceptual Framework of TaD of TEL



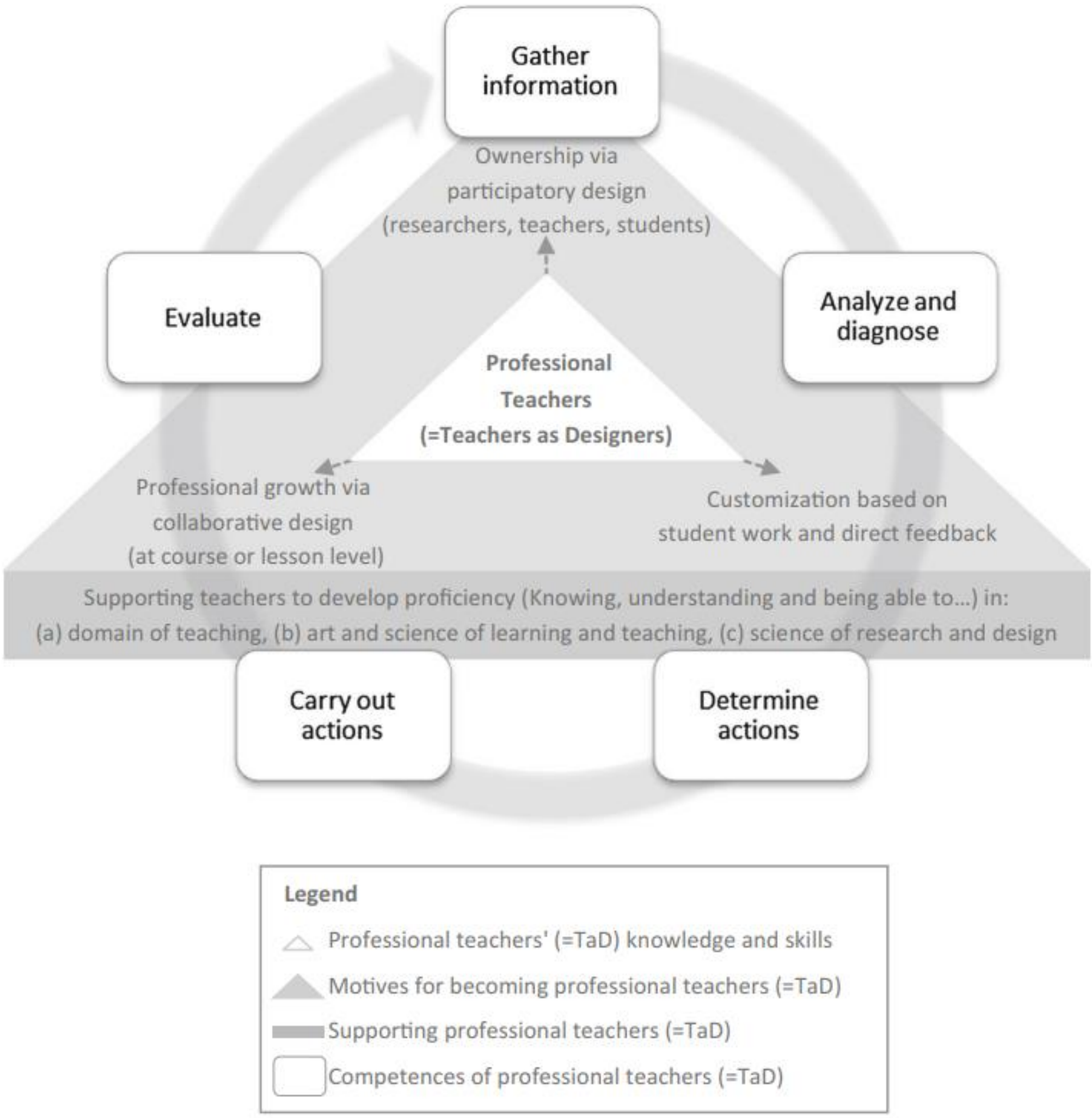
Note. From "Teachers as Designers of Technology Enhanced Learning" by Y. Kali, S. McKenney, and O. Sagy, 2015, *Instructional Science*, *43*(2), p. 173-179 (https://doi.org/10.1007/s11251-014-9343-4).

2.2.2 Teachers as Designers of TEL

In a special issue of the journal Instructional Science, Kali et al. (2015) plead for the use of teachers as designers (TaD) of TEL. They propose a conceptual framework for TaD of TEL, as seen in Figure 2.7. This is based on insights from the *International Conference of the Learning Sciences in 2012* working groups who have synthesized previous literature in this field over the course of two years. Research in this field mainly focus on TaD of TEL knowledge, motives and support. Different types of knowledge are needed depending on the nature of the design work, whether it is the re-design of existing material and activities or co-design of completely new ones. Teachers may have various motives for design of TEL, whether it is ownership via participatory design, customizations to optimize student learning,

Figure 2.8

Revised Conceptual Framework of TaD of TEL



Note. From "Do We Need Teachers as Designers of Technology Enhanced Learning?" by P. A. Kirschner, 2015, *Instructional Science*, *43*(2), p. 309-322 (https://doi.org/ 10.1007/s11251-015-9346-9).

or learning by collaborative design. Various supports such as work structuring, common goals, and process guidance can be very helpful for the design of TEL by teachers. Within the same special issue, Kirschner (2015) challenges and adds on to this conceptual framework. He argues that knowledge, motivation, and support alone are not enough to ensure that learning takes place, but that competence or the ability to take action based on these elements is just as important. He argues that all professionals must possess a set of competencies that unlike knowledge and motives do not change over time or depending on the field. As shown in Figure 2.8, they are (a) gathering information, (b) analyzing and diagnosing, (c) determining actions, (d) carrying out actions, and (e) evaluating actions. Further, he contends that TEL is “an artefact of the times” (p. 312) and only one mean amongst many. Therefore, it should not be a goal in and of itself, but only in so far as it supports learning. It is only a small piece of the tools and spaces part of the education ecosystem, with teachers and students as the two other parts. Especially the last-mentioned tends to be overlooked in TaD for TEL, which can be counteracted by considering students as the end user that the design is for.

2.2.3 Criticism of TEL

Despite the apparent promise of TEL, there has been a lot of criticism of it as well. As Kirkwood & Price (2014) points out, “it is not evident that a shared understanding has been developed (…) of what constitutes an *enhancement* of the student learning experience” (p. 6). It seems that most researchers subconsciously equate ‘enhancement’ with operational improvement such as making resources more accessible or quantitative changes in learning as opposed to qualitative ones. In a critical review of TEL intervention studies, Kirkwood & Price call attention to some of their methodological weaknesses. Experiments comparing TEL test groups with non-TEL control groups cannot infer causality due to the possibility of confounding variables and limited independent variable manipulation. In addition, although not necessarily addressed directly in the article, self-report surveys may suffer from courtesy, demand, or social desirability response biases. Last but not least, many research papers focus on the application of a specific technology, without providing details of the context in which it is used or contemplating how these results can be generalized to other contexts. More replication studies and data triangulations are needed in order to fully capture complex multifaceted processes like the introduction of technology into the education sector.

In a response to this article, Bayne (2015) contends that TEL is not in need for a clearer definition or methodological improvements, but that the term should be abandoned altogether. It is argued that it reflects oversimplified assumptions about technology that limits our capacity to be critical towards its use. First, ‘technology’ is often an expression of essentialism and instrumentalism. The first-mentioned claims that certain technologies have an immanent pedagogical value in and of itself. The last-mentioned assumes that technology is simply a tool that can be used to reach our learning goals. Such approaches tend to isolate the technical from the social and fails to see how technology can shape us instead. Second, ‘enhancement’ is often used to emphasize transhumanistic views over posthumanistic ones. The former builds upon the idea that humans are rational, autonomous, and free. However, it also inadvertently promotes colonialist, patriarchal, and capitalist power structures. The latter criticizes and challenges this idea and maintains that humans must be viewed within the larger social, political, and hierarchical context. Third, ‘learning’ often turns out to actually refer to ‘education’ and ‘teaching’. The problem with such discourse is that it hides administrative goals of assessments and feedback such as more standardized national tests under the positive term ‘learning’ which prevents us from asking critical questions about them.

Dror (2008) takes on a very different approach when it comes to criticizing TEL. He believes that the use of technology for learning should be rooted in research on human cognition. Too often, TEL simply presents educational material to students for passive consumption which results in minimal learning. He maintains that in order to encourage active participation, it is important for TEL to promote control, challenge, and commitment. Pupils should be able to decide which modules to learn or at least the order in which they learn. Learning should be made to be fun and motivating by being presented as digital puzzles and games. By using the previous two strategies, the student will hopefully also become more dedicated to education. However, it is also important to avoid cognitive overload through the use of quantitative measures such as reducing the amount of information or qualitative ones like intentional use of colors, font sizes, animations, and so on.

2.2.4 TEL in Developing Countries

The issues surrounding TEL are magnified in low-resource contexts where gaps between rich and poor, urban and rural, educated and uneducated are bigger. In a comprehensive review of TEL in low- and middle income countries, Gulati (2008) highlights the complexities of such issues:

> Technologies, on one hand (…) are seen as an answer to the limitations and rigidities of conventional education because they reach a larger number of learners, while on the other hand, where these technologies are being used (…) the digital divide between the privileged and deprived groups continues to widen the educational gap. (p. 1-2).

The first example she uses is a prominent demonstration of this. In the 1980s, the National commission of Nomadic Education in Nigeria broadcasted educational radio programs for the rural nomadic Fulbe people. Topics that were taught included the prevention of animal disease, the production of goods from crops and the cooking of nutritious foods. However, they failed to reach the women of the community, because their husbands controlled access to the radios. Although the project was aimed at reducing inequality, due to a lack of cultural sensitivity, it had actually ended up increasing it instead. A similar occurrence happened in 1963, when the World Bank started to invest in secondary, vocational, and technical education. However, the project was mostly economically motivated, and the funds ended up being spent on the wealthy educated elite in urban areas. In addition, the 1978 Distance Teaching Experiment at the University of West Indies highlights yet other conflicts of interest:

> Distance education (…) was still being sidelined by face-to-face teaching faculty who viewed it as an add-on, and voicing fears that they were 'giving away' their intellectual property. They did not view their distance teaching experience as a positive influence on their careers, nor did they identify with distance education's goal of widening access. (p. 5)

However, not all open and distance learning endeavors fail. Since 1968, Mexico has been successful in teaching lower secondary school children in far-away communities subjects like Spanish, mathematics, and chemistry using televised lessons known as Telescundaria. In 2007, the University of Terbuka in Indonesia had also experienced significant success with distance education:

> The widespread use of print, audiocassettes, face-to-face tutorials, and intermittent, government-controlled radio and television broadcasts supports over 600,000 urban, sub-urban, and rural learners attached to the Bangladesh Open University. These traditional methods may be more reliable, but they are often one-way and lack teacher-student and student-student interactivity. (p. 10)

When it comes to ICT, the challenges are perhaps even more daunting. In the 1990s, recognizing the importance of technology in a modern global economy, the Botswanan government attempted to supply computers to primary and secondary schools. However, there is a shortage of teachers who are able to teach IT skills, a lack of infrastructure needed to participate in rural areas, and a strong reliance on western funding and learning materials. These are often not culturally sensitive and have not been adapted to suit local needs. It is also difficult to justify spending on computer education when there are more pressing needs in terms of fundamental living conditions. In 1999, UNESCO reported that the proportion of the population that has access to telecommunications was around ten times higher in developed countries (42%) than in developing countries (4.5%). In rural areas, telephone connections are often shared amongst the entire communities as opposed to single households and internet costs are often extremely high which makes it an unaffordable luxury. However, Gulati concludes by remarking that this does not mean that we should stop developing IT infrastructure or providing open and distance learning in low-resource contexts. It simply means that donors, governments, and other stakeholders should be more conscientious and intentional when it comes to introducing technology into the educational sector in order to ensure equitable access for all.

2.3 Educational Data Mining

Multiple low- and middle income countries have acknowledged the importance of providing students with technology in order to increase their IT skills so that they can learn on their own (Gulati, 2008). Currently, educational data is gathered through large-scale once-a-year paper-based school censuses and standardized national tests (Fischer et al., 2020; MoBSE, n.d.). These typically are about attendance or grades and course-grained enough to be stored in a traditional relational database management system that uses simple SQL-based queries (Romero & Ventura, 2013; Sin & Muthu, 2015). However, the digitization of education can completely transform these processes:

> [Student] learning (…) activities [will] leave abundant digital traces, including, but not limited to, mouse clicks and keystrokes in front of computers, tabs and swipes on smartphones and tablets (…). While big data movement is still in infancy, the promising prospects of [it] in education are already present. (Wang, 2016, p. 382)

The next section will look at the definition of educational big data and educational data mining, in addition to how they are similar to or differ from related concepts.

2.3.1 Definitions and Related Concepts

Big data is characterized by its volume, velocity, and variety:

> The volume of the data is so massive that it might exceed the computing and processing power of one computer, hence, a cluster of computers are often needed for data processing. The velocity of data generation is so fast that 90% of the world's data today were generated during the past 2 years. The variety of data is so diverse that the data sources range from log files of using digital devices, web browsing data, social media data, geo-located images, to audio data. (Wang, 2016, p. 381).

Value and veracity have also later been added to its defining features (Joubert et al., 2021). The former refers to the extent to which economic benefits can be extracted from the data which is a key motivation for using it. The latter refers to the extent to which the data sources can be trusted to provide quality data which is essential for the data analysis to be meaningful. EBD is the result of gathering big data within the education sector and can be used to inform educational policymaking or improving student learning outcomes.

Data mining (DM) can be described as "the process of extracting previously unknown, comprehensible, and actionable information from large databases and using it to make crucial (…) decisions" (Zekulin, n.d., as cited in Friedman, 1998, p. 1). DM is often confused with statistics, machine learning, and artificial intelligence. Whereas statistics use probability theory, real analysis, measure theory, asymptotic analysis, etc., data mining uses decision trees, neural networks, rule induction, nearest neighbors and so on (Friedman, 1998). While machine learning assumes "that there is an underlying interesting concept or mechanism that produces the data", data mining "does not necessarily assume that there would be any sensible structure behind the data" (Mannila, 1996, p. 3). Finally, artificial intelligence is used to describe "any device that perceives its environments and takes actions that maximize its chance of success at some goal" (Ongsulee, 2017, p. 1). These capabilities are not usually possessed by traditional DM techniques. EDM "can be defined as the application of data mining techniques to [big data] that come from educational environments to address important educational questions" (Romero & Ventura, 2013, p. 12).

The term DM is often used interchangeably with Knowledge Discovery in Databases (KDD), but the latter is actually a process in which the former is only one step of many (Romero & Ventura, 2013). KDD starts with an (educational in this case) environment from which raw data is extracted and then preprocessed until it becomes clean data. This involves deciding on an appropriate data structure or format, converting the heterogeneous data into homogeneous data, and anonymizing the data in order to protect the privacy of the students. After that, data mining techniques are applied on the clean data in order to extract models or patterns from it which then are used to interpret or evaluate the phenomena that the data represents. These interpretations and evaluations are used to made changes and adjustments to the environment and the results of this are analyzed by extracting raw data again and repeating the entire process.

2.3.2 Introduction to EDM and its Methods

EDM can be used to analyze hierarchical, contextual, fine-grained, continuous data generated by any type of IS supporting learning or education (Romero & Ventura, 2013). This could be student, administrative, or demographic data gathered by schools, colleges, or universities from MOOCs, ITSs or LMSs. It can be used by *learners* to improve learning performance, to get feedback and recommendations, and to respond to their needs. It can be used by *teachers* to improve teaching performance, to understand student learning processes, and to reflect on their own teaching methods. It can be used by *administrators* to evaluate the best way to organize institutional resources. EDM is an interdisciplinary field commonly associated with visual data analytics, pedagogics, psychology, psychometrics, and so on. There is considerable research interest in this area, as evident by the International Conference on EDM, the Journal of EDM, the International Working Group on EDM, and multiple Workshops on EDM in various conferences. Scholars are typically concerned with issues like framework development, educational software design, personalized learning, and student interventions amongst others.

Some of the common methods used in EDM include (a) prediction, (b) clustering, (c) outlier detection, (d) relationship mining, (e) discovery with models, and (f) knowledge tracing. Prediction specifically involves projections of future values of a dependent variable like student performance or behaviors based on one or more independent variables. This can be done using *classification* if the predicted variable is a categorical value, *regression* if it is a continuous value, or *density estimation* if it is a probability density function. The goal of clustering is to find groups of entities that are similar, which is determined by a distance measure, and it can be used to find similar course materials for linking or similar students for group work. Outlier detection consists of identifying data points that are significantly different than the rest of the data which can be used to find students with learning difficulties or give an early warning to failing students. Relationship mining aims to identify and encode links between variables and can be used to connect learner behavior patterns with frequently occurring mistakes. This can be done using *association rule mining* for finding any relationships, *sequential pattern mining* for finding temporal links, *correlation mining* for linear connections between variables, and *causal data mining* for cause-effect relationships. Discovery with models involves using a previously validated model to analyze a phenomenon, for instance contextual influences or research questions. The purpose of knowledge tracing is estimating student mastery of a skill by using a cognitive problem-solving model on logs of the student's correct and incorrect answers.

2.3.3 The Challenges of Big Data and EDM

In a comprehensive review, Fischer et al. (2020) also look at some of the challenges surrounding EDM. Only empirical studies of real-world digitally recorded or archived data collected at scale from educational contexts were considered. In other words, simulation, replication, meta-analytical, methodological, conceptual, qualitative, descriptive statistical, summative data, and sensor data studies were all excluded. They differentiate between microlevel (e.g., clickstream) data, mesolevel (e.g., textual) data, and macrolevel (e.g., institutional) data, although they are not necessarily mutually exclusive. Microdata can be used to identify knowledge components, metacognitive and self-regulating skills, affective states, in addition to evaluating student knowledge, using data for action, and personalizing learning. Though, "since microlevel big data are easy to collect, many research projects focus solely on them, potentially neglecting important related phenomena that are more coarse-grained" (p. 139). Mesolevel data can be used to support and evaluate cognitive functioning and social processes, as well as detecting behavioral engagement, and examining affective constructs. Although, "the applicability of various tools has not been tested extensively in all educational settings (…) [and] researchers cannot ignore contextual factors such as the stimuli to which students are responding" (p. 143). Macrolevel data can be used for early-warning systems, course guidance systems, and administration-facing data analytics. However, curricular requirements cannot easily be generalized across schools, data about employment outcomes should be leveraged but will raise privacy concerns and sharing results with students may have unintentional adverse effects on their performances.

Furthermore, there are some challenges when it comes to (a) accessing, (b) analyzing, and (c) using big data in general. First, many educational software companies are not interested in giving public access to their data. Amongst the few who do, they might not have high quality data suited for research or they might require programming skills in order to extract them. In addition, finding the right balance between sharing data and protecting the privacy of the students can be challenging. Even if the data goes through deidentification, it is difficult to prevent reidentification. This is especially true for mesolevel and macrolevel data. A possible solution for this could be institutional review board protocols that only allow sharing of data for research purposes. Another approach would be to encourage innovative open science practices in more traditional academic disciplines like education. Second, many academics have little experience with data mining techniques that are required to analyze big data. Confounding variables may produce high error rates and noise that is difficult to filter out. On the contrary, some important influences on student learning may not be captured. When policymakers take decisions based on these data, it can potentially harm the students instead of helping them. One way to resolve this would be to introduce interdisciplinary doctoral and postdoctoral programs involving both computer science and education or other fields. Another potential fix for this is combining and comparing macrolevel data with microlevel data. In fact, several studies have shown that "online engagement variables had significantly more predictive power than academic performance or demographic variables across a variety of classification models" (p. 149). Third, scholars tend to focus more on explanation than prediction, which results in intricate theories that are difficult to prove or disprove. This is the case in both psychology and education research. For instance, studies on literacy have long revolved around the debate on code-focused versus meaning-focused instruction and child managed versus teacher managed instruction. However, it was not until recently that data mining techniques were used to develop concrete accurate predictive models about what would work best for children at different levels. Finally, it is important to ensure that data used in predictions is not biased and just a continuation or confirmation of preexisting prejudice and discrimination.

2.3.4 Big Data Readiness in Developing Countries

Seeing a gap in the literature, Joubert et al. (2021) decided to look at big data in African governments as opposed to the typical Western organizations that dominate the field. Based on a design research approach, they developed and applied the Big Data Readiness Index

Figure 2.9

BDRI Performance of All African Countries

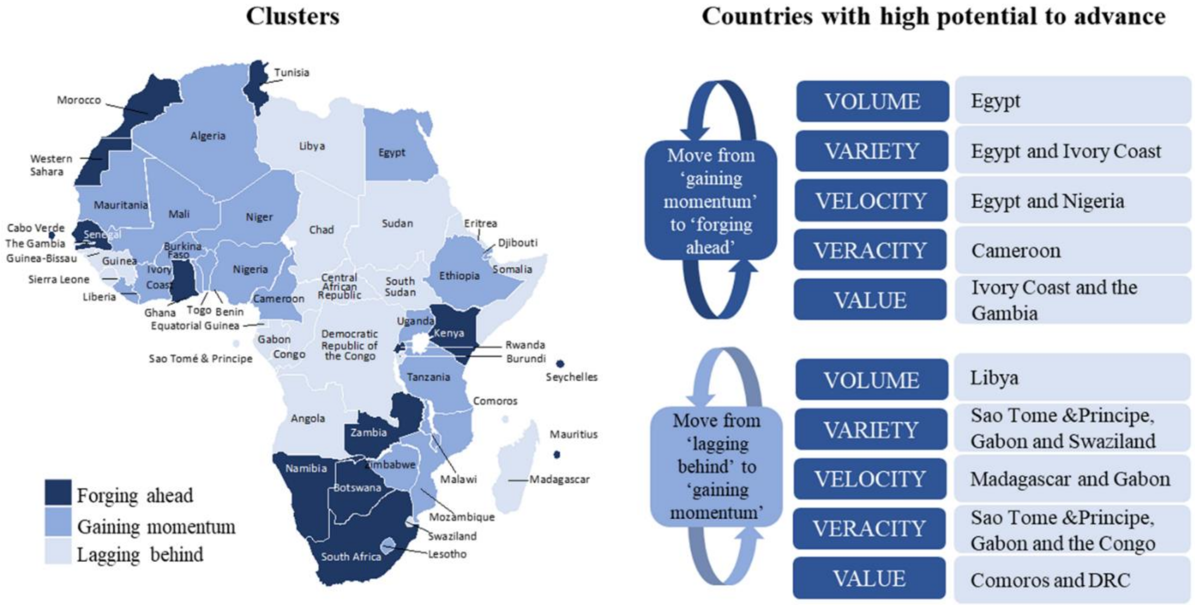
Note. From "Measuring the Big Data Readiness of Developing Countries – Index Development and its Application to Africa" by A. Joubert, M. Murawski, and M. Bick, 2021, *Information Systems Frontiers* (https://doi.org/10.1007/s10796-021-10109-9).

(BDRI) to assess whether developing countries are ready to take advantage of this technology. Following the definition of big data, the BDRI has five components – volume, variety, velocity, value, and veracity. In order to apply the index, they performed a comprehensive review of similar existing indices and their data sources, although these did not focus on developing countries. Because of this, they also did an exhaustive search for online open data sources on African countries and a rigorous data cleaning process dealing with missing, outdated, and redundant data. For each of the five V's, they found three drivers that were rescaled to a value between 0 and 1 where higher values means better achievement and lower values indicate worse outcomes. When applied to the countries, the BDRI shows that the highest scoring countries are Mauritius, South Africa, Seychelles, Rwanda, Kenya, Namibia, Morocco, Tunisia, Cabo Verde, and Ghana. South Africa is the highest performing region (amongst the top 10), followed by Eastern Africa (amongst the top 20) and then West Africa (amongst the top 30). Coastal countries and islands generally perform very well, and the former tends to have a positive influence on their neighbors. All countries tend to score high on velocity and value, low on variety and veracity, but vary widely when it comes to volume.

An overview of how the top ten performing countries scored in each component is shown in Figure 2.9.

Figure 2.10

BDRI Performance of Top Ten Countries



Note. From "Measuring the Big Data Readiness of Developing Countries – Index Development and its Application to Africa" by A. Joubert, M. Murawski, and M. Bick, 2021, *Information Systems Frontiers* (https://doi.org/10.1007/s10796-021-10109-9).

Both Mauritius and Seychelles score highly on all areas with opportunity for improvements in variety due to low technology adoption and export. The former is actually one of the leading nations in East Africa when it comes to the introduction of IT and the Internet. South Africa perform very well on all areas with potential for advancement in veracity because of issues surrounding trust and security, social inequality, and crime. However, it has the most digitally open economy, high availability of latest technologies, and substantial use of online services. Rwanda and Kenya do exceptionally well in value, with room for progress in volume due to a limited number of tech hubs, low rates of Internet subscriptions, and less computer ownership. The first-mentioned has a high investment in technology as well as widespread ICT usage within the government, education system and business sector. Most notably, hierarchical clustering analysis shows that The Gambia is falls under the group 'gaining momentum' as opposed to 'lagging behind' or 'forging ahead', as evident in Figure 2.10. However, it has the potential to move towards the last-mentioned if the value component is improved. The BDRI results implicates that policies should maximize opportunities of big data by using incentives like subsidies for digitalization of sectors and opening existing data sources for free use by the public. In addition, they should minimize risks of big data by using regulations like laws for control and privacy, enforcing openness, and transforming isolated silo systems into interoperable and integrated ISs.

2.4 Learning Analytics

LA is most commonly defined as "the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purposes of understanding and optimizing learning and the environments in which it occurs" (LAK, 2011, as cited in Siemens, 2013, p. 1382). At first glance, it may be very reminiscent of EDM, though there are some important differences. Whereas LA "is focused on (…) integrating the technical and the social/pedagogical dimensions" of data analytics, EDM "is generally looking for new patterns in data and developing new algorithms and/or models" (Romero & Ventura, 2013). There are other disparities when it comes to techniques, origins, and discovery too:

> In LA, the most used techniques are statistics, visualization, [social network analysis], sentiment analysis, influence analytics, discourse analysis, concept analysis, and sense-making models. In EDM, the most used techniques are classification, clustering, Bayesian modeling, relationship mining and discovery with models. LA has stronger origins in Semantic Web, intelligent curriculum, and systemic interventions. EDM has strong origins in educational software, student modeling, and predicting course outcomes. (…) In LA, leveraging human judgment is key; automated discovery is a tool used to accomplish this goal. In EDM, automated discovery is key; leveraging human judgment is a tool used to accomplish this goal. (p. 13)

LA is also very similar to academic analytics, which is "more concerned with improving organizational processes, such as personnel management or resource allocation, and improving efficiency within the university" (Siemens, 2013, p. 1384).

2.4.1 The Past, Present and Future of LA

In an introductory article, Siemens (2013) takes us through the past, present and future of LA. The inception of LA is multifaceted and has been influenced by citation analysis, social network analysis, user modeling, cognitive modeling, tutoring systems, adaptive hypermedia and e-learning amongst other fields. These can also be viewed upon as the data sources, processing methods, or final outcomes of LA. Citation analysis, which was pioneered by Garfield in 1995, considers how tracking associations between science articles is important for mapping knowledge domains. Social network analysis, which was introduced by Gransovetter and Milgram and adapted for education by Wellman, looks at how relationships between actors can form complex structures. User modeling focuses on how individual personalities, traits, and motivations influences interactions with computers and how they can be designed better. Cognitive modeling is concerned with tracing how learners develop knowledge and solve problems in order to develop good software tutors. Tutoring systems can be categorized under three levels of 'intelligence' – namely domain knowledge, learner knowledge evaluation, and pedagogical intervention. Adaptive hypermedia is able to personalize learning content and interaction based on the profile of the user. Finally, e-learning refers to any educational processes that occur online and has been especially common in higher education. Taken together, this highlights the interdisciplinary and sociotechnical aspects of the LA.

The history of LA has had an impact on the tools that it presently uses, which can be categorized as commercial or academic. Commercial tools include ones developed by large corporations like IBM and SAS such as SPSS, as well as ones created by smaller market vendors like Desire2Learn. The education sector is already using SISs, MOOCs and LMSs. Some recommender systems and analytics tools have been built on top of them like Degree Compass or Purdue University's Signals which uses data from Blackboard. Research tools include R and Social Networks Adapting Pedagogical Practice which do not have the systems-level adoption, integration, and support that the commercial ones do. In addition to tools, LA also has a variety of techniques and applications. Whereas "techniques involve the specific algorithms and models for conducting analytics", "applications involve the ways in which techniques are used to impact and improve teaching and learning" (p. 1386). The former includes prediction, clustering, relationship mining, and discovery with models while the latter includes creating profiles of users, modeling knowledge domains, and personalization & adaption.

When it comes to the future of LA, Siemens argues that data collection will ideally occur unobtrusively when students are absorbed in genuine learning. The *scope of data capture* should be widened to include sensors on wearable computing devices in order to obtain rich and detailed social, affective, cognitive, and behavioral information. This coupled with *knowledge domain modeling* can be used to personalize learning or to recommend courses for the student. Additionally, it can also be combined with performance outcomes to provide feedback for designers of course content. For this to happen, *organizational capacity* as in individuals with a deep understanding of programming, statistics, and the domain must be strengthened. Furthermore, cross-departmental and intra-institutional collaborations are necessary to triangulate and integrate data from different sources.

2.4.2 Stakeholder Attitudes in LA

In a study, Drachsler and Greller (2012) investigated attitudes towards six critical dimensions of LA within the education sector. The dimensions were identified in a prior literature review and consists of (a) stakeholders, (b) objectives, (c) data, (d) method, (e) constraints, and (f) competencies. Within each dimension, there are different 'hard' (i.e., technical) and 'soft' (i.e., social) issues. Stakeholders include data subjects (e.g., students, parents) who are the suppliers or providers of data and data clients (e.g., teachers, institutes) who are the demanders or receivers of data. Objectives can be reflection which is used for self -evaluation and self-regulation or prediction which is used for interventions and drop-out prevention. Data can either be open and integrated from different sources or protected and secured for the privacy of the students. Methods normally fall under either statistical analysis, EDM, or machine learning – this includes visualization techniques and theoretical approaches. Constraints are usually related to privacy or ethical issues that can be resolved through policies or guidelines and informed consent amongst other things. Lastly, competencies include the ability to interpret and the ability to critically evaluate the data and the results. These dimensions formed the basis of a digital questionnaire which was promoted online in various LA-related communities and received especially many responses from the UK and the US. Most of the respondents were teachers and researchers from higher education.

When it comes to stakeholders, the main beneficiaries of LA was perceived to be students and teachers, and institutions whereas its biggest impact is believed to be on relationships between students and teachers. The main objectives was understood to be reflection and not prediction, such as 'more timely information about the learning progress' and 'better insight by institutions on what's happening in a course'. The most important aspects of data were viewed upon as ‘added context information’ amongst researchers, ‘sharing within the institution’ amongst managers, and ‘containing meta-data’ amongst teachers. The methods trusted to be the most accurate were ‘summarize the student learning progress’, ‘predict relevant learning resources’, and ‘assess learner state of knowledge’. In terms of constraints, although most of the participants indicated that an ethical review board was in place and that anonymization was regarded as relatively safe, sharing student data with all staff was not deemed appropriate. Finally, important learner competencies were considered to be self-directedness and critical reflection, although most respondents felt that students would not be able to interpret LA reports without additional support.

2.4.3 Ethical Issues and Dilemmas in LA

In a literature review, Slade & Prinsloo (2013) take a sociocritical approach to ethical issues and dilemmas surrounding LA. They argue that the unequal power relations between students and higher education institutions is similar to the ones in Foucault’s Panopticon which is a structural prison design that enables pervasive surveillance. Moreover, they point out that “learning as a moral practice [involves] recognizing students as participatory agents with developmental (…) identities”. This stands in stark contrast to common practice today which entails “using student data in service of neoliberal consumer-driven market ideologies” (p. 1511-1512). Legal frameworks tend to focus on how data is used outside rather than inside an institution, and higher education has an inclination to let valuable data sit instead of using it efficiently to improve student learning. How the data is used will depend on what the main goal of the educational institution is – whether it is to increase completion rates or simply maximize profits. “Although students are increasingly aware of the growing prevalence of data mining to monitor and influence buying behavior, it is not clear they are equally aware of the extent to which this occurs within an educational setting” (p. 1516). It is important to maintain good relations with students by managing their perceptions of the school and being transparent about how their data is being used. In some cases, parts of the dataset may be missing or incomplete, leaving the analyst to subconsciously fill the gaps with biased interpretations. In addition, misdirected interventions and unfavorable predictions can have detrimental effects on the students. Generalized models of behavior may oversimplify intrapersonal complexity or interpersonal differences and recommendation engines may not correspond with the students own preferences or goals. Even though “pattern recognition has huge potential for delivering custom-made and just-in-time support to students, there is a danger (…) that pattern recognition can result in keeping individual prisoner to past choices" (p. 1517). It is important to keep in mind that the data only captures one aspect of a person's life at a specific time and place.

In an attempt to solve some of these privacy and ethical issues, Drachsler and Greller (2016) propose a checklist for policymakers and institutional managers named DELICATE. According to them, "the widespread adoption of [LA and EDM] has somewhat stagnated recently, (…) fears and realities are often indistinguishably mixed up, leading to an atmosphere of uncertainty" (p. 89). Although everyone seems to agree that privacy and ethics is important, there are few who have concrete suggestions on to how to uphold it. Drachsler and Greller define argue that "ethics is a moral code of norms and conventions that exists in society externally to a person, whereas privacy is an intrinsic part of a person's identity and integrity" (p. 90-91). The first principles of ethical research resulted from the Nuremberg trials of the cruelties committed by Nazi medics. They include obtainment of voluntary informed consent, avoidance of unnecessary physical and mental harm, and freedom to immediately withdraw from the experiment at any point. Two infamously unethical experiments, the Milgram experiment and the Stanford Prison experiment, remind us of why such ethical codes and committees are needed. Recently, Kramer and colleagues did an experiment in which Facebook users were emotionally manipulated without informed consent which has caused a lot of controversy. In fact, concerns about data misuse amongst the general public has led to the shutdown of inBloom, a LA project that received more than 100 million USD funding by the Gates and Carnegie foundations. When it comes to privacy, the European Commission's report "New Modes of Learning and Teaching in Higher Education" states the following:

> Member States should ensure that legal frameworks allow higher education institutions to collect and analyse learning data. The full and informed consent of students must be a requirement and the data should only be used for educational purposes. (…) Online platforms should inform users about their privacy and data protection policy in a clear and understandable way. Individuals should always have the choice to anonymize their data. (European Commission, 2014, as cited in Drachsler & Greller, 2016, p. 92)

DELICATE is based on a document review of legal texts like these, in addition to a literature review of ethics and privacy in LA, and several Ethics and Privacy for LA workshops. It consists of eight steps: (a) *determination* of data collection goals, (b) *explanation* of data collection process, (c) *legitimization* of the data collectors, (d) *involvement* of stakeholders and data providers, (e) *consent* obtainment before data collection, (f) *anonymization* of individuals, (g) *technical* procedures for privacy, and (h) *external* use accountability.

2.4.4 LA for the Global South

In a newly published article, Prinsloo and Kaliisa (2022) look at privacy regulations in Africa and the implications it has for the institutionalization of LA. As a result of the COVID-19 pandemic, higher education institutions (HEI) have increasingly turned to digital instruction, and it is not likely that Africa will be any exception to this:

> The digitalization and datafication of African HEIs are in many respects, inevitable. Africa will also not escape the impact of the [4IR.] Africa is the world's most populous continent, with estimations that its population will double by 2050, and (…) 19 out of the world's 20 youngest countries [are] on the African continent[.] Africa is seen as *terra nullis*, a new data frontier to conquer. (…) African HEIs will have to consider the implications of agreements with [online program management] providers, many of whom will be located in the Global North. Amid the broader concerns of the implications of the commodification and subsequent asetisation of higher education, concerns about Africa being re-colonised and its data exported and capitalised may be legitimate. (p. 3)

From 2010 to 2019, the number of Africans using the Internet has increased from 81 to 294 million and there has been a 12.98% growth in Internet use from 2000 to 2021. Currently, 43% of the population in Africa has Internet access and this number is projected to increase to 75% by 2030. So far, Moodle is the fastest growing LMS and the most popular one on the African continent. However, as shown by the Cambridge Analytical scandal, the absence of privacy regulations can have some serious consequences. In order to map legal and regulatory policies and frameworks in Africa, Prinsloo and Kaliisa looked at data on 32 African countries mainly based on 6 websites and 15 research articles. They found that Data Protection Africa had the most comprehensive database. Of all 32 countries that data were available for, 13 have enforced legislations, 8 have unenforced legislations, 4 have legislations, and 6 have no legislations. Most countries refer to the GDPR, clearly define personal and sensitive data, require consent prior to data collection, restrict automated data collection, and prohibit sharing data with third parties. However, a number of countries have no regulations for cross-border data transfer or for notification of security breaches. Furthermore, Mauritius is the only state that specifically mentions the right to withdraw consent at any time as well as the right to access, correct, erase, or restrict collected personal data. Prinsloo and Kaliisa conclude that there is a need for higher education institutions to fill gaps in national regulations, for governments to cooperate on the African Union's Data Protection Convention, and for African LA practitioners to develop their own LMSs.

Chapter 3

Context

3.1 DHIS2 Sociotechnical Aspects

3.1.1 HISP & DHIS2

HISP is a global action research network that aims to strengthen the health sector of low- and middle-income countries (LMIC; HISP Centre, n.d.-a; DHIS2, n.d.-a). It is coordinated by the Department of Informatics at the University of Oslo which is also where the HISP Centre resides. In addition, the network consists of several HISP groups around the world (Uganda, Tanzania, South Africa, Rwanda, Mozambique, Malawi, Ethiopia, Kenya, West & Central Africa, Nigeria, Vietnam, Sri Lanka, India, Bangladesh, Indonesia, and Colombia; DHIS2, n.d.-g).

In order to reach their goal, HISP has developed DHIS2, a "[software] platform most commonly used as a health management information system (HMIS)" (DHIS2, n.d.-a). It is the largest HMIS platform in the world, being used by 73 LMIC and affecting around 2.4 billion people or 30% of the global population (HISP Center, n.d.-a). Only at the University of Oslo, there are more than 100 staff members, 60 PhD graduates, and 500 master students working and researching on DHIS2. They frequently collaborate with and receive funding from WHO, UNICEF, GAVI, Norad, the Global Fund, and the Bill and Melinda Gates Foundation amongst many others (HISP Center, n.d.-b). In 2022, they received more than 23 million USD in combined investments.

3.1.2 History

HISP was established in 1994 in post-apartheid South Africa as an effort to decentralize and integrate health services (Adu-Gyamfi et al., 2019). DHIS1 was conceived of and created in 1995, tested and piloted in various Cape Town districts in 1998, and had become the national standard by 2001. As a result of this, it was also set up in other countries like India in 2006 and Mozambique in 1999. However, because the system relied on the Microsoft Access database, it required the Windows OS and a full Microsoft Office stack. Being a standalone installation, it also had to be maintained by a team who had to travel to each health facility in order to keep the systems updated and free from viruses. Because of this, DHIS2 was developed in 2005 as an "open-source software (…) with client-server architecture [which] supports distributed software development (…) and centralized maintenance" (p. 74).

3.1.3 DHIS2 Use Cases

DHIS2 is used for Covid-19 surveillance and vaccine delivery as well as other infectious disease tracking and immunization programs (DHIS2, n.d.-h). It is also utilized for individual-level case-based patient follow-up and data capture by health workers on smartphones with offline functionality. DHIS2 is formally recognized as a WHO Collaborating Center and offers a DHIS2 Heath Data Toolkit which is aligned with the WHO Toolkit for Routine Health Information Systems Data (DHIS2, n.d.-k). It offers analytical, aggregate, and tracker metadata packages for HIV, malaria, tuberculosis, non-communicable diseases, maternal health, birth & mortality rates, nutrition, and more. It is also compatible with the WHO Global Health Observatory Indicator Metadata Registry which contains health-related statistics from almost every country in the world (DHIS2, n.d.-j; WHO, n.d.). DHIS2 can also be integrated with iHRIS which is an internationally recognized software for managing health workers and OpenMRS which is a medical record system that is used all over the globe (DHIS2, n.d.-j; iHRIS, n.d.; OpenMRS, n.d.).

Despite their names, HISP and DHIS2 have also recently been expanded to the water & sanitation, agriculture, and education sectors (DHIS2, n.d.-a). With regard to the last-mentioned, DHIS2 functions as an EMIS in The Gambia, Togo, Uganda, Mozambique, Eswatini, and Sri Lanka (DHIS2, n.d.-d). It is used in "the collection, analysis, visualization, and use of individual and aggregate data from institutions of learning" to "facilitate (…) improvement in learning outcomes and equitable access to education". It includes functionality for school management, student enrollment and attendance, infrastructure management, nutrition and hygiene, resource allocation, and multi-sectoral analysis.

3.1.4 DHIS2 as a Digital Public Good

DHIS2 has been recognized as a Digital Public Good (DPG) by the Digital Public Goods Alliance whose aim is to reach the UN Sustainable Development Goals (SDG) (DHIS2, n.d.-

e). The 17 SDGs are (a) no poverty; (b) zero hunger; (c) good health and well-being; (d) quality education; (e) gender equality; (f) clean water and sanitation; (g) affordable and clean energy; (h) decent work and economic growth; (i) industry, innovation, and infrastructure; (j) reduced inequalities; (k) sustainable cities and communities; (l) responsible consumption and production; (m) climate action; (n) life below water; (o) life on land; (p) peace, justice, and strong institutions; and (q) partnerships for the goals (United Nations, n.d.). The UN Secretary General's Roadmap for Digital Cooperation describe DPGs as "open-source software, open data, open AI models, open standards and open content that adhere to privacy and other applicable laws and best practices" (UN High-Level Panel on Digital Cooperation, 2020).

3.1.5 DHIS2 as an Information System

DHIS2 can be defined as an IS, though the meaning of this term is somewhat vague. IS as a research area has suffered an identity crisis for this exact reason, lacking a conceptual foundation that is not borrowed from management, economics, marketing, psychology, or sociology. As Watson (2014) points out, since the field was conceived under the dawn of the computer age, most people tend to associate IS with digital technology. However, humans have exchanged information in the form of body language, spoken language, written language, mathematics, printed books, morse code, and semaphore towers before computers became an object of everyday life. Watson broadly delineates three criteria for IS; it should "improve an entity's ability to attain its goals (…) [,] improve the ability of entities to cooperate on shared goals (…) [and] transform entities in intended and unintended ways" (p. 517). As we will see in section 3.2.1, DHIS2 fulfills all of these criteria.

3.1.6 DHIS2 as a HMIS and an EMIS

As previously mentioned, DHIS2 is often used as a HMIS or EMIS. The former can be defined as an IS in which the goal is to "produce relevant information that health system stakeholders can use for making transparent and evidence-based decisions for health system interventions" (Sanner, 2021-a, p. 6). The latter can be defined as an IS that aims to "produce, manage and disseminate educational data and information as part of their IT infrastructure" (Martins et al., 2019, p. 184). However, the fragmented nature of the IS landscape makes reaching these objectives somewhat difficult (Hasselbring, 2000). Many systems were created in isolation – they use different hardware, operating systems, programming languages, and data models – which makes sharing information between them very challenging. The keyword in HMIS and EMIS is *management*, which is not possible without a complete overview of the respective sectors. Section 3.2.1 will take a look at how DHIS2 addresses integration, interoperability, and standards.

3.1.7 DHIS2 Ownership and Governance

DHIS2 is open-source, free of charge, and released under the BSD 3-clause license (DHIS2, n.d.-a). In practice, "anyone can access to [sic] the source code and may even modify and redistribute the software. The only restrictions are that the copyright notice must be maintained in the source code". Also, "neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission". Each organization is given an instance of the DHIS2 platform that they fully own, including any applications that they develop and any data that they collect (DHIS2, n.d.-a). In other words, unlike the case of some proprietary systems, they do not need to pay a fee in order to gain access to their own data. However, this also implies that the organization needs to determine how to securely store these data. Local laws, privacy concerns, and server maintenance capacity are some of the factors that must be taken into consideration. The HISP network can help by offering their expertise and weighing in with their opinion. Another proposed solution is detailed in section 3.2.2.

3.2 DHIS2 Technical Aspects

3.2.1 DHIS2 as a Digital Platform

DHIS2 is also a digital platform. It consists of a platform core provided by the platform owner upon which complementors can develop their own applications (Tiwana, 2013, Chapter 1). It is multisided in that it mediates between different stakeholder groups, such as Ministries of Health, non-governmental organizations, policymakers, donors, entrepreneurs, consultants, researchers, and students (de Reuver et al., 2018; Adu-Gyamfi et al., 2019). It comes with network effects, meaning that its value increases with the number of users it has (de Reuver et al., 2018). As it encourages both cooperation and generativity, it can be considered to be both a transaction platform and an innovation platform (Li, 2021). A third-party developer can either use the core modules and apps that come with the platform, customize and combine modules to create a new app, or use the modules and apps that other developers have created from the DHIS2 App Hub (DHIS2 Documentation, n.d.-e). DHIS2 fulfills the criteria of IS because the core modules and apps help the user reach their goals, the modules and apps of other users help them cooperate, and the flexibility and scale and of it leads to intended and unintended consequences. The platform also ensures integration, interoperability and standards through its boundary resources such as APIs, SDKs, documentation, and licenses (Sanner, 2021-c).

3.2.2 DHIS2 as a Data Warehouse

Since DHIS2 is used as a HMIS/EMIS in low resource contexts, it offers data warehouse services (DHIS2 Documentation, n.d.-a). A data warehouse differs from a traditional database in that it is specifically designed for analysis and that it gathers large amounts of data from multiple sources. More specifically, these sources would typically be transactional systems, often also termed operational systems, that collect fine-grained data points on a daily basis. The data then goes through an operational data store where it is transformed into a unified format. Finally, the data is processed in the data warehouse, using data mining techniques or online analytical processing. The efficiency of the data warehouse is highly dependent on the similarity of the operational systems. If the transactional systems are very dissimilar, data transformation becomes more resource intensive, thus the data will not be updated as frequently. Therefore, it has become more and more common to combine the operational system and data warehouse into one. DHIS2 is an example of this, which makes it suitable for both data capture and data analysis. It takes a dimensional approach to data storage, meaning that data objects are stored as facts with dimensions. The former refers to the data value (numerical, textual, or binary) of the object, whereas the latter consists of the data element (e.g., measurement unit or description), time period, organizational unit, and categories of the object. These dimensions are required for all object and cannot be skipped, although the user can add any number of custom dimensions according to their needs. DHIS2 uses PostgreSQL as their relational database management system (DHIS2, n.d.-j).

3.2.3 DHIS2 as an API

DHIS2 also offers a REST-based Web API that makes it easier to communicate with the data warehouse or any other data storage solutions (DHIS2 Documentation, n.d.-c). All data objects exist as resources with a URI that can be represented in different formats including HTML, PDF, and XLSX. Resources can be added (POST), retrieved (GET), updated (PUT), or removed (DELETE) using their URI and they can be linked to each other using dimensions or attributes. In order to keep track of which changes were made by whom, the DHIS2 API offers three options for authentication. The first option is basic authentication, in which the username and password are sent in plain text, making it only suitable for use on servers with SSL/TLS (HTTPS) encryption for security reasons. The second option is personal access tokens, which is an alternative to passwords, though it requires considerable configuration and cannot be separated from the user it belongs to. The third option is OAuth2, in which a third-party client is given a reusable bearer token and allowed access on behalf of a DHIS2 user, using either a password, refresh token or authorization code.

3.2.4 DHIS2 as a Docker Image

DHIS is written in Java and can thus run on any OS for which there is a Java Runtime Environment, including Windows, Linux, and MAC OS X (DHIS2, n.d.-j). It can be downloaded as an image and spun up in a container using Docker (DHIS2 Developers, n.d.-b), which is very similar to virtual machines except that it does not require downloading large OSes (Docker, n.d.). Containers essentially make it possible to run an application developed on one computer on a completely different computer regardless of infrastructures, runtime environments, or installed dependencies. Images are instructions for what the container should include and how it should be built (Docker Documentation, n.d.). Docker containers can be run on one or multiple servers, physical or virtual machines, locally in a data center or remotely on cloud providers or a combination of both, making it easy to scale as needed. Last but not least, Docker "significantly reduce the delay between writing code and running it in production (…) [and is] great for continuous integration and continuous delivery workflows". Because DHIS2 uses Docker as a part of their delivery pipeline, it reaps all the benefits that comes with using its containers and images. It comes with standardized data packages for national and international use and standard configurations based on WHO recommendations which can be customized to meet local needs (DHIS2, n.d.-j). DHIS2 only requires an online connection for initial setup and after that it continues to function even when offline which makes it very suitable for LMIC with lacking infrastructure. More specifically, it can be used for data entry even if there is no internet or if it is very slow or unstable, the data will simply be uploaded and updated when connectivity is regained. This also applies to the DHIS2 Android Capture App. In addition, data entry can also be configured for SMS.

Figure 3.1

DHIS2 Web Application

Note. From *DHIS2 Overview*, by DHIS2, n.d.-f (https://dhis2.org/overview/). Copyright by DHIS2.

3.2.5 DHIS2 as a Web or Android Application

DHIS2 is currently only available on Android and not iOS, although the last-mentioned can still use a web browser to access it (DHIS2 Community, 2017), as shown in Figures 3.1 and 3.2. As a web application, DHIS2 is written in JavaScript, and it follows the W3C standards for HTML and CSS (DHIS2, n.d.-j). It can be run on most web browsers including Chrome, Firefox, Opera, Safari, and Edge. DHIS2 also has a design system that offers principles, patterns, and components for programmers with little experience in design (DHIS2 Developers, n.d.-a; Cooper, 2019). It also comes with a UI library that includes buttons, menus, tables, layouts and much more. It is important that developers use these resources diligently in order to create an enjoyable user experience, make the apps intuitive to use, and maintain an overall cohesive look. As an Android application, the DHIS2 Android SDK is a library written in Java and Kotlin that can be installed as a dependency and comes with built in methods to interact with the web API (DHIS2 Documentation, n.d.-d). It uses SQLite as a local offline database, RxJava for asynchronous operations, Dagger for dependency injection, Jackson for JSON parsing, Retrofit and OkHttpClient for API communication, and SQLBrite for database migrations. At the moment, DHIS2 only comes in English, French, Spanish, Portuguese, Hindi, Vietnamese, Chinese, and Norwegian (DHIS2, n.d.-j). However, translations into other languages are ongoing projects that can be found on the Transifex localization platform.

Figure 3.2

DHIS2 Android Application

Note. From *Android Capture App V2.5 Overview*, by DHIS2, n.d.-b (https://dhis2.org/android/version-2-5/). Copyright by DHIS2.

3.2.6 DHIS2 Developer Resources

DHIS2 offers a Developer Portal which is a sort of quick start guide for third-party complementors (DHIS2, n.d.-c). It includes instructions on how to setup a local instance of the DHIS2 web app, how to use the UI library components, how to perform simple data queries and mutations, how to set up automatic code formatting, and how to solve common development issues, how to set up continuous updates to the App Hub (DHIS2 Developers, n.d.-b). More detailed information on the use, implementation, development, and management of DHIS2 can be found in the documentation, including comprehensive descriptions of the API and the Android SDK (DHIS2 Documentation, n.d.-b, n.d.-c, n.d.-d). There is also a Community of Practice forum in which questions can be answered, various topics can be discussed, and networks can be built (DHIS2, n.d.-c). Moreover, there is an Academy that offers training and capacity building on two levels in the form of both digital courses and physical conferences (DHIS2, n.d.-i). In addition, DHIS2 has published a Development Roadmap on Jira with an overview of future updates as well as insight into how these are prioritized, and eager developers are welcome to beta test these (DHIS2, n.d.-c). Furthermore, the DHIS2 source code is available on GitHub. Programmers can fork the repositories in order to explore and play with the code without affecting the codebase or open pull requests to contribute and add on to it.

3.3 Leap Learning

Leap Learning, formerly known as Alphabet King, is a Norwegian company founded in 2015 (Oppvekstportalen, 2018) whose "mission is to be a game-changer in education… globally (…) by developing innovative educational apps and resources" (Leap Learning, n.d.-c). They do this in collaboration with wide range of organizations, most notably with Africa Startup, which is also a Norwegian company founded in 2008 (Africa Startup, n.d.-a). Their aim is to "improve livelihoods in Africa and around the world, through education in agricultural innovation, environmental protection, entrepreneurship, and new innovative learning methods for basic literacy, numeracy, logic and concepts". The last-mentioned is achieved through Leap Learning, whereas the first-mentioned is done through MyFarm, which is a garden and educational center for sustainable agriculture (Africa Startup, n.d.-b). Leap Learning directly impacts around 50 000 children and adults at more than 300 schools in over 20 countries (Leap Learning, n.d.-a).

3.3.1 Leap Learning Partners

In addition to this, Leap Learning is also one of the corporate social responsibility projects of Hydro, a leading industrial aluminum and renewable energy company, with a revenue of approximately 150 million NOK or 16 million USD (Leap Learning, n.d.-c; Hydro, 2022). Leap Learning is also one of UNESCO's partners in the Empowering Adolescent Girls and

Young Women through Education in Tanzania project (Leap Learning, n.d.-c). Furthermore, they receive backing from Innovation Norway, the official business support system of the Norwegian government (Innovation Norway, 2020). They also cooperate with Save the Children, a humanitarian non-governmental organization that has helped over 80 million people in more than 100 countries through their fundraisers in 2020 alone (Save the Children, n.d.). Moreover, they collaborate with the Ministry of Basic and Secondary Education of The Gambia, the Adventist Development and Relief Agency in Niger, and Fandema which works with education and occupational training of women in The Gambia (Leap Learning, n.d.-c).

Figure 3.3

Leap Learning Lab

Note. From *Leap Learning Labs*, by Leap Learning, n.d.-b (https://leaplearning.no/learning-labs). Copyright by Leap Learning.

3.3.2 Leap Learning Labs

Leap Learning offers "a new educational method (…) combining technology and hands on games into a flipped classroom (…) with stations, to ensure that all children can learn at their own level and in their own pace" (Leap Learning, n.d.-a). It is called Leap Learning Labs and consists of a collection of educational apps and a collection of educational printable posters and cards that are meant to teach school children reading, writing, and mathematics, as shown in Figure 3.3. The collections can be divided into two main categories – literacy and numeracy (Leap Learning, n.d.-b). The former can be further broken down into the subcategories of (a) logic and concepts, (b) letter sounds, (c) listening to stories, (d) letter shapes, (e) writing, (f) blending, (g) words, (h) sentences and reading comprehension, and (i) reading stories. The latter can be further split into the subcategories of (a) logic and concepts, (b) numbers, (c) addition, (d) subtraction, (e) multiplication/division, (f) measure, (g) concepts, and (h) word games and mixed math. There are currently more than 70 labs set up around the world that are used by schools, humanitarian organizations, governments and private people (Leap Learning, n.d.-a; Oppvekstportalen, 2018).

Figure 3.4

Leap Learning Apps

Note. From *Leap Learning App Universe*, by Leap Learning App Universe, n.d. (https://webapps.leaplearning.no/en/num). Copyright by Leap Learning.

3.3.3 Leap Learning Hotspots

If a classroom is not available, Leap Learning also offers Hotspots, which consists of a preloaded tablet embedded inside a metal box that can be put up rural areas with minimal infrastructure (Leap Learning, 2018a). The tablet also comes with a local WiFi which allows users to stream the content on their own device. A pilot for these was set up in The Gambia, which was very successful. “We saw that the tablets were still in use 24/7, they were not stolen and there was no problem with the power supply. The biggest challenge has actually been that so many children are coming at the same time so there are big arguments over who's turn it is to play". There are more than 30 hotspots set up all over the world.

3.3.4 Leap Learning Apps

There are more than 500 apps which are available on both Android, iOS, and on their website. In order to keep them organized, there is also a Leap Store app and a Leap Launcher app, as displayed in Figure 3.4. A free printable book that tracks the learning progress of the students is also available. In order to gain access to the apps, a license must be purchased for 99-299 NOK or 11-33 USD, depending on the country/language in which it will be used. The apps are "tailor-made by pedagogical experts" and can be used "without guidance from qualified teachers (…) [or] access to a proper school system" (Leap Learning, 2018).

3.4 The Gambia

3.4.1 General

The Gambia has a population of 2,268,000, making it one of the most densely populated countries in Africa (Britannica, 2022). Most people live in larger cities and urban areas that are concentrated near the coast, including the capital Banjul and the largest city Serekunda (CIA World Factbook, 2022). The official language is English, although various ethnic groups also have their own languages. The largest ones are Mandinka, Fulani, Wolof, Diola, and Soninke as shown in Figures 3.5 and 3.6 (Britannica, 2022). The overwhelming majority of the Gambian population is Muslim, although there are a select few Christians as well. With high infant mortality and population growth rates, two thirds of the population is under the age of 30 as seen in Figure 3.7 (Britannica, 2022), and life expectancy is around 66 years (CIA World Factbook, 2022). The biggest health risks that the country is facing are lacking sanitation, insufficient drinking water, hepatitis A, typhoid fever, malaria, dengue fevers, rabies, meningitis, and tuberculosis. Infrastructure such as housing, plumbing, pavement, electricity, and internet is poor, especially in rural areas (Britannica, 2022).

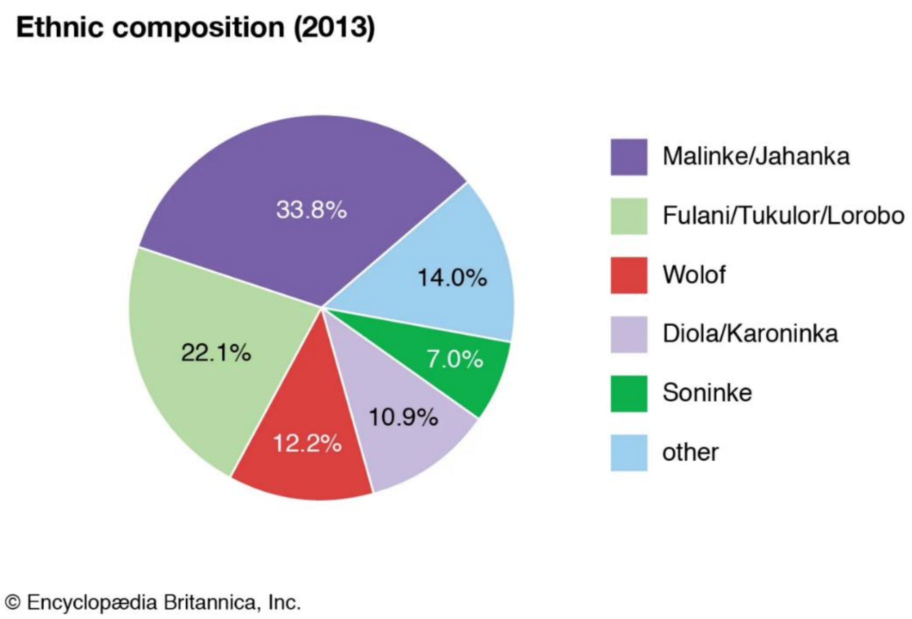
Figure 3.5

Ethnic Group Spellings

Note. There are multiple spellings for the same ethnic group depending on the language and dialect used (Britannica, 2017b, 2021, 2017c, n.d., 2012).

Figure 3.6

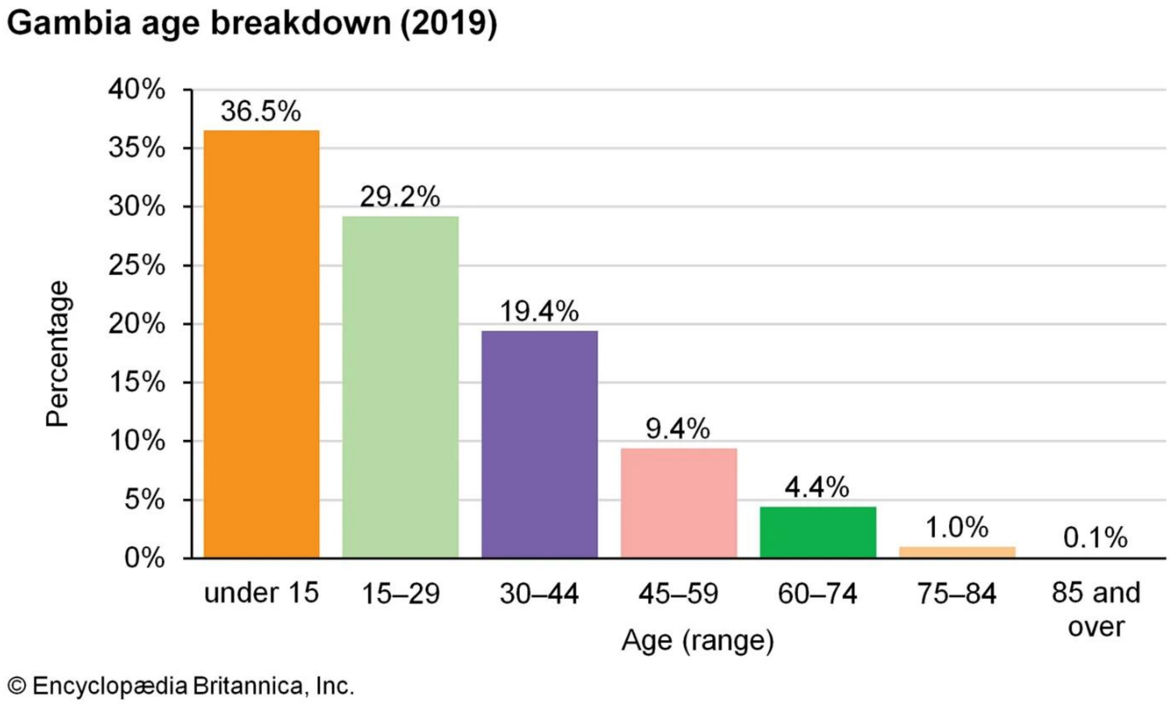
Ethnic Composition



Note. From *The Gambia*, by Britannica, 2022 (https://www.britannica.com/place/The-Gambia). Copyright by Encyclopædia Britannica, Inc.

Figure 3.7

Age Distribution



Note. From *The Gambia*, by Britannica, 2022 (https://www.britannica.com/place/The-Gambia). Copyright by Encyclopædia Britannica, Inc.

3.4.2 Geography

With a total area of only 11,295 square kilometers, The Gambia is the smallest country of mainland Africa (Country Reports, n.d.-a). As displayed in Figure 3.8, it consists of a narrow strip of land 480 kilometers long and 25 to 50 kilometers wide surrounding a river by the same name (Britannica, 2022). Located northwest in the continent, it is almost completely surrounded by Senegal on all sides, except for where the river meets the Atlantic ocean, as presented in Figure 3.9. Situated near the equator, it enjoys a subtropical climate with a rainy season lasting from June to October when it is prone to droughts and a dry season from November to May when dust winds blow in from the Sahara desert (Country Reports, n.d.-a). The land is mainly made up of flood plains surrounded by low hills – specifically sandy beaches, mangrove swamps, open savannah, sparse woodlands, and red-iron stone cliffs (Country Reports, n.d.-a; Nations Encyclopedia, n.d.).

Figure 3.8

Map of Africa

Note. The Gambia is located on the far left hand side of the map. From *Africa Map and Satellite Image*, by Geology, n.d. (https://geology.com/world/africa-satellite-image.shtml). Copyright by Geology.

Figure 3.9

Map of The Gambia

Note. The Gambia River starts in Guinea and flows westwards into the Atlantic Ocean (Britannica, 2017a). From *The Gambia Map and Satellite Image*, by Geology, n.d. (https://geology.com/world/the-gambia-satellite-image.shtml). Copyright by Geology.

3.4.3 Economy

Agriculture makes up most of the economy in The Gambia, followed by industry and then services (Country Reports, n.d.-b). Most of the cultivated crops are peanuts, though there are also yams, eggplant, tomatoes, rice, lentils, and citrus (Britannica, 2022). Similarly, a majority of the industry is devoted to peanut processing. Other manufactured products include food and beverage products, textiles and footwear, and wood crafts. Recently, there has also been an increase in tourism and construction. Additional contributions to the economy include animal husbandry and fishing. There is also a potential for petroleum, natural gas, and hydroelectricity extraction. Main exports include peanuts, cotton, rice, and cattle whereas critical imports comprise manufactured items, petroleum products, lumber, and cement.

3.4.4 Government

The governmental form in The Gambia is that of a democratic multi-party presidential republic, in which the president is directly elected by universal adult suffrage in five-year terms (Britannica, 2022). Major political parties include Alliance for Patriotic Reorientation and Construction, Gambian People's Party, Progressive People's Party, and United Democratic Party (Access Gambia, n.d.-a). The president appoints members of the Cabinet from the House of Representatives, which consists of the vice president and various ministers, and together they hold executive power. The Ministry of Basic and Secondary Education (MoBSE) and the Ministry of Higher Education, Research, Science and Technology (MoHERST) are together responsible for education. Legislative power resides with the National Assembly whose 53 members also serve five year terms and in which 48 are directly elected by the people and 5 are appointed by the president. Judiciary power lies with the Khadis Courts, District tribunals, Magistrates Courts, Supreme Court, and Court of Appeal. Administratively, the country is divided into six regions: (a) Banjul, (b) West Coast, (c) North Bank, (d) Lower River, (e) Central River, and (f) Upper River (Statoids, 2017). These correspond to the following eight Local Government Areas: (a) Banjul/Kanifing, (b) Brikama, (c) Kerewan, (d) Mansakonko, (e) Kuntaur/Janjanbureh, and (f) Basse. The regions can again be sub-divided into 43 districts (MoBSE, 2021).

3.4.5 History

One of the first ethnic groups that settled around the Gambia river were the Jola people, possibly as early as 5,500 years ago (Access Gambia, n.d.-b). Between the 5th and 8th century, the Soninke tribe was dominant and around the 1500s the Malinke migrated and brought with them the religion of Islam. Partly as a result of trans-Saharan trade and the Mali empire, the Arabs have been an important influence of Gambian culture. The first contact with imperialist Europe occurred in the 1400s, when the Portuguese established trading routes exchanging salt, iron, pots and pans, firearms and gunpowder for ivory, ebony, beeswax, and gold. They scaled up slave trade heavily and made substantial profits from trade which enticed other European countries to enter. By the 1600s, they were driven out by the French and British who were competing for control over Senegambia until the end of the nineteenth century in which The Gambia was settled as a colony of Britain. In 1807, slavery which had dominated trade since the mid-seventeenth century was abolished, and commerce centered around peanut oil.

As a result of World War II, a powerful movement for nationalism and independence swept across Africa. Millions of Africans were being drafted to fight to free European countries while still not being free themselves and receiving very little credit for it. During the war, they were armed with military knowledge and leadership skills, and they watched the Europeans bleed, die, and be defeated just like any other people. After the war, Europe was in financial ruin and unable to maintain the colonies or fight back resistance in them. In 1965, The Gambia become politically independent from Britain, and in 1970 it became a republic with Dawda Jawara as president, who was continually re-elected after that. However, as a result of economic crisis, a military coup in 1994 instated Yahya Jammeh as president. During this period, infrastructure like airports, hospitals, roads and schools were improved, although the reign became increasingly authoritarian, oppressive, and violent. In 2016, Adama Barrow won the election, and after considerable military intervention from the Economic Community of West African States, Jammeh eventually conceded and went into exile. Barrow was reelected in 2021 and earlier victims of human rights abuses were compensated by the new government.

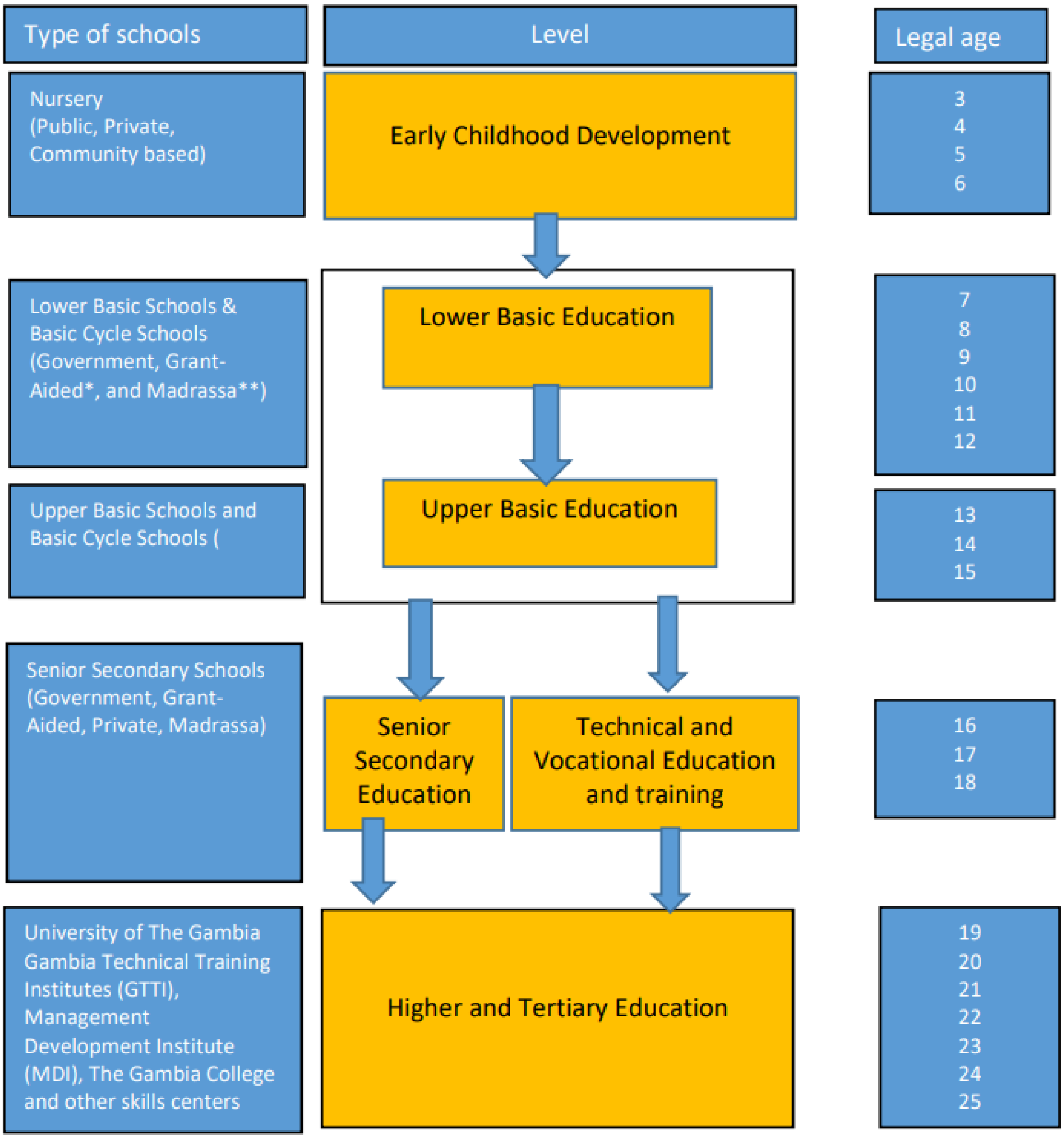
3.5 The Gambian Educational System

3.5.1 Structure

As seen in Figure 3.10, the formal educational system of The Gambia can be divided into four blocks (MoBSE & MoHERST, 2017). Early childhood education consists of early childhood development centers (ECD, level 1-3, age 4-6, duration of 3 years). Primary education consists of lower basic schools (LBE, grade 1-6, age 7-12, duration of 6 years) and upper basic school (UBE, grade 7-9, age 14-16, duration of 3 years). These may sometimes be combined into basic cycle schools (BCS/BCE, grade 1-9, age 7-16, duration of 9 years). Secondary education consists of senior secondary school (SSE, grade 10-12, age 16-18, duration of 3 years) or technical and vocational education and training. Tertiary education consists of tertiary institutions, college, or university. This includes The Gambia Technical Training Institute, the Management Development Institute, The Gambia College, and the University of The Gambia. There are two types of schools with two types of management each. Public schools are generally government-owned or grant-aided, whereas private schools are either commercial or Madrassahs. The last-mentioned are educational institutions of Islamic theology (Britannica, 2020-a).

Figure 3.10

The Gambian Educational System



Note. From *Education Sector Strategic Plan 2016-2030*, by MoBSE & MoHERST, 2017 (https://www.globalpartnership.org/content/2016-30-education-sector-plan-gambia).

3.5.2 Calendars and Timetables

The school year is normally divided into three terms; the first one lasts from September to December, the second term lasts from January to March, and the last one from April to July (Banjul American International School, 2020; Marina International School, n.d.; The Gambia Academy, n.d.). However, due to the various ethnicities celebrating distinct holidays, there is a different but adaptable calendar for each region, which is distributed before the end of the last term of the school year (MoBSE, 2020). The administration of each school ensures that local holidays are accounted for in terms of lost instruction time and this is usually done once every school year or every term. There should be a minimum of 880 hours per year and 23 hours per week of instruction time, and in order to achieve this there should be at least 39 weeks of regular school time each year. In addition, each lesson should be approximately 30-45 minutes in length. Due to a high student to school ratio in urban areas, there are usually two classroom shifts, one in the morning and one in the afternoon. On the flip side, due to a low student to school ratio in rural areas, there may be some multi-grade schools. These are called double-shift and multi-grade schools respectively.

3.5.3 Curriculum

At grades 1-3, common subjects such as languages (English or French and a national language), mathematics, and integrated studies take up 70% of the teaching period, as presented in Table 3.11 (MoBSE, 2011b). Specialized subjects such as physical education, life skills, religious knowledge, and Arabic take up 15% of the teaching period. Decentralized school curricula such as creative arts & handicraft, music, home economics, technology & engineering, and ICT take up another 15% of the teaching period. At grades 4-6, integrated studies are replaced by science and social & environmental studies in common subjects, and national languages become a part of specialized subjects. At grades 7-9, literature is added into common subjects, agricultural science is added into specialized subjects, ICT becomes a part of specialized subjects. Furthermore, metal work, wood work, and technical drawing is added to decentralized school curricula.

Table 3.11

The Gambian Curriculum Framework

Common Subjects	School-Based Curriculum	
	Specialized Subjects	Decentralized School Curricula
	Grades 1-3	
English/French	Physical Education	Creative Arts & Handicrafts
National Language	Life Skills	Music
Mathematics	Religious Knowledge	Home Economics
Integrated studies	Arabic	Technology & Engineering
		ICT
	Grades 4-6	
English/French	Physical Education	Creative Arts & Handicrafts
Literature	Life Skills	Music
Mathematics	National Languages	Home Economics
Science	Religious Knowledge	Technology & Engineering
Social & Environmental Studies	Arabic	ICT
	Grades 7-9	
English/French	Physical Education	Arts & Crafts
Literature	Life skills	Music
Mathematics	National Languages	Home Economics
Science	Religious Knowledge	Technology & Engineering
Social & Environmental Studies	Arabic	Metal Work
	ICT	Wood Work

Agricultural Science	Technical Drawing

Note. From *Curriculum Framework for Basic Education*, by MoBSE, 2011 (http://www.ibe.unesco.org/fileadmin/user_upload/archive/curricula/gambia/gm_befw_2011_eng.pdf).

3.5.4 Assessment

There are five national standardized assessments, tests, and examinations (MoBSE, 2017). The Early Grade Reading Assessment (EGRA) and Early Grade Mathematical Assessment (EGMA) are competence-based and conducted biannually on a small but representative sample of school children at grades 1, 2 and 3. EGRA measures pre-reading skills such as phonetic awareness and letter names, reading skills such as narrative reading and reading comprehension, and listening comprehension (UNESCO, 2021-a). EGMA measures number identification, quantity discrimination, missing numbers, shape identification, addition, subtraction and word problems. The National Assessment Test (NAT) is curriculum-based and administered to all pupils biannually at grades 3 and 5 as well as annually at grade 8. For grade 3, it tests English, mathematics and integrated studies. For grades 5 and 8, it tests English, mathematics, science, and social & environment science. The Gambia Basic Education Certificate Examinations (GABECE) and the West-African Secondary School Certificate Examinations (WASSCE) are curriculum-based and taken by all students each year at the end of grade 9 and 12 respectively (MoBSE, 2017). They are tested in the 4 core subjects of English, mathematics, science, and social & environmental science as well as 3-5 elective subjects (as described in section 3.2.3) where they only receive final grades for the best 2. The GABECE and WASSCE must be passed in order for the student to receive UBE and SSE diplomas and their credits will determine if the student is admitted into SSE and tertiary institutions.

3.5.5 Management

At the national level, the ministries of education are responsible for policy development, management, and coordination (MoBSE, 2017). MoBSE has directorates for (a) planning, policy analysis, research and budgeting; (b) human resources, (c) basic and secondary education programs; (d) standards and quality assurance; (e) science and technology education (STED); and (f) curriculum research, evaluation and development (CREDD). In addition, there is a project coordination unit who is in charge of donor partnerships and grant disbursement. Moreover, there is also an advisory council on education who makes sure that the ministry is aligned with national laws. MoHERST has similar directorates and support structures, albeit focused more on research, science and technology. At the local level, there are regional education directorates who are responsible for implementing and evaluating policies. Under these are the cluster monitors who are responsible for quality assurance in schools. At the school level, the headteacher or principal, the Senior Management Team (SMT), the Parent Teacher Association (PTA). the School Management Committee (SMC) and Teacher Trainee (TT) mentors are responsible for administration and management (MoBSE, 2020). The SMT consists of experienced teachers that assist the headteacher or principal in large schools. The PTA is responsible for school premises, teaching environment, and pupil welfare. The SMC is elected by the PTA and act on their behalf in order to reach those goals. It can be further divided into subcommittees for (a) leadership and management, (b) community participation, (c) curriculum management, (d) teacher's professional development, (e) teaching and learning resources, and (f) learner welfare and school environment. The subcommittee for teacher's professional development appoints mentors amongst experienced teachers to guide TTs who are currently in training and not qualified teachers yet.

Chapter 4

Method

4.1 Research Questions

As previously mentioned in Section 1.2, the research questions were:

- RQ1: What are the opportunities and challenges or digitizing the education sector in developing countries?
  - RQ1a: What are the existing strengths and weaknesses of the Gambian education system and how can technology be used to leverage it?
- RQ2: How can data from educational apps be used to inform school management and planning in a low resource context?
  - RQ2a: How can DHIS2 be used for educational data collection in The Gambia?
  - RQ2b: How can data from Leap Learning be used to inform educational policymaking by the Gambian Ministry of Education?

4.2 Data Collection

Document analysis was used to answer RQ1, and mainly reports, policies, and statistics from the MoBSE website was used to answer RQ1a. Presentations and workshops from conferences as well as interviews were used to answer RQ2. The DHIS2 Level 2 Academy for Education was used to answer RQ2a. The conference took place on the 25th – 29th of April 2022 in Banjul, The Gambia. Interviews with Leap Learning and directorates of MoBSE were used to answer RQ2b. Informal and unstructured interviews was used. A nonprobability snowball convenience sample was used to find suitable interview candidates.

4.3 Data Analysis

Inductive semantic descriptive thematic analysis was used on both the document analysis, conference workshops, and interviews. As previously mentioned in section 1.3, this thesis is rooted in critical research and critical theory. The document analysis was therefore used to determine a chronology of achievements and difficulties in the Gambian educational system. It was of special interest to see the extent to which policymaking have had an effect on education statistics. The goal of the interviews and conference was to establish a timeline of the use of DHIS2 and Leap Learning in The Gambia and find out how they can be used to inform policymaking. The role of technology in supporting the education system will be discussed in light of critical theory and the literature review in section 6.

Chapter 5

Results

5.1 Document Analysis

The documents used were primarily the Education Sector Strategic Plan 2016-2030 (ESSP), the Education Statistics Yearbooks of 2015/2016 (ESY16) and 2020/2021 (ESY21). The secondary documents that were used were the Cluster Monitor Manual 2011 (CMM) and the School Management Manual 2020 (SMM). The ESSP provides some background info about The Gambia and its education system, before detailing enrollment, assessment, financing, policy development, management, and planned interventions. The ESY16 and ESY21 supply comprehensive statical data on schools, classes, enrollment, teaching staff, facilities, and textbooks divided by year, level, gender, region, and district. The CMM describes the basic expectations of cluster monitors, how they can establish good relationships with schools, and what their various tasks and responsibilities are. The SMM details leadership & management, community participation, curriculum management, teacher's professional development, teaching & learning resources, and learner welfare & school environment.

The ESY16 was compared to the ESY21 in order to assess the effects of ESSP. However, it is important to note that the ESSP includes many policies whose effects are difficult to measure, and for that reason, there are not always statistics for them in the ESY16 and ESY21. Therefore, this section will focus on policies and statistics that overlap, whereas Chapter 6 will consider the implications the ones that were omitted. The CMM and SMM were mainly used to fill in gaps and resolve unanswered questions from the conference. Here, it is also worthy to note that some topics in the CMM and SMM were not covered in the conference and vice versa. Again, the emphasis will be on the topics that did overlap, and all of the documents will also be used to supplement Section 5.2.

Before moving on, it is imperative that the reader is familiarized with some statistical terms. Enrollment is usually represented by the number of students enrolled. However, there is also gross enrollment ratio (GER) and net enrollment ratio (NER). Both are the number of students that are enrolled into a school at a particular grade divided by the total number of children in the population at the corresponding age expressed as a percentage (The World Bank, n.d.). However, the former accounts for students that have repeated or moved up a year, whereas the latter only accounts for students that are the appropriate age for their grade. Intake is represented by gross intake ratio (GIR), which is the number of enrolled students in the first grade of early childhood or primary education divided by the total number of children in the population at the corresponding age expressed as a percentage (MoBSE, 2017). Completion is represented by completion rate (CR), which is the number of enrolled students in the final grade of primary or secondary education divided by the total number of children in the population at the corresponding age expressed as a percentage.

5.1.1 Education Access

First, a general theme of "education access" was found which consists of the specific themes "management", "schools", and "enrollment".

The ESSP aims to expand ECD centers to all communities, train facilitators to cater for different native languages, provide furniture and teaching/learning material, monitor and evaluate the effectiveness of the program, and continue to integrate it with formal education. Universal LBE has been achieved, and it will now be expanded to UBE, where a NER and CR of 100% will be the target. In addition, a minimum transition rate of 67% and an idealistic transition rate of 70% by 2030 will be the goal. Moreover, support and cooperation with Madrassas will be continued through provision of facilities, teacher salaries, and customized calendars. Opportunities for accommodation of Madrassa graduates into tertiary and higher education institutions will be explored. All school taxes, fees, and levies will be replaced with carefully managed, monitored, and considered school improvement grants. Out-of-school children will be assisted with conditional cash transfer to encourage them to participate in non-formal education literacy and numeracy classes. Regular audits will be administered to examine and evaluate these programs.

In 2016, there were 1,141 ECD centers, of which 33% were managed by the government, 50% were private, 15% were Madrassas, and 1% were other. This indicates that mainly children from more affluent families have the opportunity to participate in preschools and get a head start on their education. There were 100,349 children enrolled, of which 47% were at level 1, 29% at level 2, 19% at level 3, and 6% at level 4. This suggests that most of the children need simpler classes. There were 3,076 classes, of which 98% were morning classes and only 2% were afternoon classes. This illustrates that there is little need for double-shift ECD centers. The GER was only at 45.8%, which supports this.

In 2021, there were 1,460 ECD centers, of which 32% were managed by the government, 45% were private, 20% were Madrassas, and 3% were grand-aided. Although there is a lower portion of children enrolled in private preschools, there is still a way to go in terms of providing equal opportunity of participation for children from less affluent families. There were 126,573 children enrolled, of which 45% were at level 1, 29% at level 2, 20% at level 3, and 5% at level 4. The distribution is a little bit more even. There are 3,958 classes, of which 99% are morning classes and only 1% are afternoon classes. It indicates that there is still a low demand for double-shift ECD centers. The GER was at 53.1%, which supports this.

In 2016, there were 958 LBE schools, of which 51% were managed by the government, 15% were private, 30% were Madrassas, 4% were grant-aided. This signifies that children from less affluent have equal opportunities of LBE attainment. In addition, 25% of them were basic cycle, 49% were double-shift, and 16% were multi-grade schools. The high portions of the two last-mentioned suggests that pupils are unevenly distributed amongst the schools. There were 308,729 pupils enrolled, of which 22% were grade 1, 20% were grade 2, 18% were grade 3, 15% were grade 4, 13% were grade 5, and 12% were grade 6. There are slightly fewer classes higher grades which may indicate some dropout, repeating, grade skipping or simply different cohort sizes. There were 8,586 classes, of which 73% were morning classes and 27% were afternoon classes. The afternoon classes are generally evenly distributed amongst the different grades. The GIR was at 121.6%, the GER at 104.0%, the NER at 84.6%, the CR at 75.4%, and there were 5.2% repeaters. There are remarkably few repeaters, but potential for improvement in NER and CR.

In 2021, there were 1,229 LBE schools, of which 45% were managed by the government, 18% were private, 33% were Madrassas, and 4% were grant-aided. This may signify that more private and Madrassa schools have been registered by MoBSE. In addition, 27% of them were basic cycle, 57% were double-shift, and 16% were multi-grade schools. There is still an uneven distribution of students amongst different schools. There were 401,333 pupils enrolled, of which 20% were grade 1, 19% were grade 2, 17% were grade 3, 16% were grade 4, 14% were grade 5, and 13% were grade 6. Again, this may indicate dropout, repeating, grade skipping, or just cohort differences. There were 12,298 classes, of which 64% were morning classes and 36% were afternoon classes. There is a clear trend towards more older pupils taking the afternoon classes. The GIR was at 124.5%, the GER at 120.4%, the NER at 96.6%, the CR at 94.7% and there were 4.8% repeaters. There has been a marked increase in GER and CR.

In 2016, there were 390 UBE schools, of which 41% were managed by the government, 15% were private, 36% were Madrassas, and 8% were grant-aided. There are fewer government and more Madrassa and grant-aided schools compared to the LBE level. In addition, 62% of them were basic cycle and 55% were double-shift schools. This may suggest that the need for multi-grade schools have subsided. There were 90,838 pupils enrolled, of which 36% were grade 7, 33% were grade 8, 30% were grade 9. These ratios are about the equivalent to those of LBE schools. There were 2,246 classes, of which 67% were morning classes and 33% were afternoon classes. The afternoon classes are actually commonly attended by lower grades. The GER was at 66.8% and the CR at 61.0%, which could be better. There were only 3.1% repeaters.

In 2021, there were 524 UBE schools, of which 38% were managed by the government, 16% were private, 38% were Madrassas, and 8% were grant-aided. The portion of government and Madrassa schools are roughly equal, which demonstrates the importance of Islam in The Gambia. In addition, 63% of them were basic cycle and 49% were double-shift schools. These numbers have not changed much. There were 119,405 pupils enrolled, of which 37% were grade 7, 34% were grade 8, and 29% were grade 9. There has not been a lot of changes here either. There were 3,544 classes, of which 61% were morning classes and 39% were afternoon classes. There is a slightly larger portion of afternoon classes. The GER was at 75.9% and the CR at 65.1%, which signifies that the former has improved. There were 2.6% repeaters, which is also better.

In 2016, there were 160 SSE schools, of which 33% were managed by the government, 23% were private, 23% were Madrassas, and 22% were grant-aided. The portions of different management types are surprisingly similar and there has been a conspicuous increase in grant-aided schools. In addition, 52% of them were double-shift schools. This has been consistent through all levels of education. There were 56,001 students enrolled, of which 39% were grade 10, 33% were grade 11, 27% were grade 12. Similar to the other levels, there is a higher portion of lower grades. There were 1,302 classes, of which 71% were morning classes and 29% were afternoon classes. The portion of morning classes is slightly higher in comparison with UBE, perhaps because of the significantly lower number of schools, classes, and students. The GER was at 44.0% and the CR at 36.6%, which is somewhat low. There were only 4.4% repeaters though.

In 2021, there were 213 SSE schools, of which 31% were managed by the government, 19% were private, 28% were Madrassas, and 21% were grant-aided. There has been more schools though the numbers have dramatically decreased by more than half in comparison with UBE which were less than half of that of LBE. In addition, 47% of them were double-shift schools. There has been a slight decrease here, which may indicate that demand has been better met. There were 76,537 pupils enrolled, of which 36% were grade 10, 34% were grade 11, and 30% were grade 12. The distribution of students amongst classes has evened out a little. There were 2,108 classes, of which were morning classes and 35% were afternoon classes. There are curiously enough more afternoon classes though. The GER was at 51.1% and the CR at 46.2%, both have increased. There were 3.2% repeaters, which is also a bit better.

5.1.2 Equitable Education

Second, an overall theme of “equitable access” was found which consists of the distinct themes “regional differences”, “gender differences”, and “special needs children”.

The ESSP plans to provide children in far-away rural areas with school transport to and from their nearest LBE schools. Curriculum, assessment, and teaching will be granted in the different native languages as well as in English, French, and Arabic. Governance will become more decentralized in order to empower regional, district, and local education. Support and incentives will be given to female students in science, mathematics, and technical vocational subjects at higher and tertiary institutions. Dropout of girls at will be addressed with zero tolerance policies for gender-based violence and resumption of studies after pregnancy or early marriage will be encouraged. Community participation and cooperation with idealistic organizations in gender equity initiatives will continue to be strengthened. User-friendly access to all school building for differently-abled students will be provided. They will be granted appropriate teaching/learning equipment and a range of alternative modes of education will be offered. Existing special schools will be strengthened, inclusive teacher training will be provided at all schools, and periodic surveys will be carried out to uncover accommodation needs.

In 2016, of the 1,141 ECD centers, 20% were in region 1, 34% in region 2, 13% in region 3, 7% in region 4, 13% in region 5, and 14% in region 6. This indicates a dominance of ECD centers in urban areas, but very few opportunities for preschool education in rural areas. Moreover, 96% used English, 33% used Arabic, 62% used Mandinka, 43% used Wollof, 38% used Pullar, 16% used Jolla, and 10% used other languages. It may signify that Pullar children are in need of more ECD centers that uses their native language. Of the 100,349 children enrolled, 51% were girls and 49% were boys. This signifies that gender differences are minimal. Finally, 1,132 were special needs children of which 23% have visual, 16% have hearing, 26% have speech, 21% have physical, 6% have mental, 4% have other, 4% have multiple, and 1% have unknown impairments. The last-mentioned indicates that the field of the form was left blank, which might suggest that some assistance in filling the forms or classifying the disability is needed.

In 2021, of the 1,460 ECD centers, 17% were in region 38% in region 2, 13% in region 3, 7% in region 4, 12% in region 5, and 14% in region 6. The urban and rural gap remains although there are fewer centers in region 1 and more centers in region 2. Moreover, 95% use English, 38% use Arabic, 62% use Mandinka, 42% use Wollof, 42% use Pullar, 23% use Jolla, and 13% use other languages. It seems that the number of ECD centers in Pullar has increased. Of the 126,573 children enrolled, 52% were girls and 48% were boys. Gender equality is still intact. Lastly, 1,276 were special needs children of which 27% have visual, 13% have hearing, 27% have speech, 18% have physical, 5% have mental, 3% have other, and 4% have multiple impairments. The unknown disabilities have disappeared which may suggest that assistance in filling the forms or classifying the disability has been provided.

In 2016, of the 958 LBE schools, 14% were in region 1, 30% in region 2, 16% in region 3, 8% in region 4, 15% in region 5, and 16% in region 6. There is a high number of schools in region 2, but very few available in region 4. Of the 308,729 children enrolled, 51% were girls and 49% were boys. Finally, 4,615 were special needs children of which 40% have visual, 22% have hearing, 12% have speech, 10% have physical, 5% have mental, 7% have other, and 3% have multiple impairments. Compared to ECD centers, there is a lot more children with visual impairments and a lot fewer with speech or physical impairments.

In 2021, of the 1,229 LBE schools, 14% were in region 1, 32% in region 2, 16% in region 3, 8% in region 4, 15% in region 5, and 14% in region 6. The urban and rural gap remains and may even have increased a bit. Of the 401,333 children enrolled, 52% were girls and 48%

were boys. Lastly, 5,948 were special needs children of which 38% have visual, 18% have hearing, 13% have speech, 14% have physical, 6% have mental, 5% have other, and 5% have multiple impairments. This indicates a high need for visual aids in primary education.

In 2016, of the 390 UBE schools, 20% were in region 1, 34% in region 2, 14% in region 3, 7% in region 4, 11% in region 5, and 14% in region 6. There are almost five times more schools in region 2 than in region 4. Of the 90,838 children enrolled, 51% were girls and 49% were boys. Finally, 1,660 were special needs children of which 50% have visual, 17% have hearing, 13% have speech, 9% have physical, 2% have mental, 8% have other, and 2% have multiple impairments. Compared to LBE schools, there is an even larger portion of children with visual impairments.

In 2021, of the 524 UBE schools, 17% were in region 1, 39% in region 2, 14% in region 3, 7% in region 4, 10% in region 5, and 13% in region 6. Regional differences have unfortunately increased slightly. Of the 119,405 children enrolled, 54% were girls and 46% were boys. Lastly, 2,331 were special needs children of which 49% have visual, 19% have hearing, 11% have speech, 9% have physical, 3% have mental, 5% have other, and 3% have multiple impairments. There is a higher number of special needs children, but this could be due to the higher number of enrolled children overall.

In 2016, of the 160 SSE schools, 29% were in region 1, 35% in region 2, 14% in region 3, 6% in region 4, 9% in region 5, and 8% in region 6. Compared to UBE, there are more schools in region 1 and fewer schools in region 6. Of the 56,001 children enrolled, 51% were girls and 49% were boys. Finally, 780 were special needs children of which 56% have visual, 18% have hearing, 5% have speech, 8% have physical, 1% have mental, 12% have other, and 1% have multiple impairments. There are fewer students with speech impairments compared to UBE schools.

In 2021, of the 213 SSE schools, 25% were in region 1, 40% in region 2, 13% in region 3, 5% in region 4, 8% in region 5, and 8% in region 6. There are more students in region 2 and fewer in region 1, but the urban rural divide persists. Of the 76,537 children enrolled, 55% were girls and 45% were boys. There is a slightly larger portion of girls. Lastly, 1,298 were special needs children of which 57% have visual, 16% have hearing, 7% have speech, 8% have physical, 0.3% have mental, 9% have other, and 2% have multiple impairments. There is a high need for visual aids in secondary education as well.

5.1.3 Educational Resources

Third, a general theme of “available resources” was found which consists of the specific themes “capacity”, “teachers”, “textbooks”, and “facilities”.

The ESSP intends to build, rehabilitate, and maintain more schools and classrooms with kitchens, water points, and sanitary services. Routine censuses will be conducted to identify infrastructure needs and the school maintenance policy will be enforced. Home-grown school feeding programs will be promoted, students meals will be distributed, canteens will be established, and sustainable agricultural practices encouraged. Teachers will be supported through provision of performance rewards, hardship allowances, and living quarters especially in remote areas. Recurring questionnaires will be performed to reveal teacher needs. Free textbooks will be provided to all types of schools at all levels, but there will be a transition from print to digital learning materials, mainly due to challenges in distribution and replacement. School libraries will be supported through mobilized resources and materials, staff training and management, and new operational policies. E-books, e-libraries, and e-learning programs will be provided in both LBE, UBE, SSE schools. The curriculum framework and teaching syllabus will be updated and the customization and localization of them will be encouraged.

In 2016, there were 1,141 ECD centers with 3,076 classes and 100,349 children enrolled in them. There were on average 88 children enrolled at each ECD center and 33 in each ECD class. Moreover, the child to facilitator ratio was 33 whereas the child to trained facilitator ratio was 46. This numbers are somewhat high even though most of the children are between 3 and 6 years old, as special attention should be given towards the few who are under the age of 2. There were 3,072 facilitators, of which 2,182 or 71% were trained. In addition, 1,614 or 53% were women and 1,458 or 47% were men. Further, 2,800 or 91% of all facilitators are Gambians and 272 or 9% are non-Gambians. This indicates a good balance of male and female staff, although the level of training could be higher. This is especially important for the vulnerable children that are less than 2 years old. Of the ECD centers, 54% are housed in permanent structures, 63% have safe drinking water, 36% have adequate separate toilets, 57% have playgrounds, 52% have furniture, and 27% are linked to the health care sector. These figures are somewhat low and there is definitely a need for improvement in this area.

In 2021, there were 1,460 ECD centers with 3,958 classes and 126,573 children enrolled in them. There were on average 87 children enrolled at each ECD center and 32 in each ECD class. Moreover, the child to facilitator ratio was 31 whereas the child to trained facilitator ratio was 42.This indicates that there has been almost no changes in terms of capacity although most of the children are still between the ages of 3 and 6. There were 4,046 facilitators, of which 3,018 or 75% were trained. In addition, 2,600 or 64% were women and 1,446 or 36% were men. Further, 2,876 or 94% of all facilitators are Gambians and 245 or 6% are non-Gambians. There has been a slight increase in Gambian teachers, an increase in the portion of female staff, and a slight increase in trained staff. Of the ECD centers, 65% are housed in permanent structures, 56% have safe drinking water, 37% have adequate separate toilets, 53% have playgrounds, 48% have furniture and 21% are linked to the health care sector. Although there has been improvements in permanent structures, safe drinking water, and adequate toilets, there has also been decreases in playgrounds, furniture, and health care links.

In 2016, there were 958 LBE schools with 8,586 classes and 308,729 children enrolled in them. There were on average 322 pupils enrolled at each school and 36 in each class. Moreover, there are 47 students per classroom, 1.9 per desk, and 1.8 per seat. The class sizes are on the larger side and children may need to share desks and seats. The pupil to teacher ratio was 35 whereas the pupil to qualified teacher ratio was 41. There were 8,730 teachers, of which 7,592 or 87% were qualified. In addition, 2,953 or 34% were women and 5,777 or 66% were men. Further, 8,118 or 93% of all teachers are Gambians and 612 or 7% are non-Gambians. There are somewhat more qualified teachers in comparison with ECD, though there are also a lot fewer female teachers. There are 247,881 textbooks, of which 109,864 or 44% are in English, and 138,017 or 56% are in math. Grades 3 and 4 tend to have the fewest number of textbooks. Of the LBE schools, 68% have fences, 80% have safe drinking water, 80% have adequate separate toilets, 33% have electricity, 54% have libraries, and 18% have computer labs. There is a lot more safe drinking water and adequate separate toilets compared to the ECD level, but very few schools have electricity or computer labs.

In 2021, there were 1,229 LBE schools with 12,298 classes and 401,333 pupils enrolled in them. There were on average 327 pupils enrolled at each school and 33 in each class. Moreover, there are 46 students per classroom, 2.1 per desk, and 1.9 per seat. Although students per classroom has decreased, there is an increase in students per desk. The pupil to teacher ratio was 33 whereas the pupil to qualified teacher ratio was 38. There were 12,116 teachers, of which 10,614 or 88% were qualified. In addition, 5,005 or 41% were women and 7,111 or 59% were men. Further, 11,521 or 95% of all teachers are Gambians and 595 or 5% are non-Gambians. There are a lot more teachers, both qualified and unqualified, though the ratio between them remains unchanged. Students per teacher ratio and gender equality amongst teachers has improved somewhat. There are 466,810 textbooks, of which 228,395 or 49% are in English, and 238,415 or 51% are in math. Grades 5 and 6 tend to have the fewest number of textbooks. Of the LBE schools, 66% have fences, 86% have safe drinking water, 83% have adequate separate toilets, 40% have electricity, 50% have libraries, and 22% have computer labs. There has been a decrease in libraries and an increase in electricity and computer labs, which might indicate some digitization of the education sector.

In 2016, there were 390 UBE schools with 2,246 classes and 90,838 children enrolled in them. There were on average 233 pupils enrolled at each school and 40 in each class. The pupil to teacher ratio was 23 whereas the pupil to qualified teacher ratio was 25. Moreover, there are 45 students per classroom, 1.5 per desk, and 1.5 per seat. The improved pupil to teacher ratios are better which might be indicative of more individualized learning. There were 3,901 teachers, of which 3,652 or 94% were qualified. In addition, 801 or 20% were women and 3,100 or 80% were men. Further, 3,543 or 91% of all teachers are Gambians and 358 or 9% are non-Gambians. There are many more qualified teachers though somewhat fewer female teachers at this level compared to the LBE level. There are 100,841 textbooks, of which 48,505 or 48% are in English, and 52,336 or 52% are in math. Textbooks are evenly distributed across grades. Of the UBE schools, 77% have fences, 84% have safe drinking water, 80% have adequate separate toilets, 56% have electricity, 39% have libraries, and 28% have computer labs. In comparison to LBE, there seems to be an increasing focus on computer skills.

In 2021, there were 524 UBE schools with 3,544 classes and 119,405 pupils enrolled in them. There were on average 228 pupils enrolled at each school and 34 in each class. Moreover, there are 43 students per classroom, 1.8 per desk, and 1.7 per seat. The pupil to teacher ratio was 20 whereas the pupil to qualified teacher ratio was 21. Classes have become smaller and more personalized, although there is slightly more desk and seat sharing. There were 6,083 teachers, of which 5,787 or 95% were qualified. In addition, 1,599 or 26% were women and 4,484 or 74% were men. Further, 5,710 or 94% of all teachers are Gambians and 373 or 6%

are non-Gambians. There is a very high portion of qualified teachers, but a very small portion of female teachers. There are 177,959 textbooks, of which 92,176 or 52% are in English, and 86,861 or 48% are in math. The number of textbook slightly decrease at higher grades. Of the UBE schools, 78% have fences, 90% have safe drinking water, 85% have adequate separate toilets, 58% have electricity, 61% have libraries, and 41% have computer labs. Facilities have been greatly improved in most areas.

In 2016, there were 160 SSE schools with 1,302 classes and 56,001 children enrolled in them. There were on average 350 pupils enrolled at each school and 43 in each class. Moreover, there are 39 students per classroom, 1.3 per desk, and 1.3 per seat. The pupil to teacher ratio was 24 whereas the pupil to qualified teacher ratio was 25. There has been a slight increase in overall capacity compared to UBE schools. There were 2,312 teachers, of which 2,218 or 96% were qualified. In addition, 242 or 11% were women and 2,070 or 89% were men. Further, 1,825 or 79% of all teachers are Gambians and 487 or 21% are non-Gambians. There are fewer female and Gambian teachers in comparison with UBE. Of the SSE schools, 83% have fences, 89% have safe drinking water, 77% have adequate separate toilets, 66% have electricity, 77% have libraries, and 61% have computer labs. The portion of schools with libraries has increased markedly when compared to UBE.

In 2021, there were 213 SSE schools with 2,108 classes and 76,537 pupils enrolled in them. There were on average 359 pupils enrolled at each school and 36 in each class. Moreover, there are 49 students per classroom, 1.6 per desk, and 1.4 per seat. The pupil to teacher ratio was 23 whereas the pupil to qualified teacher ratio was 24. Although there are fewer students in each class, there are more students per classroom. There were 3,326 teachers, of which 3,211 or 97% were qualified. In addition, 443 or 13% were women and 2,883 or 87% were men. Further, 2,871 or 86% of all teachers are Gambians and 455 or 14% are non-Gambians. Although there has a little bit of an increase in female teachers, it has only been declining and is at its lowest at this level. Of the SSE schools, 79% have fences, 66% have safe drinking water, 63% have adequate separate toilets, 67% have electricity, 64% have libraries, and 457% have computer labs. Most of these figures have unfortunately decreased the past few years.

5.2 Conference

The program for the DHIS2 Level 2 Academy for Education 2022 consisted of presentations, panel discussions, roundtables, and workshops. There were technical sessions and strategic sessions that were held in parallel on day 3 and 4 of the conference. Ministries of education and health, non-governmental organizations, and other relevant stakeholders from mainly African countries were invited and participated. First, an overall theme of "organizational aspects" was found which consists of the distinct themes "strategy", "systems" and "data". Second, a general theme of "technical aspects" was found which consists of the specific themes "features" and "implementation". These themes apply to any management information system in any country. However, the two last-mentioned themes will be specifically about DHIS2 in The Gambia.

5.2.1 Strategy

The digitalization of education in low-resource contexts requires a carefully considered strategy in order to maximize opportunities and minimize difficulties. There are limited resources in developing countries, and governments must ensure that they are used efficiently. This requires learning from, comparing to, and cooperating with other sectors such as health when it comes to digitization knowledge transfer, resource sharing, and data exchange. As a panelist pointed out, there is a Gambian saying that "if you want to go really fast, you need to go alone, but if you want to go really far, you need to go with other people". In order for this to work, all relevant stakeholders at all levels must be involved. This includes technical experts, domain specialists, and government administrators at both state, region, and district levels. Management of the education sector should be more decentralized so that schools and teachers are not only obligated to collect data, but also included in data analysis and involved in legislations. Moreover, it is not enough to simply introduce the technology – capacity building have to take place in the form of training. The staff needs to be equipped with not only basic digital literacy but also advanced computer skills. Fundamental national standards should be developed and existing international standards should be adopted to ensure best practices and easy comparisons. Finally, there is a need for assessment and documentation of the efficacy of different projects and programs. Measurement of key performance statistics both pre- and post-intervention is essential in order to be able to establish a baseline and infer causation.

5.2.2 Systems

Due to a high number of silo systems, the landscape of management information systems is very fragmented in all sectors, and there is a pressing need to integrate them to avoid duplication. The introduction of any new system needs to be adequately justified; it has to add something of value that did not exist before. As a minister of education said, “people come and go, but systems remain”. Flexibility is important, as the systems need to be modular enough to be able to adapt to changing requirements. Another issue is sustainability – there is often a desire to have real-time systems tough it may not always be necessary. It is also important to keep in mind the demands of such a system in terms of time, money, and connectivity. There is a considerable need for student identification systems, daily attendance registration, and geographic information systems. Each school tends to have their own identification system, making it difficult to track the transfer history of a student and identify at-risk children in order to plan interventions for them. Similarly, absenteeism is often an indicator of dropout and attendance registration is imperative in the prevention of this. Due to low wages and economic incentives for private tutoring, some teachers will not show up to class and there is a need for an overview of these. Finally, having a common cross-sector geographic information system in place can help greatly in finding out where to build schools based on population density and nearby hospitals, markets, and sanitation.

5.2.3 Data

In order for decision-making to be effective, it has to be evidence-based and data-driven. As a donor partner said, “making decisions without data is like walking with a blindfold – you might move forward, but it might be painful”. However, it is not enough only to gather the data, they also need to be processed, analyzed, and interpreted by people who are qualified to do so. Oftentimes, schools will receive grants based on the performance outputs that they report. There is a need to make sure that this data is in fact correct and this can only be done through triangulation. In addition, the data needs to be accurate, complete, consistent, valid, unique, and timely in order to ensure a high quality. There is a distinction between aggregate data such as how many students are present, and tracker data such as which students are present and many governments want to move towards the latter. Moreover, raw data like counts must be differentiated from analyzed data like ratios; in order for the last-mentioned to be calculated, the first-mentioned must have a fine level of granularity. However, this also means that more time needs to be spent on reporting data. In traditional paper-based systems, the daunting logistics of distributing empty forms and gathering filled-out surveys means that data collection mainly occur in annual censuses. By the time the data is centralized it has often become outdated and it is hence difficult to assess the effect of policies because of this delay. An EMIS can solve both of these problems and enable the gathering of data in both large quantities and high frequencies.

5.2.4 Features

In order to address all of these needs, an EMIS has to come with various features. DHIS2 implements various requirements that ministries of education have expressed. It can be used to both view and enter aggregate data as well as tracker data. The former is registered in the "data entry" app which comes with data quality checks for completeness, correctness, and integrity as well as an audit trail. The latter is registered in the "capture" app which comes with bulk action capabilities, possible duplication detection, and program rules execution. In other words, both apps come with data quality assurance tools that allows the user to create customized validation rules, such as "regional enrollment cannot be larger than national enrollment". Data in DHIS2 can be visualized in dashboards which can be created, edited, deleted, bookmarked, shared, opened, and closed. When shared, the user can choose between viewing rights and editing rights. When closed, a list of the most frequently used dashboards will be shown. Different dashboards can be used to display data on intake, enrollment, attendance, teachers and much more. Data can be displayed as charts using the "data visualizer" app, as tables using the "pivot tables" app, or as maps using the "maps" app. Amongst the available charts are column, stacked column, bar, stacked bar, line, area, stacked area, scatter, pie, and gauge. When column, bar, scatter dot, or pie slice is moused over, it will briefly be highlighted and a pop-up with additional information such as an in-depth explanations can be shown. Tables can be made to show cells in different colors depending on custom rules, for example, data values for absenteeism can be green if less than ten percent, orange if less than 30 percent, and red if above that. Moreover, all columns are sortable and the tables can be flipped. Maps backgrounds can be shown in various types of drawings, satellite images, and mixes between the two. Map foregrounds consists of overlays that can show regional and district borders filled with different colors. Multiple overlays can be stacked and the opacity of both the backgrounds, borders, and colors can be set. Furthermore, the same data can be viewed as both a chart, table, and map and conversion between them is both quick and simple. All the charts, tables, and maps of the dashboard can be resized and rearranged as the user likes. In addition, the zoom level and position in maps and similar preferences are remembered. Data can be filtered by periods (both relative and fixed) and area (both regions and districts). Also, it can be interpreted and discussed through comments that are shared between users who can tag each other and receive notifications. Finally, it can be downloaded as a PNG, PDF, JSON, XML, XLSX, CSV, and raw data SQL and the two former can be printed and shared with schools without electricity or internet connectivity.

5.2.5 Implementation

DHIS2 has been implemented as a part of the EMIS data cycle of The Gambia. Normally, paper-based forms are distributed, manually filled with aggregate data, and gathered for centralized data entry and analysis in excel. However, education policies and the fourth SDG increasingly stress the importance of equitable access for vulnerable children which requires tracker data. Because of this, MoBSE wants to decentralize data collection and management to the school level. Cooperation with the HISP center started in January 2019 when a Memorandum of Understanding was established. The DHIS2 core team then started to look at the integration of EMIS forms in March and a Learner Tracker Pilot was launched in August. Although in March 2020, the project was put on hold due to COVID-19. It resumed in January 2021 when the Learner Tracker was scaled all the way up to the national level. Before, daily attendance was registered in a system from 2009 using SMS messages that had to strictly adhere to an inconvenient configuration. Now, SMS messages are still being used but they are automatically formatted by the DHIS2 android app. Moreover, schools are now able to see the data that they and others have reported on student and teacher attendance over the past year and a half.

5.3 Interviews

Directors from the STED and the CREDD as well as a Leap Learning manager were interviewed. Although there is good reason to believe that Leap Learning apps can lead to new pedagogical possibilities, the goal of these interviews was to find out if they can have administrative applications as well.

5.3.1 Interview with the Directorate of Science and Technology Education

First, a general theme of “learning apps” emerged which consists of the specific themes “Leap Learning” and “Play to Learn”.

There are currently 18 schools or three in each region that have Leap Learning labs, and the STED is looking to implement in 4 more schools which would bring the total up to 22. There is usually only one lab in one classroom for either literacy or numeracy at the school or two labs in two classrooms with both literacy and numeracy. Different classes use the labs and classrooms at different times, and one class can have anywhere from 25 to 40 students though the average class has roughly 35 students. The schools are bigger and have more students in the cities and they are smaller and have fewer students in villages. So, a school in region 1 might have 6 classes at grade 1, whereas a school in region 5 might have 2 classes at grade 1. All of the schools have the physical learning stations as well as the preloaded and secured tablets, although the number of tablets is capped at 25 per classroom. One class uses the lab around 2-3 times a week for a double period of 70 minutes with a break in between or a single period of 35 minutes without any break. There is usually only one teacher in each lab session and their job is to try and connect what is happening in the classroom with what is happening in the learning apps. The children are free to move around as they like and find a station that is appropriate to their level while the teacher tries to intervene as little as possible. The Leap Learning labs fall under the jurisdiction of the STED. In addition, each region has an IT focal point that can be contacted if there are any issues with the apps or the tablets.

Only one other similar project involving learning apps on tablets has been initiated in The Gambia before. It was a system called Play To Learn which was implemented at one school in every region for a total of 6. Each school was given 40 tablets preloaded with grade 3 and 4 national math curriculum. The app was a game with a map of the county, in which the user starts at region 6 and works his/her way to region 1. There were maybe four topics (e.g., addition, subtraction, etc.) with three levels each that had to be successfully completed in a consecutive order. The pilot was a success, only a few tablets were stolen or broken, so the system was expanded to 18 schools. However, the tablets were of low quality – they were slow and froze up, the screens were easily scratched, and there were blotches on the display. Moreover, it had two chargers: one pin and one micro-USB. Despite that, the charging ports would often be ruined and then the tablets would run out of power and be unusable. On the other hand, the Leap Learning tablets come locked and secured in wooden boxes and hence will not break easily. Although it does not follow the national curriculum, it has more apps, topics, and games with numerous levels that can be used in multiple grades. A unique feature is its combination with physical learning stations and its emphasis on analytical thinking and critical reflection.

5.3.2 Interview with Leap Learning Management

Second, an overall theme of “management” appeared which is made up of the distinct themes “cooperation”, “training”, and “challenges”.

All of the Leap Learning labs are administered from the Africa Startup MyFarm center in The Gambia. In addition to the 18 labs associated with MoBSE, there are many more labs in private schools, totaling more than 40 schools. The cooperation between the two stakeholders has only recently begun, though they are looking to eventually associate all schools with the STED. Every quarter, LL checks in on all of the labs, aiming to spend one day in every region. When a new school enters the program, they download all of the apps on one device and then transfer them to the others as needed. In the first session, only the most basic apps and stations are in use. After that, the selection is slowly expanded until everything is covered. The schools tend to be a bit apprehensive towards the labs in the beginning, though it does not take long before there is a big shift in their attitudes. It seems like the children have fun and enjoy learning by using the LL labs. The labs are meant to build up and not replace the existing educational system.

There is a training guide booklet for the labs and even a learning app on how to use them. Although only a few teachers at each school facilitate the labs, all teachers are trained to use them. This is in the case of a teacher being sick, another one can take over the facilitation. Training is administered by the Teacher Training Committee. Recently, LL has opened up their platform to the public and because it is low-code/no-code, anyone can create apps on it without any programming experience. All schools have their own WhatsApp group with LL where they can receive technical and pedagogical support for the apps. They can make a video call in order to get feedback on how well they are using the lab, for instance.

Some of the challenges are that the students tend to focus too much on the tablets. Sometimes components are missing from the learning stations, posters are not in the right position, cards are not in the right order, or the launcher app is not downloaded. Any of these can have an adverse effect on the learning experience. Another issue is that the children take too many photos with the tablets that occupy storage space. Finally, each school has a unique license, but individual level data capture is not possible. This is partly because many schools do not have stable electricity and internet, so data is only reported when LL visits the school and brings electrical power and network connectivity to them. This was confirmed on a field trip school visit to Bakoteh Lower Basic School where there had been a power outage for three days and they were unable to charge their tablets. Unfortunately, we visited on the 26th of April when the third school term had just started and we were thus unable to conduct class observations due to a lack of available students and teachers.

5.3.3 Interview with the Directorate of Curriculum Research Evaluation and Development

Third, a broad theme of "considerations" surfaced which is comprised of "evaluation" and "localization".

The first LL literacy lab was introduced in 2007 and the first numeracy lab in 2012. In Somalia, there was a pilot in which they compared a school with a lab to a school without a lab. They found that LL led to a big difference in performance. Similarly, in Ghana the tablets were even used at home and parents even wanted to buy them. The first LL pilot in Gambia was performed at two schools in region 2, as it is a very typical use case scenario. Due to the small scale, they would be able to quickly identify any problems and properly respond to them before expanding it to the national level. After an initial period of skepticism, the tablets generated a lot of engagement, in students as well as teachers. The conclusion of the pilot was quite emotional as the children would cry and not let go of the tablets. Similar success was seen with Play to Learn which was introduced in collaboration with the World Bank. With it, school performance in region 6 soared and they even exceeded that of region 1 on NSTs like EGMA and NAT. The next logical step will be to systematically compare a treatment group and a control group that are similar in which one has LL and the other does not. A baseline should be established after which changes in EGRA and EGMA performance should be measured. Currently, both students and teachers have expressed some concern about LL taking time away from the curriculum. They question how important or relevant it is when there is no evidence of increase in entrance exams performance. Because of this, the CREDD wants data on anything that can be an indication of learning such as number of successful attempts, number of failed attempts, time spent on task, and so on. When it comes to tracker data, only the app books that are printed out and filled in to track learner progress is capable of providing it, though the data quality is questionable. They are not always filled out or not always filled out correctly, and it is impossible to confirm or disconfirm them. Scanning the app books would be too labor intensive and cannot realistically be automated. Login would not be possible to implement due to intermittent connectivity and limited storage. Another concern is if the students would be able to log in to the right account, remember to log out, refrain from sharing account info, and so on.

MoBSE plans to increase use of ICT and the Progressive Science Initiative/Progressive Mathematics Initiative has been administered at the UBE and SSE levels for this. They want to increase computer literacy and technical capacity in both students and teachers, especially in rural areas. There are IT focal points in each region that the schools can contact for assistance and the government has even sponsored motor cycles for them so that they can easily travel around. However, a problem is donor partners who often have strict requirements on how to collect data, what data to collect, and which tools to use. Some are more preoccupied with branding than on for instance improving literacy or strengthening the native languages. The CREDD is very concerned with reading skills, as the EGRA toolkit shows that it is the most important prerequisite for learning which mostly occurs in text. The longer it takes for a child to acquire proficiency, the more difficult it becomes to achieve it. Moreover, the difference between more and less literate children will increase dramatically over time. In order to use global indicators and compare performances with other countries, the EGRA was administered in The Gambia in 2016 and 2022. However, there are 7 national languages, one for each ethnic group. For example, although Mandinka and Wollof are the most spoken languages in regions 1 and 2, Fulani is the most common language in region 6. Many students are most comfortable in their native tongue, and this has an adverse effect on learning that mainly occurs in English which they are not as fluent in. The education sector has a policy that aims to translate and digitize the whole curriculum in accommodation for distance learning. They are currently in a dialogue with the World Bank about funding for this. This problem also extends to teachers who often speak two or three languages in addition to English, French, or Arabic. They might be most comfortable teaching in their main language, but they are often rotated between all schools because of staff shortages. In the future, MoBSE wants to only transmit teachers to the schools that use the same language as them. In other words, there is a need to map languages spoken by the students with languages spoken by the teachers in order to amplify learning outcomes.

Chapter 6

Discussion

6.1 Representation Theory

Weber's RT can be applied to both LL and DHIS2. The former would represent the state of literacy and numeracy in children, whereas the latter would represent the existing paper-based forms of the Gambian EMIS data cycle.

The surface structure of LL would be the images, sounds, and buttons of the apps; the physical structure would be the rugged tablets in wooden boxes; and the deep structure would be the literacy and numeracy tasks. The representational model would be how well the apps represents the speaking, listening, reading, writing, addition, subtraction, multiplication and division skills of the child. The state-tracking model would be how well it catches improvements, stagnation, and deterioration in these skills of the child. The good-decomposition would be how well these skills are decomposed into tasks, levels, and games within the apps. Further, in terms of the interpretive structure introduced by Fossum and Kempton (2020), the correctness dimension would be the extent of causality between acquiring literacy or numeracy skills and increased levels in the apps. The subjectivity dimension could be evaluation that performance on a literacy skill component is lower than those of other children in the same grade. The situation dimension could be an indication of learning disability or other form of special need. In addition, when it comes to the effective use model by Burton-Jones and Grange (2013), the transparent interaction could be ease of which the tablets and apps can be used. The representational fidelity could be the pedagogical validity of the apps. The informed actions could be the effectiveness of the apps in improving literacy and numeracy. The user friendliness of the apps increase the pedagogical validity of the apps which improves literacy and numeracy skills.

The surface structure of DHIS2 would mainly be its interactable dashboards, charts, tables, and maps, but maybe also the "data entry", "capture", and "data visualization" apps amongst other things depending on the user. The physical structure would be the computers on which it is accessed, the smartphones in which data is entered, as well as any servers storing the data and so on. The deep structure would be how well its data represents the educational statistics it represents. The representational model would be how well its data objects represent concepts like student enrollment, teacher absence, school facilities and so on. The state-tracking model would be how well it captures changes in the education sector. The good-decomposition model would be the way in which categories formed. Additionally, in relation to the interpretive structure introduced by Fossum and Kempton (2020), the correctness dimension could be the extent of causality between updates in educational statics and data objects. The subjectivity dimension could be a remark that cases of dropout are higher than normal. The situation dimension could be economic difficulties. Moreover, with reference to the effective use model by Burton-Jones and Grange (2013), transparent interaction could be the ease of data entry through SMS. Representation fidelity could be the accuracy of student attendance counts. Informed actions could be the role of the dashboard in advising legislation. The ease of data entry through SMS increase the accuracy of student attendance counts which informs policymaking.

6.2 Technology-Enhanced Learning

Both DHIS2 and LL can be viewed upon as TEL as both lead to the enhancement, streamlining, and transformation of learning using technology. The former can be used as a LMS whereas the latter can be viewed upon as a learning management content system (LMCS). Whereas the first-mentioned specializes in administration and include features such as registration systems, performance assessment, and human resources, the last-mentioned focuses solely on the content (Sejzi & Aris, 2013).

Currently, these systems are either not evaluated or not done so in a formal and systematic way. Although it might be tempting to resort to the usual questionnaire, it is worth considering design-based research approaches. As Mor & Winters (2007) and Wang & Hannafin (2005) point out, they have the potential to reveal rich and nuanced data with a lot of depth in a way that traditional approaches cannot do. There is a lot of merit in the principles and values that it is based on such as focusing on the end user (the students) and other relevant stakeholders that might often be overlooked (the teachers). This stands in contrast to current TEL research which has a tendency to focus on researchers and administrators. On the one hand, design-based research can be impractical in a low-resource setting and it is understandable if it is not a top priority of ministers of education in LMICs.

On the other hand, it can be more resource-effective than resource-intensive and it can contribute to design and research at the same time and provide insight into vulnerable student groups.

Another interesting topic is the use of TaD of TEL, as promoted by Kali et al. (2015) and dissected by Kirschner (2015). As previously mentioned, they have received relatively little attention considering that learning occurs through the relationship between the teacher and the student. The last-mentioned is another important stakeholder group that are rarely consulted and are often represented in literature through intermediaries aside from evaluations of the occasional TEL tool. On this note, Kirkwood and Price (2014) attempt to define enhancement of learning through technology by reviewing existing TEL intervention studies. They found that most of them either end up replicating or supplementing traditional teaching practices, whereas only a few attempt to transform them. The two former tend to define enhancement as quantitative changes in performance, whereas the latter tends to describe it as qualitative change in understanding. While it can be argued that both DHIS2 and LL are attempting to transform education, they both define enhancement in terms of numbers and not experiences, narratives, or discourse. It is important that the effects of LL and DHIS2 on learning and education are not only measured in terms of statistical outputs but also in lives touched. This can be done using design-based research approaches.

6.3 Educational Data Mining

EDM is the application of KDD on any large set of data characterized by volume, variety, velocity, veracity and value. Whether data from LL apps alone can be considered to be big data or not is up for debate. However, there is no doubt that it can fulfill these criteria when they are combined with EMIS data and integrated into DHIS2.

On the one hand, EMIS data include multiple aspects of education, but only in one country such as annual school censuses, population census, school surveys, household surveys, and cluster monitor reports. On the other hand, LL data is present in multiple countries, but focus solely on the teaching aspect of education such as Sierra Leone, Liberia, Ghana, Nigeria and many other low-resource contexts. They are both voluminous and varied, though they suffer a bit when it comes to velocity and veracity. This is mainly due to poor infrastructure such as electricity shortages or network disconnection and a lack of capacity building such as training teaching staff to report the data accurately and finding data analysts who can use the data correctly. However, there is no doubt that there is value in EMIS and LL data which can be used to inform data-driven policy-making and evidence-based decision-taking. While LL can be viewed upon as microlevel clickstream data, EMIS can be viewed upon as macrolevel institutional data. Although DHIS2 mainly focuses on simple data analytics and visualization as opposed to DM and ML, its ability to gather, integrate, and standardize data from different sources makes it invaluable. Although it may not be suitable for the DM stage of KDD, it can certainly be used in the preprocessing stage or in the interpretation & evaluation stage.

There are some challenges when it comes to accessing, analyzing, and using LL and EMIS data in DHIS2. First, there are possibilities for using multiple authentication methods and creating intricate authorization hierarchies in DHIS2. The high level of customization makes the system very versatile and suitable for a range of different use cases depending on the needs of the ministries. However, administrators at the national level should be encouraged to share data with regional, district, and school managers. As they are the ones collecting the data and the system would not be possible without them, they should at least get to view their own data and ideally be heard in the decision making process. Second, some users of DHIS2 may lack the data analytics skills that are required to analyze and evaluate the data correctly. Confounding variables, statistical noise, disruptive outliers makes it difficult to ensure that there are no misinterpretations and that policymaking remains effective. In addition, as there is a risk of falsely reported data based on self-serving motives. Therefore, it is important that robust data triangulation and validation rules are in place. Finally, striking the right balance between using the available data to improve learning outcomes versus not using it in order to protect the privacy of the students can be difficult. Most countries do not have a properly enforced privacy legislation in place, many students do not really have any choice as to whether their data is collected, and if the data is not properly anonymized it can be used for discriminatory purposes. Because of this, it is important to make sure that vulnerable groups like girls, ethnic minorities, and special needs children are protected.

6.4 Learning Analytics

Leap Learning, EMIS, and DHIS2 can be used for learning analytics as they include its six critical dimensions of stakeholders, objectives, data, method, constraints, and competencies.

First, relevant stakeholders would be students, teachers, principals, and managers who can all both data providers and data receivers. Although an interesting question would be to explore perceptions of who the main beneficiaries are. Second, objectives include both self-evaluation and self-regulation. The former comes in the form of NSTs for students, job performance for teachers, school improvements for the principals, and policy effectiveness for managers. The latter comes in the form of increased efforts for students, feedback for teachers, financial administration for principals, and funding directions for managers. Third, the data is both open and flowing as a result of integration and standardization as well as protected and secured due to data warehousing capabilities. Fourth, data can be exported to spreadsheets like excel and then to programs like SPSS or R for the application of various statistical, EDM, and LA methods. Firth, constraints are enforced through terms of use and privacy policies. Sixth, competencies are enhanced through detailed documentation, developer resources, community of practices, and open academies.

Finally, DHIS2 implementations should follow the DELICATE checklist of Drachsler and Greller (2016). First, the data collection goals should be determined whether they are mainly for administrators at the top of the educational system or the students and teachers at the bottom of the hierarchy. Second, the data collection process should be transparent and properly explained to all relevant stakeholders, including the children who might not fully understand what is going on. Third, data collectors should be legitimized by justifying the collection process and ensuring that the pros outweigh the cons. Fourth, all relevant stakeholders should be involved so that different opinions and perspectives are considered. Fifth, informed consent should be obtained before the data collection and be clear and understandable. Sixth, anonymization of individuals should be ensured though by using general demographic characteristics if not then equivalent dummy data should be used instead. Seventh, technical procedures for privacy should be performed such as secure data storage. Eighth, responsibilities should be delegated so that the different stakeholders can be held accountable.

Chapter 7

Conclusion

While the introduction of technology in education brings with it a lot of promises of inclusion and efficiency, it also comes with a lot of privacy concerns, data storage problems, and increased social inequality. This is especially the case in low-resource contexts like developing countries. For instance, while the Gambian ministry of education has robust data collection practices in place and have made enormous strides in providing basic primary education for everyone, it lacks policies for the protection of student data, the infrastructure needed for scaling, and large regional differences.

Sometimes, starting with a something simple like learning apps can help greatly in providing literacy and numeracy skills for children. Pre- and post-intervention measures of on national tests often show that they can make a big difference in performance. Sometimes, they are even able to overcome large social challenges like the urban rural divide. Although not very useful for providing student identification, higher enrollment rates, and improved school facilities, they address the core objective of all education – which is learning.

References

Access Gambia. (n.d.-a). *Gambia government information*. https://www.accessgambia.com/information/government.html

Access Gambia. (n.d.-b). *History of The Gambia*. http://www.accessgambia.com/information/history.html

Adu-Gyamfi, E., Nielsen, P., & Sæbø, J. I. (2019, November 12-13). *The dynamics of a global health information systems research and implementation project* [Paper presentation]. Proceedings of the 17th Scandinavian Conference on Health Informatics, Oslo, Norway.

Africa Startup. (n.d.-a). *Africa Startup - Improving livelihoods through innovation*. https://africastartup.org/

Africa Startup. (n.d.-b). *MyFarm – An educational journey from seed to business*. https://africastartup.org/myfarm-2/

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research*, *25*(3), 443-448. https://doi.org/10.1287/isre.2014.0546

Bayne, S. (2015). What's the matter with 'technology-enhanced learning'? *Learning, Media and Technology*, *40*(1), 5-20. https://doi.org/10.1080/17439884.2014.915851

Britannica. (n.d.). *Dyola*. https://www.britannica.com/topic/Dyola

Britannica. (2012, August 31). *Soninke*. https://www.britannica.com/topic/Soninke

Britannica. (2017a, May 8). *Gambia River*. https://www.britannica.com/place/Gambia-River

Britannica. (2017b, December 10). *Malinke*. https://www.britannica.com/topic/Malinke

Britannica. (2017c, December 28). *Wolof*. https://www.britannica.com/topic/Wolof

Britannica. (2020-a, September 22). *Madrasah*. https://www.britannica.com/topic/madrasah

Britannica. (2021, February 15). *Fulani*. https://www.britannica.com/topic/Fulani

Britannica. (2022, January 22). *The Gambia*. https://www.britannica.com/place/The-Gambia

Burton-Jones, A., & Grange, C. (2013). From use to effective use: A representation theory perspective. *Information Systems Research, 24*(3), 632-658. https://doi.org/10.1287/isre.1120.0444

Burton-Jones, A., Recker, J., Indulska, M., Green, P., Weber, P. (2017). Assessing representation theory with a framework for pursuing success and failure. *MIS Quarterly, 41*(4), 1307-1333. https://doi.org/10.25300/MISQ/2017/41.4.13

Central Intelligence Agency World Factbook. (2022, March 3). *The Gambia*. https://www.cia.gov/the-world-factbook/countries/gambia-the/

Cooper, J. (2019). *DHIS2 design system*. https://github.com/dhis2/design-system/blob/master/readme.md

Country Reports. (n.d.-a). *Gambia, the geography*. https://www.countryreports.org/country/thegambia/geography.htm

Country Reports. (n.d.-b). *Gambia, the economy*. https://www.countryreports.org/country/thegambia/economy.htm

de Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: A research agenda. *Journal of Information Technology, 33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3

District Health Information Systems 2. (n.d.-a). *About DHIS2*. https://dhis2.org/about/

District Health Information Systems 2. (n.d.-b). *Android Capture App V2.5 Overview*. https://dhis2.org/android/version-2-5/

District Health Information Systems 2. (n.d.-c). *Developer resources*. https://dhis2.org/development/

District Health Information Systems 2. (n.d.-d). *DHIS2 for education*. https://dhis2.org/education/

District Health Information Systems 2. (n.d.-e). *DHIS2 news: DHIS2 featured by the Digital Public Goods Alliance*. https://dhis2.org/dpga-digital-public-good/

District Health Information Systems 2. (n.d.-f). *DHIS2 Overview*. https://dhis2.org/overview/

District Health Information Systems 2. (n.d.-g). *Global HISP network*. https://dhis2.org/hisp-network/

District Health Information Systems 2. (n.d.-h). *DHIS2 in action*. https://dhis2.org/in-action/

District Health Information Systems 2. (n.d.-i). *Level 2 academies*. https://dhis2.org/academy/level-2/

District Health Information Systems 2. (n.d.-j). *Technology platform*. https://dhis2.org/technology/

District Health Information Systems 2. (n.d.-k). *WHO Health Data Toolkit*. https://dhis2.org/who/

District Health Information Systems 2 Documentation. (n.d.-a). *What is DHIS2?* https://docs.dhis2.org/en/use/what-is-dhis2.html

District Health Information Systems 2 Documentation. (n.d.-b). *DHIS2 as data warehouse*. https://docs.dhis2.org/en/implement/implementing-dhis2/dhis2-as-data-warehouse.html

District Health Information Systems 2 Documentation. (n.d.-e). *Developing with the Android SDK: Overview*. https://docs.dhis2.org/en/develop/developing-with-the-android-sdk/overview.html

District Health Information Systems 2 Documentation. (n.d.-f). *Using the API: Introduction*. https://docs.dhis2.org/en/develop/using-the-api/dhis-core-version-237/introduction.html

District Health Information Systems 2 Developers. (n.d.-c). *Getting started with DHIS2*. https://developers.dhis2.org/docs/

District Health Information Systems 2 Developers. (n.d.-d). *DHIS2 UI library*. https://developers.dhis2.org/docs/tutorials/ui-library/

Docker. (n.d.). *What is a container?* https://www.docker.com/resources/what-container/

Docker Documentation. (n.d.). *Docker overview*. https://docs.docker.com/get-started/overview/

Drachsler, H., & Greller, W. (2012). The pulse of learning analytics understandings and expectations from the stakeholders. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Canada*, 120-129. https://doi.org/10.1145/2330601.2330634

Drachsler, H., & Greller, W. (2016). Privacy and analytics – It's a DELICATE issue: A checklist for trusted learning analytics. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge, USA*, 89-98. https://doi.org/10.1145/2883851.2883893

Dror, I. E. (2008). Technology enhanced learning: The good, the bad, and the ugly. *Pragmatics & Cognition*, *16*(2), 215-223. https://doi.org/10.1075/pc.16.2.02dro

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining big data in education: Affordance and challenges. *Review of Research in Education*, *44*(1), 130-160. https://doi.org/10.3102/0091732X20903304

Fossum, K., & Kempton, A. M. (2020, June 15-17). *Faithfulness in practice: Towards an expansion of representation theory* [Paper presentation]. 28th European Conference on Information Systems, online event.

Friedman, J. H. (1998). Data mining and statistics: What's the connection? *Proceedings of the 29th Symposium on the Interface Between Computing Science and Statistics, USA*, *29*(1), 3-9.

Gašević, D. (2018). Include us all! Directions for adoption of learning analytics in the Global South. *Learning Analytics for the Global South*, 1-22. https://core.ac.uk/download/pdf/479891415.pdf

Geology. (n.d.). *Africa map and satellite image*. https://geology.com/world/africa-satellite-image.shtml

Geology. (n.d.). *The Gambia map and satellite image*. https://geology.com/world/the-gambia-satellite-image.shtml

Gulati, S. (2008). Technology-enhanced learning in developing nations: A review. *International Review of Research in Open and Distance Learning*, *9*(1), 1-16. https://doi.org/10.19173/irrodl.v9i1.477

Hasselbring, W. (2000). Information system integration. *Communications of the ACM*, *43*(6), 32–38. https://doi.org/10.1145/336460.336472

Health Information Systems Programme Centre. (n.d.-a). *Facts and figures*. https://www.mn.uio.no/hisp/english/about/facts/index.html

Health Information Systems Programme Centre. (n.d.-b). *HISP strategy*. https://www.mn.uio.no/hisp/english/about/strategy/

Hydro. (2022, February 28). *Key facts*. https://www.hydro.com/en-GB/about-hydro/this-is-hydro/facts/

Innovation Norway. (2020, February 3). *Our mission*. https://www.innovasjonnorge.no/en/start-page/about/our-mission/

Integrated Human Resource Information System. (n.d.). *iHRIS health workforce information systems software*. https://www.ihris.org/about

Joubert, A., Murawski, M., & Bick, M. (2021) Measuring the big data readiness of developing countries – Index development and its application to Africa. *Information Systems Frontiers*, 1-24. https://doi.org/10.1007/s10796-021-10109-9

Kali, Y., McKenney, S., & Sagy, O. (2015). Teachers as designers of technology enhanced learning. *Instructional Science*, *43*(2), 173-179. https://doi.org/10.1007/s11251-014-9343-4

Kirkwood, A., & Price, L. (2014). Technology-enhanced learning and teaching in higher education: What is 'enhanced' and how do we know? A critical literature review. *Learning, Media and Technology*, *39*(1), 6-36. https://doi.org/10.1080/17439884.2013.770404

Kirschner, P. A. (2015). Do we need teachers as designers of technology enhanced learning? *Instructional Science*, *43*(2), 309-322. https://doi.org/10.1007/s11251-015-9346-9

Leap Learning. (2018, November 15). *Leap Learning Hotspot* [Video]. YouTube. https://www.youtube.com/watch?v=ViIiLlPmoEY

Leap Learning. (n.d.-a). *Home*. https://leaplearning.no/home

Leap Learning. (n.d.-b). *Leap Learning Labs*. https://leaplearning.no/learning-labs

Leap Learning. (n.d.-c). *Our partners*. https://leaplearning.no/partners-1

Leap Learning App Universe. (n.d.). *Leap Learning App Universe*. https://webapps.leaplearning.no/en/num

Li, M. (2021, October 21). *Platform ecosystems fundamental concepts* [PDF slides]. Department of Informatics, University of Oslo. https://www.uio.no/studier/emner/matnat/ifi/IN5320/h21/timeplan/in5320-21-3-platform_fundamental_concepts.pdf

Mannila, H. (1996). Data mining: Machine learning, statistics, and databases. *Proceedings of 8th International Conference on Scientific and Statistical Database Management, Sweden*, 2-9. https://doi.org/10.1109/SSDM.1996.505910

Martins, J., Branco, F., Gonçalves, R., Au-Yong-Oliveira, M., Oliveira, T., Naranjo-Zolotov, M., & Cruz-Jesus, F. (2019). Assessing the success behind the use of education management information systems in higher education. *Telematics and Informatics*, *38*, 182–193. https://doi.org/10.1016/j.tele.2018.10.001

Ministry of Basic and Secondary Education. (n.d.). *Data sources*. http://www.edugambia.gm/news-even/data-sources

Ministry of Basic and Secondary Education. (2021). *Year book 2021*. http://www.edugambia.gm/data-area/publications/year-book-2021.html

Mor, Y., & Winters, N. (2007). Design approaches in technology-enhanced learning. *Interactive Learning Environments, 15*(1), 61-75. https://doi.org/10.1080/10494820601044236

Nations Encyclopedia. (n.d.). *The Gambia*. https://www.nationsencyclopedia.com/geography/Congo-Democratic-Republic-of-the-to-India/The-Gambia.html

Ongsulee, P. (2017, November 22-24). *Artificial Intelligence, machine learning and deep learning* [Paper presentation]. 15th International Conference on ICT and Knowledge Engineering, Bangkok, Thailand.

Open Medical Record System. (n.d.). *OpenMRS medical record system*. https://openmrs.org/

Oppvekstportalen. (2018, March 7). *Leap Learning*. https://oppvekstportalen.no/leap-learning/

Pardos, Z. (2017). Big data in education and the models that love them. *Current Opinion in Behavioral Sciences*, *18*, 107-113. http://dx.doi.org/10.1016/j.cobeha.2017.11.006

Prinsloo, P., & Kaliisa, R. (2021). Data privacy on the African continent: Opportunities, challenges and implications for learning analytics. *British Journal of Educational Technology*, 1-20. https://doi.org/10.1111/bjet.13226

Romero, C., & Ventura, S. (2013). Data mining in education. *WIREs Data Mining and Knowledge Discovery, 3*(1), 12-27. https://doi.org/10.1002/widm.1075

Sanner, T. (2021-a, March 1). *Health system, health management information system (HMIS), and integrated health information architecture (IHIA)* [PDF slides]. Department of Informatics, University of Oslo. https://www.uio.no/studier/emner/matnat/ifi/IN5330/v21/lecture-notes/2_hmis_ihia.pdf

Sanner, T. (2021-b, March 10). *Health information architectures: Integration, interoperability and standards* [PDF slides]. Department of Informatics, University of Oslo. https://www.uio.no/studier/emner/matnat/ifi/IN5330/v21/lecture-notes/integration.pdf

Sanner, T. (2021-c, September 21). *Understanding digital platforms: Key concepts and definitions* [PDF slides]. Department of Informatics, University of Oslo. https://www.uio.no/studier/emner/matnat/ifi/IN5210/h21/presentations/platform_lectures.pdf

Save the Children. (n.d.). *Annual report 2020*. https://www.savethechildren.org.uk/about-us/who-we-are/annual-report

Sejzi, A. A., Aris, B. (2013, October 7-10). *Learning management system (LMS) and learning content management system (LMCS) at virtual university* [Paper presentation]. 2nd International Seminar on Quality and Affordable Education, Johor Bahru, Malaysia.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380-1400. https://doi.org/10.1177/0002764213498851

Sin, K., & Muthu, L. (2015). Application of big data in education data mining and learning analytics – A literature review. *ICTACT Journal on Soft Computing, 5*(4), 1035-1049. https://doi.org/10.21917/ijsc.2015.0145

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 57(10), 1510-1529. https://doi.org/10.1177/0002764213479366

Spencer, D. (2017). Work in and beyond the Second Machine Age: The politics of production and digital technologies. *Work, Employment and Society*, *31*(1), 142-152. https://doi.org/10.1177/0950017016645716

Staring, K. (2017, January 17). *Obviously, you can use the browser on an iPhone or iPad to access a Dhis server instance, but that is probably not what you need* [Comment on the online forum post *Using DHIS2 with IOS (Ixxxx Devices)*]. DHIS2 Community. https://community.dhis2.org/t/using-dhis2-with-ios-ixxxx-devices/33575

Tiwana, A. (2013). The rise of platform ecosystems. In *Platform Ecosystems: Aligning Architecture, Governance, and Strategy* (1st ed., pp. 3–21). Morgan Kaufmann.

Statoids. (2017, June 30). *Regions of Gambia*. http://www.statoids.com/ugm.html

Sørensen, C. (2016). The curse of the Smart Machine? Digitalisation and the children of the mainframe. *Scandinavian Journal of Information Systems, 28*(2), 57-68.

United Nations. (n.d.). *Secretary-general's roadmap for digital cooperation*. https://www.un.org/en/content/digital-cooperation-roadmap/

United Nations High-Level Panel on Digital Cooperation. (2020). *Report of the Secretary-General: Roadmap for digital cooperation*. United Nations. https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf

Walsham, G. (2012). Are we making a better world with ICTs? Reflections on a future agenda for the IS field. *Journal of Information Technology*, *27*(2), 87-93. https://doi.org/doi:10.1057/jit.2012.4

Wang, Y. (2016). Big opportunities and big concerns of big data in education. *Tech Trends*, *60*, 381-384. https://doi.org/10.1007/s11528-016-0072-1

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53*(4), 5-23. https://doi.org/10.1007/BF02504682

Watson, R. T. (2014). A personal perspective on a conceptual foundation for information systems. *Journal of the Association for Information Systems*, *15*(8), 514–535. https://doi.org/10.17705/1jais.00368

World Health Organization. (n.d.). *The global health observatory*. https://www.who.int/data/gho/info/about-the-observatory

Xu, M., David, J. M., & Kim, S. H. (2018). The Fourth Industrial Revolution: Opportunities and challenges. *International Journal of Financial Research, 9*(2), 90-95. https://doi.org/10.5430/ijfr.v9n2p90

Appendix