Using Data Science and Artificial Intelligence to Improve Teaching and Learning

Anass Bayaga
ORCID iD: https://orcid.org/0000-0003-0283-0262

Abstract

The current research examines how the domains of data science and artificial intelligence (A.I.) (collectively termed data-based A.I.) could improve teaching and learn in higher education. The current research raises awareness of the paths, dangers, and opportunities of data-based A.I. with a thrust on teaching and learning. Thus, it reacts to the teaching of programming to non-science disciplines through A.I., data science, and big data processing. The paper links A.I. and data science with pedagogy and curriculum design. From a multidisciplinary perspective, the paper explores the applications of data-based A.I. to inform students' learning and how higher education institutions teach and develop. Connecting with and reacting to the challenges faced, the author examines some models for teaching, learning, student support, and administration. Conclusively, the author argues for a data-based AI-enabled pedagogical approach. Instead of replacing teachers or administrators or using teacherbots for teaching and learning, data-based A.I. in higher education should extend human abilities in teaching, learning, and research with relevant administrative and leadership roles. Given the interconnectedness of data-based A.I., pedagogy, and curriculum design, the implycation from the findings of the research is thus that instead of education being a technology-centric endeavour, it should be human-centric, with human-centric-machine solutions. This approach allows humans to identify and critique human-centric risks and solutions continuously; hence, the universities would have to encourage and ensure that it nurtures creativity by maintaining academic skepticism as a health-check process in education.

Keywords: Artificial intelligence, Data Science, Machine learning, Teaching, Higher Education**,** Teacherbots, Curriculum Design, Pedagogy.

Print ISSN 1023-1757; Electronic ISSN: 2519-5476; DOI https://doi.org/10.29086/2519-5476/2022/sp39a6

1 Introduction

One drawback of data science and artificial intelligence (A.I.) solutions (i.e., massive open online courses - MOOCs) and Teacherbots, which are interventions in automated teaching, as Bayne (2015: 8) observed, is that '… current perspective of using automated methods in teaching is driven by a productivity-oriented solutionism', not by pedagogical or charitable reasoning, so we need to re-explore a humanistic perspective for mass education to replace the 'cold technocratic imperative'. Meanwhile, the advent of data science and A.I. (collectively termed data-based A.I.) coupled with the massification of education have, in some cases, resulted in defunding universities to find economically viable solutions to address depleted teaching, learning, and research funds. This phenomenon is largely seen in data-based A.I. replacing an increasing number of administrative roles and teaching staff. Thus, we now need to examine the effects of such nuances faced with a demand for creativity, entrepreneurial abilities, and new initiatives. Notably too, due to research connoting key sources of funding in top ranking universities locally and internationally, the MOOC form of teaching and learning has led university administrators to reduce academic staffing, which is aggressively being observed in countries such as the United Kingdom (University of Warwick) and Australia with a persistent move towards casual and short-term contracts (Andrews, Bare, Bentley, Goedegebuure, Pugsley & Rance 2016; Gallagher 2015).

One implication is that across various disciplines (whether human or techno-centric disciplines), while enhancing human cognition through teaching and learning is not a new phenomenon, certainly, its pathways, models, and how long it takes us to arrive at human-machine-centric solutions have become concerned and source for debate (Bostrom & Nick 2014; Popenici & Kerr 2017: Shulman 1987, Tuomi 2018). To that effect and in search for solutions, Popenici and Kerr (2017: 5) highlight that,

> Complex computing systems using machine learning algorithms can serve people with all abilities and engage to a certain degree in human-like processes and complex processing tasks employed in teaching and learning. This opens a new era for institutions of higher education. This type of human-machine interface presents the immediate potential to change how we learn, memori[s]e, access,

> and create information. The question of how long it will take to use this type of interface to enhance human memory and cognition is one which we are currently unable to answer.

Anchored upon Popenici and Kerr's (2017) view, examining the need to use data science and artificial intelligence (A.I.) to improve teaching and learning in human endeavors and thus higher education requires exploration and the understanding of data-based A.I. Arguably too, the understanding should not simply be focussed only on teaching and learning in higher education, but also on research in a technological-enhanced learning environment.

Hinged upon Popenici and Kerr's (2017) position, the principal intent of the current paper on the use of data-based A.I. to improve teaching and learning is guided by several reasons: The intent is to bring together students and lecturers to explore issues related to but not limited to pedagogy, curriculum development together with how to create environments for learning computation and programming. Accordingly, the paper focuses on teaching, learning, student support, and programming administration in non-science disciplines. It also intends to provide some simple and practical examples of predictive/mathematical modelling concerning data-based A.I. The intent is to stimulate a deeper level of thinking among university administrators, academics, students, and policymakers. The key to the discussion is integrating computational tools into university-related subjects (health, engineering, applied, and social sciences disciplines) with an emphasis on undergraduate teaching and the challenges in teaching programming to non-science disciplines. Effectively, it is envisioned that the current paper further lays a foundation in science and non-science teaching with a sound appreciation of scientific computing and big data processing.

2 Research aims

In order to achieve the intents mentioned above regarding important but possibly technical developments of data science and A.I., several issues need to be understood. This is because thorough knowledge and expertise in A.I. technology appear scarce. It is also coupled with stakeholders (educators and policymakers) struggling with getting up to speed with the operational and functional knowledge of the domain. Sadly too, A.I. technologies are inundated with … severely limited, and there are technical, social, scientific, and conceptual limits to what they can do (Tuomi 2018: 3). Even though there are competing A.I. uses and competencies depending on the Field of study. Some skills, competences and A.I. capabilities sourced from Tuomi (2018: 24) include application both in science, technology engineering and mathematics (STEM) as well as non-STEM disciplines collectively termed as health, engineering, applied and social sciences disciplines as depicted in Figure 1 below.

As with A.I., data science's key skills, competencies, and capabilities tend to have different goals and foci, which intend to assist in identifying and understanding sources and foci together with the context in which there are applied.

Therefore, it is useful to evaluate the use of data science and A.I. to improve teaching and learning, and to piece together the evolution of health, engineering, applied, and social sciences disciplines.

Consequently, the current chapter makes the following contributions by connecting and exploring A.I. and data science research's influence on pedagogy and curriculum development and shift from one to another. The author describes emerging features that have gained prominence in the A.I. and data science research community of practice in creating environments for learning computation and programming disciplines. A focus on the A.I., data science, and big data processing paths, dangers, and challenges in teaching programming to non-science disciplines is considered. Further, the author highlights and proposes pedagogical models suitable for teaching and learning in data-based A.I. and practical examples of predictive/algorithmic-based modelling. In essence, the contribution of the current paper lies in examining the following aims:

- A.I. and data science research influence pedagogy and curriculum development and simultaneous shift from one to another
- The emerging features which have gained prominence in A.I. and data science research community of practice in creating environments for learning computation and programming
- A.I., data science and big data processing paths, dangers, and challenges in teaching programming to non-science disciplines and
- Models for teaching and learning in data-based A.I. and practical examples of the use of predictive/algorithmic-based modelling

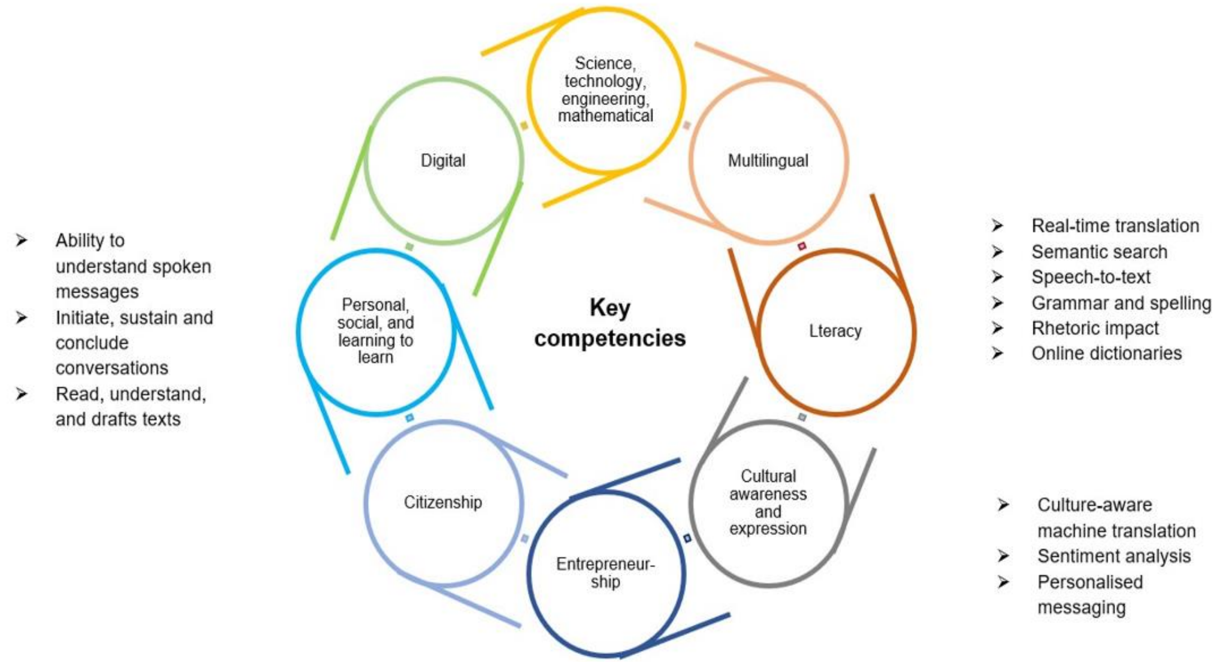

Figure 1: Data-based A.I. Key skills, competencies, and capabilities (adopted from Tuomi 2018: 24).

By examining the aims mentioned above, we could understand the nature and scope and provide a spur for the initiation of further discussions on teaching and learning in data-based A.I.

3 Methodology

Data Science, A.I., and higher education policymakers searching for reliable sources of information for learning and, consequently, paths in teaching progressively turn to systematic reviews, with the principal reason to achieve a standardised summary of evidence. There are several other reasons for using systematic reviews. Principally and for the current study, it was used to identify, select, assess, and synthesize the findings of comparable but distinct analyses.

Very often, uncertainty or poor quality tends to be associated with systematic reviews. Additionally, there appears to be no standard of measure nor universally accepted standard to develop systematic reviews. The consequence is some variability related to handling conflicts of interest, biases, and evaluation evidence, but keenly the general scientific rigour associated with the process. In addressing such drawbacks, the whole systematic review process begins with formulating the topic, building the intent (as noted earlier) to producing comprehensive aims that synthesise what the evidence shows and where knowledge gaps remain.

Drawn from the discussion so far, in addressing what works, thus using data science and A.I. to improve teaching and learning, the author used systematic review in helping to clarify two key issues, that is the known as well as the unknown, regarding using data science and A.I. to improve teaching and learning. Thus, anchored upon both the intents and aims, systematic review helped in integrating previous research findings related to,

(1) A.I. and data science influence on pedagogy and curriculum development;
(2) Features to consider in creating environments for learning computation and programming;
(3) Data-based A.I. paths, dangers, and challenges in teaching programming to non-science disciplines; and
(4) Predictive/ algorithmic, practical use cases and theories for teaching and learning in data-based A.I.

4 Background

4.1 AI and Data Science Influence Pedagogy and Curriculum Development

Over the past 70 years, seminal work related to models of curriculum designs and pedagogical developments has evolved. Fundamentally, three crucial issues stand out (Ralph & Tyler 1949; Walker 1971; Saylor Alexander, William, Lewis & Arthur 1981). While it is not the intent or within the scope of the current paper to espouse the different types of curriculum design models, principally for all curricula, the consensus is that curriculum design is one of continuous process and thus predicated upon a time with an intent to serve a dynamic society (for in-depth analysis of different types of models, see Ralph Tyler's model, Wheeler's cyclic model, Nicholls and Nicholls, Giles, Walker's model, and Hilda Taba's model).

Though there are many other factors, disciplines evolve from their content structure, the pedagogy used to deliver the content, and the knowledge associated with technologies. To this point, as with curriculum design, in the past three decades, one framework which has influenced the relationship, complexities, and dynamics linking the three main parts of knowledge (technology, pedagogy, and content) is the technological pedagogical content knowledge (TPACK - Figure 2) (Koehler & Mishra 2008; Mishra & Koehler 2006; Schmidt, Baran, Thompson, Mishra, Koehler & Shin 2009; Shulman 1986).

Essentially, TPACK tends to build on Shulman's (1987; 1986) explanation of pedagogical content knowledge (PCK) (see the work of Shulman 1987; 1986 for details). While PCK describes the close relation with pedagogical content knowledge, TPACK extends the understanding of PCK to using educational technologies (i.e., A.I.) to produce effective teaching and learning environment.

For instance, it has been argued that all three parts (technology, pedagogy, and content) are essential components for teaching and learning (Franklin 2004; Koehler & Mishra 2008; Schmidt *et al.* 2009). Koehler and Mishra's (2008) assessment is built on the fact that TPACK serves as a frame- work capable of integrating technology into teachers' knowledge to assist learning.

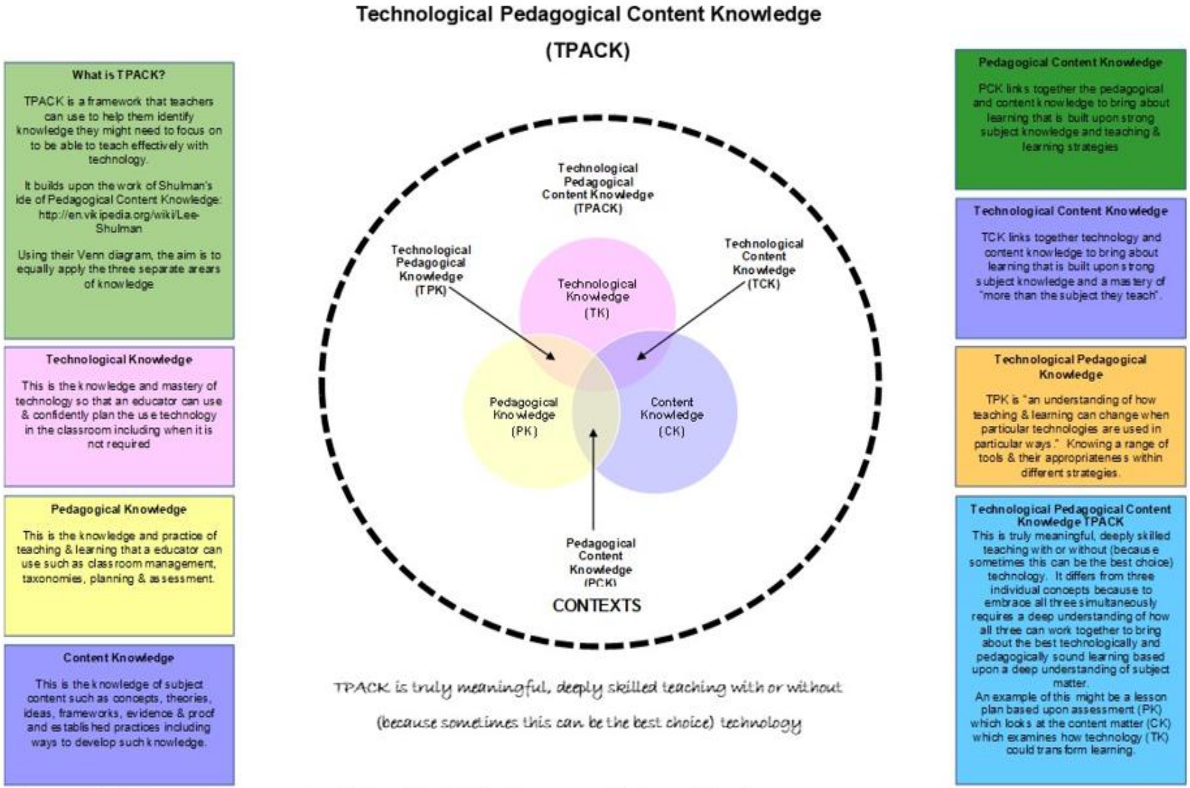

Figure 2: The components of the TPACK framework (graphic from https://ictevangelist.com/technological-pedagogical-and-content-knowledge/)

Based on the studies, it is conceivable to argue that integrating educational technologies such as A.I. can improve teaching and learning. In essence, we can confront the problems of teaching and learning with educational technologies such as A.I. and eventually establish the foundations of the new reforms. In the manner now being indicated, such a framework inclusive of A.I., through TPACK, could generate the needed knowledge to integrate successfully into teaching and learning. Resultantly, teaching and learning largely depend on an integration of knowledge resulting from different domains as well as *educational- technology reforms.*

One such reform may be drawn from the guidance from McTear *et al*. (2016: 1) related to machine learning (branch of A.I.), a collective term for which '… a range of different algorithms that a computer can use to learn from existing data rather than being specifically programmed to understand it'. Take into account the machine learning algorithms for personal assistance, namely Siri, Cortana, or Alexa. These algorithms tend to '… first turn our spoken request into text and then to analyse this text to understand what action to carry out' (McTear *et al*. 2016: 1).

Connecting studies this far, there appears to be a critical need for educational-technology reforms for teaching and learning environments. The essence is not only restricted to the inclusion of A.I. as advocated by TPACK, but the need to understand the scope and nature of processed data, particularly through the Field of data science *(*Donaldson, Ntarmos & Portelli 2017: 2). Informed by the need to improve our teaching and to learn through TPACK and hinged on Donaldson et al.'s (2017: 2) work in exploring the potentials of technology (machine learning (branch of A.I.)), it would mean that data science in teaching and learning becomes paramount. It is paramount because, as Donaldson *et al*. (2017: 2) allude, data science is,

> … an interdisciplinary area that brings together different people, processes, systems, and scientific methods to extract insight from data in various forms. This data can be highly structured with related values such as text, dates or numbers grouped or unstructured with many individual values recorded separately.

Though there is an overabundance of operational definitions of AI, Russell and Norvig (2010) provided one that is still relevant to the fledgling data science domain. This definition links A.I. and data science and how they influence pedagogy and curriculum development**.** In Russell and Norvig's (2010:34) view, A.I. generally proceeds '… based on the conjecture that every aspect of learning or any other intelligence feature can in principle be so precisely described that a machine can be made to stimulate it'.

Several implications arise: Thus, connecting the work of Russell and Norvig (2010) and other studies (i.e., Gane *et al.* 2018; Popenici & Kerr 2017; Porayska-Pomsta & Kaska 2015), new technologies are explicitly linked with the core functions of higher education. So, do their new computing proficiencies, with their developments explicitly linked with teaching, learning, and research. However, while there is a need to be cautious of superintelligence paths and dangers, for instance, in moving towards AI-based assessment systems (Luckin 2017), there are simultaneously new challenges and opportunities for teaching and learning (Bostrom & Nick 2014). For example, data-based A.I. have permeated and, to this point, influenced the domains of health, engineering, applied, and social sciences disciplines. Nevertheless, the variety of definitions, operational development, and understanding remains unclear and disputable.

As the subsequent section elaborates, it is difficult to efficiently create environments for learning computation and programming. Such difficulties are primarily true due to the alternative approaches of A.I. -+

Fundamentally, there are three alternative approaches to A.I.: data-based, logic-based, and knowledge-based (Tuomi 2018). Mostly, data-based neural A.I. is built on mathematical neural network models. Though research proceeded on neural networks, A.I. research soon transitioned to symbolic processes. It was soon noticed that computers could make logical inferences, which heralded the logic-oriented model of A.I. being the dominant form in the late 1950s to early 1970s. However, it was also noticed that human thinking could not sufficiently be simulated and guided fully by formal manipulation of logical statements. Consequently, domain-specific knowledge guided by different forms of representing knowledge inevitably was seen as the focal point of research in A.I. -thus leading to what is now expert systems, more generally known as knowledge-based systems.

The evolution of the different approaches has led to several potential capabilities of data-based A.I. For instance, such capabilities are informed by Freire's (1972) notion of the pedagogy of the oppressed. In essence, being in the era of data-based A.I., as highlighted by EPSC (2018), requires human-centric machines approach as opposed to being solely dependent on A.I. and data science only (Bayne, 2015; Gane *et al.* 2018; Popenici & Kerr 2017; Porayska-Pomsta & Kaska 2015; Scassellati, Brian, Henny Admoni & Matarić 2012; Scardamalia & Bereiter 2006; Woolf, Beverly & Park 2009). For instance, Popenici and Kerr (2017) highlighted the potential research and practice in the technology-enhanced learning environment, which should effectively influence data-based A.I. teaching and learning in higher education. Some studies are currently focusing on teacherbot, an intervention in automated teaching in higher education (Bayne 2015). Others use technology to assess multidimensional learning (Gane *et al.* 2018). Others have built intelligent interactive tutors for student-centered strategies for e-learning (Woolf 2009). As far back as 2012, robots were used to improve the learning abilities of patience with autism (Scassellati *et al.* 2012).

A few implications regarding data-based A.I. influence on pedagogy and curriculum development could be arrived at from the studies mentioned and the use cases. The first is that data-based A.I. in higher education could improve teaching and learning and thus enable an evidence-based approach to pedagogy and curriculum development (Porayska-Pomsta & Kaska 2015). Another key inference, as suggested by Scardamalia and Bereiter (2006), is the interconnectedness of knowledge building, theory, pedagogy, and curriculum development, via technology with a student-centered approach. The implications mentioned above have paved the foundation for A.I. and data science (data-based A.I.) influence on pedagogy and curriculum development, thus requiring careful consideration of features to consider in creating environments for learning.

4.2 Features to Consider in Creating Environments for Learning Computation and Programming

Widely speaking, developments related to features to consider in creating environments for learning programming and non-programming disciplines tend to focus limitedly on cognition. In some cases, there is limited to no connection between psychological and philosophical wellbeing and aspects of intelligence (Bostrom & Nick 2014; Gibney 2017; Popenici & Kerr 2017). In contrast, it is advocated that improving teaching and learning via data-based A.I. and thus creating learning environments should be hinged on human-like perspectives/qualities, including cognition (Bostrom & Nick, 2014). It is because, realistically, any data-based A.I. domain should engage in human-like perspectives/qualities. These may include the ability to learn and adapt. There is also the need to be able to synthesise and even self-correct with the ability to retrieve and use data for complex processes or activities. It is widely known that most data-based A.I. solutions rely on programming by the algorithms built in them. In circumventing this challenge and creating environments for learning, far more attention should be focused on the inbuilt capacity to learn, unlearn patterns and concurrently make student-centric predictions suitable for carrying out a task.

The implication is to focus on machine learning, one that can recognise far more patterns than it does prediction but given limited data. It should be able to uncover such hidden patterns in appropriate conditions far from the intended design. From numerous sources, while not in any particular order, there has been fledgling development in education and practical examples of the use of such features in predictive/algorithmic-based modelling (Alkhaldi, Pranata & Athauda 2016; Andrea, Holz, Sellers & Vaughan 2015; Bayne 2015; Chen, Wang, Nakanishi, Gao, Jung and Gao 2015; Daems, Erkens, Malzahn & Hoppe 2014; Deakin University 2014; Feild, Lewkow, Zimmerman, Riedesel, Essa 2016; Fyfe 2016; Gallagher 2015; Slavuj, Meštrović & Kovačić 2017; Zeiler & Fergus 2013).

For instance, Andrea *et al.* (2015) explored independent home use of brain-computer interfaces, which is the decision-based algorithm for selecting potential end-users. Bayne (2015), on the other hand, used teacherbot as an intervention in automated teaching in higher education. In other cases, Chen *et al.* (2015) developed a high-speed spelling programme with a noninvasive brain-computer interface. Deakin University (2014) used IBM Watson to power Deakin, aiming at exceeding students' needs. Recently, the University of Warwick launched a new department to employ all temporary or fixed-term teaching staff, focussing mainly on research productivity (Gallagher 2015) while ensuring that much of the teaching is carried out by teacherbot. In the past decade, Zeiler and Fergus (2013) focussed on visualising and understanding convolutional networks, while Alkhaldi *et al.* (2016) explored virtual and remote laboratory implementations. In other learning domains, Daems *et al.* (2014) highlighted the use of content analysis and domain ontologies to check students' understanding of science concepts, while Field *et al.* (2016) examined a scalable learning analytics platform for automated writing feedback.

Meanwhile, Fyfe (2016) and Slavuj *et al.* (2017) examined how to provide feedback on computer-based algebra homework and computer-assisted language learning, respectively. However, for the current study's aims, one that stands out is an embryonic type of data-based A.I. called IBM's supercomputer Watson at Deakin University in Australia to support students throughout the year (Deakin University 2014). Though primitive, a very efficient form of A.I. is also Siri on our iPhones. Google has also secretly tested an A.I. bot. For instance, with limited data, DeepMind, the A.I. branch of Google, developed a software called the AlphaGo (complex board game), which conditioned itself by learning patterns, thereby defeating the world's best player.

Thus far, the number of common implications could be linked with the ongoing development of environments for learning computation and programming and practical examples of using predictive/algorithmic-based modelling. One provides feedback on computer-based learning, and another could be adaptive or improve adaptivity in educational systems. Together with Apple's Siri, which might come across as '…low complexity A.I. solution or simply a voice-controlled computer interface, it is important to remember that it started as an artificial intelligence project funded in the USA by the Defense Advanced Research Projects Agency (DARPA) since 2001' (Popenici & Kerr 2017: 5). Another implication that could be drawn is algorithms suitability. Regardless, all environments or examples are based on algorithms or features that inherently need to be suitable with the ability to complete and comparatively predict tasks.

4.3 *Data-based A.I. Paths, Dangers, and Challenges in Teaching Programming to Non-science Disciplines*

As with any teaching and learning domain, Bostrom (2014) warns against the gratification derived from superintelligence's paths and strategies and cautions against the dangers and challenges. In the recent past, a vast array of studies have attempted to examine such a phenomenon concerning the use of A.I., data science, and big data processing paths, dangers, and challenges in teaching non-science and computing science disciplines (Armstrong 2015; Bostrom & Yudkowsky 2011; Donaldson, Ntarmos & Portelli 2017; González, Robbes, Góngora, Medina 2015; Drachsler & Greller 2016; Gibney 2017; Griffiths, Drachsler, Kickmeier-Rust, Steiner, Hoel & Greller 2016; Kharif, 2014; Popenici 2015; Perez 2016; Popenici & Kerr 2013; Tsur, Davidov & Rappoport 2010; U.S. National Science & Technology Council 2016; Wolpaw & Wolpaw 2012).

One of the principal concerns of Bostrom and Yudkowsky (2011) is the ethics associated with data-based A.I. What that means, for instance, as Kharif (2014) explains, is the privacy uncertainties related to students' data in terms of tracking data. In essence, what needs to be recognised is that in the domain of learning analytics, for instance, privacy concerns could impede the development of A.I., data science, and big data processing paths, thereby exacerbating the already envisioned dangers as well as challenges (Donaldson *et al.* 2017; Griffiths *et al.* 2016). Consequently, this slows the uptake of and teaching programming to non-science disciplines, which depend on A.I., data science, and big data processing. Therefore, it is binding upon such programmes to think of a human-centric machine checklist for engineering trusted learning modalities in terms of privacy (Drachsler & Greller 2016).

Several other phenomena of interest need further understanding, particularly with A.I., data science, and big data processing, dangers, and challenges. While work is ongoing, for instance, we do not yet fully comprehend the capabilities of machine learning in a human setting and hence lack sufficient technical know-how of machine learning abilities, limitations, and implications (Armstrong 2015). In augmented cognition regarding debugging and creativity tasks, limited research studies exist to measure concentration in programming with low-cost human-computer interface devices (Gibney 2017; González *et al.* 2015). In the recent past, for instance, Microsoft had to silence its new A.I. bot Tay when Twitter users thought it racism. While MOOCs hype for higher education is ever-increasing, Popenici (2015) suggests it possesses deceptive promises. Popenici and Kerr (2013) lament that the very nature of MOOC, coupled with the limitations of machine learning abilities and implications thereof, undermines higher education and hence impacts employment, economies, and one's democracy.

In essence, what could be considered from all these challenges and in such scenarios (in computational natural language learning) is the heavily developing semi-supervised recognition of human-computer devices (Tsur *et al.* 2010).

4.4 Predictive/ Algorithmic, Practical Use Cases and Theories for Teaching and Learning in Data-based A.I.

Along with the advent of A.I. and data science's influence on pedagogy and curriculum development, there is a debate regarding models for teaching and learning in data-based A.I. Routinely, along with features to consider in creating environments for learning computation and programming and practical examples on the use of predictive/algorithmic-based modelling, models for teaching and learning in data-based A.I. have also become a point of debate. Customarily, data-based A.I. paths, dangers, and challenges in teaching programming to non-science disciplines have led to many models of teaching and learning. Consequently, though there are some general forms of such models, fundamentally, models for teaching and learning in data-based A.I. are largely dependent on the domain of teaching and learning (health, engineering, applied and social sciences disciplines) (Beck & Mostow 2008; Baker Corbett & Aleven 2008; Blikstein 2011; Corbett & Anderson 1994; Dahman & Da 2019; Gonzalez-Marcos, Alba-Elías & Ordieres-Mer 2016; Köck & Paramythis 2011; Hamilton, Halverson, Jackson, Mandinach, Supovitz & Wayman 2009; Wong, Hsu, Wu, Lee & Hsu 2007). While the goal to critique and prescribe a particular model falls outside the aims of the current paper, in the main, few are discussed from various domains and different models.

In the past, an analytical method was used to measure project management competence (Gonzalez-Marcos *et al.* 2016). In the educational information technology domain, a machine learning model was used in predicting adult students' choice to continue ESOL courses (Dahman & Da 2019), while in the Field of user modelling and user-adapted interaction, Köck and Paramythis (2011) used activity sequence modelling and dynamic clustering to enhance personalised e-learning. Some researchers have, however, been able to accurately model students regarding intelligent tutoring systems using contextual estimation of slip and guess probabilities in Bayesian knowledge racing (Baker *et al.* 2008). Some have been able to model different types of students' behaviour in open-ended programming tasks using learning decomposition (Beck & Mostow 2008). Hamilton *et al.* (2001) used student achievement data to support instructional decisions in the past decade.

From the use cases, the general implication is that data-based A.I. *has an increasing influence* on deep-rooted pedagogy and curriculum development. Another important implication for discussion is that virtually all mentioned and unlisted models of A.I., data science, and big data processing systems and technologies need to rely on supervised learning modeling. Intend, developing human-centric data processing systems to be adjustable at suitable point intervals is needed.

For example, *transfer learning*, a practical variant of supervised learning (complex neural network), could be trained using a large sample of data sets. The intent is to learn to distinguish, discriminate, and separate vital components of the data set. Consequently, the system may be used and re-used for pattern recognition activities and predictive purposes. For instance, one could train the system to assist in labelling human faces, parts, or objects with a reasonable number of features and images. Upon learning or training to process, it becomes optimised for its intended use for pattern recognition and prediction with its deep layers. In the main, other training and learning processes could be assigned to its top layer to detect newly found and untrained images not found or seen before. The reason for the top layer approach is to reduce both computational and data requirements. Typical examples may include but are not limited to the GloVe vectors developed by Stanford University and Google's pre-trained inception image processing system, which is used to identify objects and similar image processing tasks.

While the current paper may not account for the advantages, one key fact is that such systems or networks produce relatively accurate statistical guesses. It does also assume the preconceived notion of the object of interest. Regardless of several disadvantages associated with pre-defined classes/ notions, particularly in race and race relations, this type of system/ network has become one of the dominant models for teaching and learning in data-based A.I. and practical examples of the use of predictive/algorithmic-based modelling. Technically, such a system powering (for example, a self-driving car) will have known the pre-defined classes of cyclist, truck, train, or child to take decisive action(s).

Nevertheless, in recent times, due to drawbacks associated with such systems and based upon Skinnerian models of operant conditioning, other A.I., data science, and big data processing-based models such as the *AlphaZero* game A.I. used to reinforce learning have begun. Used in the fashion industry, another variation of reinforcement learning known as generative adversarial networks (GANs) has been inspired by Skinnerian models of operant conditioning capable of deceiving one that the data used originates from the training data set. Another variation of such a network is Turing learning, which is aimed at interacting with its surroundings to predict or guess the possible source of the data-thus, whether the source of data is from an actual environment or a machine.

Several other ramifications need to be examined: Pedagogically and in terms of curriculum development, such a multi-layered model is fundamentally akin to norms and values that are implied or tacit, often expressing unarticulated emotional disposition. It is also implied that students and teachers may be deprived of agency power, which ought to have allowed for evolution and making decisions on their own. Regardless of such technicalities and pre-defined classes/ notions, such intelligence is predicated upon similar ones akin to humans, wherein humans tend to associate, for instance, psychological, environmental, economic, technological, and neurological circumstances or conditions with learned conducts and actions. Psychologists have named this Pavlovian phenomenon theory of reflexes or Skinnerian reinforcement learning, where Vygotsky noted it as the developmentally simplest learning model. Sadly, one key disadvantage is seeing worldview perception as a repetition of past events. Additionally, such pre-defined classes are mostly supplied by humans with their already existing pre-defined classes. Which inherently implies an additional layer of personal and cultural biases associated with A.I., data science, and big data processing.

5 Discussion

The current paper intends to examine data science, and A.I. uses to improve teaching and to learn in higher education. In response, reference was made to A.I. and data science's influence on pedagogy and curriculum development, as well as features to consider in creating environments for learning**.** Attention was also given to data-based A.I. processing paths, dangers, and challenges in teaching programming to non-science disciplines. The review also considered models/theories for teaching and learning in data-based A.I. and practical examples of the use of predictive/algorithmic-based modelling.

Based on the intent, the evidence from previous studies highlights the ongoing impact of artificial intelligence on teaching and learning in higher education (Popenici & Kerr 2017). Nevertheless, it was also evident that while we hype teaching and learning in data-based A.I., some categories of knowledge acquisition, for instance, models for teaching and learning, ought to be human-centric. It was equally evident that research does not converge adequately on what it means to improve student's cognition (Tuomi 2018). For instance, in terms of curriculum delivery and dependent on the complexity, it is quite evident that adaptive learning platforms are more suitable for machine learning or data-based A.I. regarding a wide variety of educational usages. However, few questions remain unresolved and, thus, open avenues for further exploration. For instance, it is still unclear how various contexts could be accommodated or are ready to adapt to machine learning, nor what the possibilities are for countries with poor infrastructure and limited technological resources. Nevertheless, in close link with related pedagogical activities, we need far more depth regarding stimuli such as inputs regarding multi-touch screens and virtual reality (V.R.) in university learning processes. In effect, in terms of assessment of students, for instance, and equally dependent on complexity, computer-based assessment for learning (CBAfL) coupled with personalised feedback need to be developed. That is to say, teaching and learning in data-based A.I. models should focus on present cognition levels. It should also account for the type of questions designed to articulate the right levels of difficulty index while granting the most information related to the students' present state of development and progression rate (Feild *et al.* 2016).

However, implied challenges to overcome may include barriers to adoption. Even though several barriers were noted, notably the implied challenges rested upon the lack of or the need to understand educational data and the associated fits. There was also a need to understand students and associated privacy issues and ethical concerns of students' data.

Such implications pave the way for future work in many forms. One such form of future work is the need to know what results in appropriate knowledge in a human-centric approach, coupled with the form of acquisition, creation, and learning. It would also mean examining further procedures in regulating cognitive processes. It would mean attention and emotion involved in learning processes as well as both the social and practical implycation for learning while using data-based A.I., given that they sometimes are unrealistic. It means there is a need for data-based A.I. capabilities, which should be far more closely linked with or resemble human intelligence than simplified models of learning biological intelligence. Thus, it is important to rethink A.I., data science, and big data processing, dominating behaviouristic models of learning inspired by both Pavlov and Thorndike.

Despite the recognition of the abundance of use cases of data-based A.I. and the effects of computing algorithms influencing mundane activities of education, education remains a human-centric endeavour. As such, educational-technological reforms should open new possibilities for higher education's teaching and learning, simultaneously highlighting the urgent need to be cognisance that data-based A.I. does not replace teachers. Instead, there is an opportunity for augmentation via such tools as V.R. and augmented reality (A.R.). In essence, instead of replacing teachers, data-based A.I. in higher education should extend human abilities in teaching, learning, and research together with administrative and leadership roles. For instance,

> assistive technologies, such as text to speech, speech to text, zoom capacity, predictive text, spell checkers, and search engines- are just some designed technologies to assist people with a disability. These technological solutions were later expanded, and we now find them as generic features in all personal computers, handheld devices, or wearable devices. These technologies now augment the learning interactions of all students globally, enhancing possibilities opened for teaching and design of educational experiences (Popenici & Kerr 2017: 5).

Clearly, higher education is faced with possibilities engineered by data-based A.I. teaching, learning, organisation, and governance with increasing implycations and possibilities. The implication so far is premised on the fact that education should be human-centric, with human-centric machine solutions instead of education being a technology-centric endeavor. That notion then enables us to be cautious of A.I., data science, big data processing paths, dangers, and challenges, particularly in teaching programming to non-science disciplines. Such notions also allow humans to continuously identify and critique human-centric risks and, consequently, human solutions and hence the need to nurture creativity, thus maintaining academic skepticism as a health check in education. It is because such a notion of education reduces mere information delivery and recollection. Such academic skepticism also paves the way to interrogate power structures and power of control underpinning algorithms or data-based A.I. This is because people in control of data-based A.I. certainly have an overwhelming influence over the rest of contemporary society and culture.

Additionally, massification, as well as wider participation in higher education coupled with increasing student numbers, class sizes, and staff costs, make data-based A.I. an attractive possibility that has become far more evident amid MOOCs. Similarly, the drive for personalised learning resulting from teacherbot, or cloud-lecturer used for blended delivery courses or fully online courses, is beginning to be computing solutions targeting administrative aspects of teaching, content delivery, administrative feedback, and supervision as the alternative to traditional teaching assistants. These data-based A.I. solutions, together with others used to detect plagiarism, raise questions about who is setting the agenda for teaching and learning – big tech corporate ventures in the form of techlords and the quasi-monopoly or higher education.?

What could then be ascertained is the potential to enormously change higher education administratively. Similarly, teaching and learning are faced with varieties of challenges. That is, while data-based A.I. use cases are naturally automated, there is a need to be far more cautious with more complex tasks (Tuomi 2018). This is primarily because data-based A.I. still falls short of absolute human-centric dispositions regarding emotional and psychological wellbeing. The evidence is that most attempts for data-based A.I. solutions are currently limited to superficial solutions characterised by algorithms meant for repetitive tasks.

6 Conclusion

This far, the conclusion in response to the four aims is that A.I., data science, and big data processing paths and techniques have specific technologies and domains of usage and application. For instance, it was noted that intelligent tutoring systems, adaptive learning resources, and assessments tend to be most evident in teaching and learning and thus have the potential to enhance pedagogy and curriculum design. However, the need to create an interactive platform that will bring together students and academics may provide far more support. This means there is the need to provide enabling data-based A.I. features in an interdisciplinary and human-machine-centric fashion instead of focusing on one domain tech-centric. Given the interconnectedness of data-based A.I., pedagogy, and curriculum design, the implication so far is premised on the fact that instead of education being a technology-centric endeavour, it should be human-centric, with human-centric-machine solutions, which allow for humans to continuously identify and critique human-centric risks and consequently human solutions and hence the need to nurture creativity, thus maintaining academic scepticisms as a healthy in education.

Reference

Alkhaldi, T., I. Pranata & R.I. Athauda 2016. A Review of Contemporary Virtual and Remote Laboratory Implementations: Observations and Findings. *Journal of Computers in Education* 3: 329 – 351. https://doi.org/10.1007/s40692-016-0068-z

Andrews, S., L. Bare, P. Bentley, L. Goedegebuure, C. Pugsley & B. Rance 2016. Contingent Academic Employment in Australian Universities. Melbourne: LH Martin Institute. https://www.lhmartininstitute.edu.au/documents/publications/2016-contingent-academic-employment-in-Australian-universities-updatedapr16.pdf (Accessed on 26 August 2017.)

Andrea, K., E.M. Holz, E.W. Sellers & T.M. Vaughan 2015. Toward Independent Home Use of Brain - Computer Interfaces: A Decision Algorithm for Selecting Potential End-users. *Archives of Physical Medicine and Rehabilitation* 96,3: 27 - 32. https://doi.org/10.1016/j.apmr.2014.03.036 PMid:25721544

Armstrong, H. 2015. *Machines that Learn in the Wild: Machine Learning Capabilities, Limitations, and Implications.* London: Nesta.

Baker, R.S.J.d., A.T. Corbett, A.T. & V. Aleven 2008. More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing. In *Proceedings of the 9th International Conference on Intelligent Tutoring Systems*. Berlin, Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-540-69132-7_44

Bayne, S. 2015. Teacherbot: Interventions in Automated Teaching. *Teaching in Higher Education* 20,4: 45 - 56. https://doi.org/10.1080/13562517.2015.1020783

Beck, J.E. & J. Mostow 2008. How Who Should Practice: Using Learning Decomposition to Evaluate the Efficacy of Different Types of Practice for Different Types of Students. In *Proceedings of the 9th International*

Conference on Intelligent Tutoring Systems. Berlin, Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-540-69132-7_39

Blikstein, P. 2011. Using Learning Analytics to Assess Students' Behavior in Open-Ended Programming Tasks. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2090116.2090132

Bostrom, N. 2014. *Superintelligence: Paths, Dangers, Strategies*. 1st Edition. Oxford: Oxford University Press. https://doi.org/10.1017/CBO9781139046855.020

Bostrom, N. & E. Yudkowsky 2011. The Ethics of Artificial Intelligence. In Frankish, K. & W.M. Ransey (eds.): *Cambridge Handbook of Artificial Intelligence*. Cambridge, UK: Cambridge University Press.

Chen, X. Wang, Y. Nakanishi, M. Gao, X. Jung, T.P. & S. Gao 2015. High-speed Spelling with a Noninvasive Brain – Computer Interface. *Proceedings of the National Academy of Sciences* 112,44: 6058 – 6067. https://doi.org/10.1073/pnas.1508080112 PMid:26483479 PMCid:PMC4640776

Corbett, A.T. & J.R. Anderson 1994. Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. *User Modeling and User-Adapted Interaction* 4,4: 253 - 278. https://doi.org/10.1007/BF01099821

Daems, O., M. Erkens, N. Malzahn & H.U. Hoppe 2014. Content Analysis and Domain Ontologies Check Learners' Understanding of Science Concepts. *Journal of Computers in Education* 1: 113 - 131. https://doi.org/10.1007/s40692-014-0013-y

Dahman, M.R. & H. Dag 2019. Machine Learning Model to Predict an Adult Learner's Decision to Continue ESOL Course. *Education and Information Technologies* 24,4: 2429 - 2452. https://doi.org/10.1007/s10639-019-09884-5

Deakin University 2014. IBM Watson now Powering Deakin. A New Partnership that Aims to Exceed Students' Needs. http://archive.li/kEnXm (Accessed on 30 October 2016.)

Donaldson, P., N. Ntarmos & K. Portelli 2017. A Systematic Review of the Potential of Machine Learning and Data Science in Primary and Secondary Education. http://eprints.gla.ac.uk/151244/

Cutts, Q. (leader), P. Donaldson (co-leader) & E. Cole (co-leader) 2017. Searching for Early Developmental Activities Leading to Computational Thinking Skills. Professional Learning and Networking for Computing (PLAN C). ITiCSE Working Group 9. https://www.researchgate.net/publication/317989395_Searching_for_Early_Developmental_Activities_Leading_to_Computational_Thinking_Skills

Drachsler, H. & W. Greller 2016. Privacy and Analytics: It's a DELICATE Issue: A Checklist for Trusted Learning Analytics. In Proceedings of the 6th International Conference on Learning Analytics & Knowledge, LAK '16. ACM, New York, 89 - 98. https://doi.org/10.1145/2883851.2883893

Franklin, C. 2004. *Teacher Preparation as a Critical Factor in Elementary Teachers: Use of Computers.* Harmondsworth: Penguin.

Fyfe, E.R. 2016. Providing Feedback on Computer-based Algebra Homework in Middle-school Classrooms. *Computers in Human Beha-viour* 63: 568 - 574. https://doi.org/10.1016/j.chb.2016.05.082

Freire, P. 1972. *Pedagogy of the Oppressed*. Harmondsworth: Penguin.

Gane, B.D., Z.Z. Sania & J.W. Pellegrino 2018. Measuring What Matters: Using Technology to Assess Multidimensional Learning. *European Journal of Education* 53,2: 176 - 187. https://doi.org/10.1111/ejed.12269

Gallagher, P. 2015. The University of Warwick's New Department Employs all Temporary or Fixed-term Teaching Staff. *The Independent* 24 September 2015. http://www.independent.co.uk/news/education/educationnews/the-university-of-warwick-launches-new-department-to-employ-all-temporary-or-fixed-term-teaching-10160384.html (Accessed on 1 May 2017.)

Giles, H.H., S.P. McCutchen & A.N. Zechiel 1942. *Exploring the Curriculum*. New York: Harper.

González-Marcos, A., F. Alba-Elías & J. Ordieres-Meré 2016. An Analytical Method for Measuring Competence in Project Management. *British Journal of Educational Technology* 47,6: 1324 - 1339. https://doi.org/10.1111/bjet.12364

Griffiths, D., H. Drachsler, M. Kickmeier-Rust, C. Steiner, T. Hoel & W. Greller 2016. Is Privacy a Showstopper for Learning Analytics? A Review of Current Issues and Solutions. *Learning Analytics Community*

Exchange 6-19, 45-57. Available at: https://cdn.tc-library.org/Rhizr/Files/FkE9DdrKdtH7PAQaw/files/LACE-review-6_privacy-show-stopper.pdf

Hamilton, L., R. Halverson, S. Jackson, E. Mandinach, J. Supovitz & J. Wayman 2009. *Using Student Achievement Data to Support Instructional Decision Making* (NCEE 2009-4067). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance

Kharif, O. 2014. *Privacy Fears Over Student Data Tracking Lead to InBloom's Shutdown.* Bloomberg Company. https://www.bloomberg.com/news/articles/2014-05-01/inbloom-shuts-down-amid-privacy-fears-over-student-data-tracking

Köck, M. & A. Paramythis 2011. Activity Sequence Modeling and Dynamic Clustering for Personalized E-Learning. *Journal of User Modeling and User-Adapted Interaction* 21,1: 51 - 97. https://doi.org/10.1007/s11257-010-9087-z

Koehler, M.J. & P. Mishra 2008. Introducing TPCK. In AACTE Committee on Innovation and Technology (ed.): *The Handbook of Technological Pedagogical Content Knowledge (TPCK) for Educators*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lewkow, N., J. Field, N. Zimmerman, M. Riedesel, A. Essa, D. Boulanger, J. Seanosky, V. Kumar, Kinshuk & S. Kode 2016. A Scalable Learning Analytics Platform for Automated Writing Feedback. Proceedings of the Third (2016) ACM Conference on Learning @ Scale. In *EDM* 688-693. https://doi.org/10.1145/2876034.2893380

Luckin, R. 2017. Towards Artificial Intelligence-based Assessment Systems. *Nature Human Behaviour* 1,28: 45 - 67. https://doi.org/10. 1038/s41562-016-0028

Mishra, P. & Koehler, M. J. 2006. Technological Pedagogical Content Knowledge: A Framework for Integrating Technology in Teachers' Knowledge. *Teachers College Record* 108,6: 1017 - 1054. https://doi.org/10.1177/016146810610800610

Popenici, S. & S. Kerr 2017. Exploring the Impact of Artificial Intelligence on Teaching and Learning in Higher Education. *Research and Practice in Technology Enhanced Learning* 3,4: 12 - 22. https://doi.org/10.1186/s41039-017-0062-8 PMid:30595727 PMCid:PMC6294271

Porayska-Pomsta, K. & T. Kaska 2015. A.I. in Education as a Methodology for Enabling Educational Evidence-Based Practice. In *Seventeenth International Conference on Artificial Intelligence in Education (AIED 2015)*, 52-61. Madrid.

Russell, S.J. & P. Norvig 2010. Artificial Intelligence: A Modern Approach. 3rd Edition. Upper Saddle River: Prentice-Hall.

Saylor, J.G., W.M. Alexander & A.J. Lewis 1981. *Curriculum Planning for Better Teaching and Learning*. New York: Holt, Rinehart and Winston.

Scassellati, B.. A. Henny & M. Maja 2012. Robots for Use in Autism Research. *Annual Review of Biomedical Engineering* 14,1: 275 - 294. https://doi.org/10.1146/071811-150036; https://doi.org/10.1146/annurev-bioeng-071811-150036 PMid:22577778

Scardamalia, M. & J. Bereiter 2006. Knowledge Building: Theory, Pedagogy, and Technology. In Sawyer, R.K. (ed.): *Cambridge Handbook of the Learning Sciences*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511816833.008

Schmidt, D.A., E. Baran, D. Thompson, P. Mishra, J. Koehler & T.S. Shin 2009. Technological Pedagogical Content Knowledge (TPACK): The Development and Validation of an Assessment Instrument for Pre-service Teachers. *Journal of Research on Technology in Education* 42, 2: 123 - 149. https://doi.org/10.1080/15391523.2009.10782544

Shulman, L.S. 1986. Those who Understand: Knowledge Growth in Teaching. *Educational Researcher* 15,2: 4 - 14. https://doi.org/10.3102/0013189X015002004

Slavuj, V., A. Meštrović & B. Kovačić 2017. Adaptivity in Educational Systems for Language Learning: A Review. *Computer Assisted Language Learning* 30: 64 - 90. https://doi.org/10.1080/09588221.2016.1242502

Tuomi, I. 2018. *The Impact of Artificial Intelligence on Learning, Teaching, and Education: Policies for the Future*. In Cabrera, M., R. Vuorikari & Y. Punie (eds.). A Science for Policy Report by the Joint Research Centre (JRC) of the European Union. EUR 29442 EN. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/12297

Tyler, R.W. 1949. *Basic Principles of Curriculum and Instruction*. Chicago & London: The University of Chicago Press.

Tyler, R.W. 1986. The Five Most Significant Curriculum Events in the Twentieth Century. *Educational Leadership* 4,4: 36 - 38. https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_198612_tyler.pdf

EPSC 2018. The Age of Artificial Intelligence: Towards a European Strategy for Human-Centric Machines. https://op.europa.eu/en/publication-detail/-/publication/f22f6811-1007-11ea-8c1f-01aa75ed71a1/language-en

Walker, D.F. 1971. A Naturalist Model of Curriculum Development. *A School Review* 1: 51 - 67. https://doi.org/10.1086/443014

Woolf, B.P. 2009. *Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing e-Learning*. San Francisco, CA: Morgan Kaufmann.

Wong, W.K., S.C. Hsu, S.H. Wu, C.W. Lee & W.L. Hsu 2007. LIM-G: Student-initiating Instruction Model Based on Cognitive Knowledge for Geometry Word Problem Comprehension. *Computers & Education* 48,4: 582 - 598. https://doi.org/10.1016/j.compedu.2005.03.009

Zeiler, M.D. & R. Fergus 2013. Visualizing and Understanding Convolutional Networks. *Computer Vision and Pattern Recognition* 1-11**.** https://doi.org/10.48550/arXiv.1311.2901 https://arxiv.org/pdf/1311.2901.pdf

Professor Anass Bayaga
Cognitive Mathematics/ STEM
Senior Secondary Education
Faculty of Education
Nelson Mandela University
anass.bayaga@mandela.ac.za