"No One Left Behind" – Designing A Conceptual Framework for Nurturing a Data-Literate Mindset in Higher Education Administration

Ji Hu, New York University Shanghai, China
Xuchu Meng, New York University Shanghai, China

The Southeast Asian Conference on Education 2020
Official Conference Proceedings

Abstract
As Higher Education Institutions (HEIs thereafter) are eagerly engaging data-informed decision making, attentions are mostly put on setting up specialized data teams for the job, but much less on nurturing data-literate mindset and capacity of the administrative team as a whole. This missing link leads to at least two issues that can undermine the efforts towards effective data-informed decision making. The first is "garbage in, garbage out". Most data the data people are working on comes from the seemingly non-data-related workers. The second is the loss of competencies or motivations for the non-data people to produce better quality work in today's innovative environment. As an attempt to fill the gap, a conceptual framework is proposed in this working paper to tackle the question of how to nurture a data-literate mindset – to be curious about and aware of the importance and implications of data, before being able to work with it (Bhargava & D'Ignazio, 2015) – in workers of higher education administration. Benefits of employing Backward Design Model (Wiggins & McTighe, 1998) as methodology to develop this framework is discussed. Common misconceptions around data, identified in practice, are mapped against Bloom's taxonomy of cognitive development (Bloom & Krathwohl, 1956). Last, the working paper discusses future work to operationalize the framework and to evaluate the effectiveness of such training in enhancing institutional data efforts.

Keywords: Training Framework, Data-Literate Mindset, Higher Education Administration, Backward Design Model

iafor
The International Academic Forum
www.iafor.org

1. Introduction

When more higher education institutions (HEIs), actively or passively, are brought into the information age, both the growing awareness of benefits that can be generated from data (Huron, ACE & GIT, 2019) and the increasing volume of data that is collected (Wilsdon et al. 2015) call for a certain level of data literacy of the administrative workers to inform their daily operation. However, among the data literacy training frameworks that have been proposed, which target different audiences (Wolff et al., 2016), proficiency levels (Qlik, 2018), or even organizational development phases (Sternkopf, 2017; Sternkopf & Muller, 2018), competencies and skills that are commonly required of specialists prevail (Bonikowska et al., 2019). What is often underrepresented in the established work around data literacy training is the necessity and ways to engage data non-specialists (OECD, 2017), who are mainly workers not in traditionally-viewed data-intensive functions or roles.

The same trend is seen in higher education. While HEIs have been growing teams and building capacity of "data specialists" like institutional researchers, learning analysts, or data librarians (Swing, 2016; Kim, 2018), they are also called on to place efforts elsewhere. There is rising concern that data of poor quality can greatly undermine the efforts to drive insights from analytics, as suggested in established studies. Early research work indicated direct relationship between data quality and individual work (Wu & Wang, 2006; Santos, Tokaoka & de Souza, 2010). Evidence was also found that validates the relationship of data quality and organizational outcomes (Sheng & Mykytyn, 2002). Despite its root in Information Systems research, work on data quality is also emerging in the context higher education. Jim and Chang (2018), in their work on data governance among universities, claims that "data quality is the foundation of the data-driven decision-making process", and low quality data leads to misleading and ineffective analytical results. Issues categorized into "Trusted Data" have been on the top of EDUCAUSE's (a higher education IT community) "2019 Top 10 IT Issues" (Grajek, 2019). In response, HEIs started to hire data management professionals, assemble data governance committees, or implement data quality tools (Hayhurst, 2019). However, a significant portion of the data on which specialists work often comes from the administrative by-products of work done by non-specialists (Australian Bureau of Statistics, 2011). Most data quality issues also come down to a lack of data quality assurance mechanism in non-specialists' work, or in other words, neglect of engaging them into the institutional data efforts. The illusory belief that non-specialists do not work with data only makes specialists' work harder because the quality of data cannot satisfy analytical needs (Australian Bureau of Statistics, 2011).

Developing a sense of data among non-specialists is vital, also because in higher education, student success and research development are collaborative work (Cavanagh, 2019) that assumes informed decisions at all levels and aspects of the university, not only in centralized offices. Trend of data democratization in higher education emerges that aims to extend direct access to institutional data to even non-specialists, and to empower them to utilize data (Harfield, 2017). This trend is fostered by both operational needs and technological possibility. Increasing demand of data from both government agencies and the general public is pressuring HEIs to work beyond the capacity of their traditional data-centric functions (Swing, 2016). The time lag caused by routing data requests to specialists raised concerns about the timeliness of findings to inform practice (Petrides, 2004). Meanwhile technological advancements simplified the work to query and use institution-wide data through tools such as web-based dashboards, which allows workers to access and use data with minimum training requirements (Petrides, 2004). However, insufficiency of intent or confidence to adopt data-informed decision-making beyond the "specialist community" still prevails (Laskovsky & O'Donnell, 2018). Most existing data literacy training programs are developed in the context of for-profit organizations. The trouble that HEI workers have to transpose the knowledge to the special setting of higher education further impedes data adoption among non-specialists in HEIs (Laskovsky & O'Donnell, 2018).

That said, HEIs need to find a way to effectively combat the siloed nature of traditional data practice that is overly concentrated on "data specialists". This working paper proposes development of a conceptual framework for nurturing a data-literate mindset of HEI administrative workers, especially non-specialists who used to be left out.

2. Data-Literate Mindset

The term of data-literate mindset in this article—built upon the exposition of data literacy by MIT—is defined as the awareness and curiosity of the importance and implications of "the ability to read, work with, analyze, and argue with data" (Bhargava & D'Ignazio, 2015). When most existing data literacy training programs are tailored for specialized personnel, they set natural barriers for reaching non-specialists. Both the training materials about advanced skills and techniques of working with data, and variety of prerequisites on proficiency levels of the trainees are intimidating to regular administrative workers of HEIs. Therefore, evoking the curiosity and awareness of the benefits of data to both the institution and workers themselves is a critical first step to bring non-specialists into the data world - to secure their adoption and buy-in (Qlik, 2018). In his bestseller book "Drive", Daniel Pink argues that motivations in completing tasks that require cognitive and creative skills are dominated by intrinsic factors, one aspect of which is purpose, "the desire to do something that has meaning and is important" (Pink, 2009), which is the foundation of buy-in.

For HEI administrative workers, the "meaning" and "importance" of data is that data has already been incorporated into the ways of how their daily work is shaped, and underlies the quality of their work (Sandler Training, 2019). A data-literate mindset here entails the awareness that data is an inherent part of their existing responsibilities, not additional burdens; and the curiosity about the possibility how their work and daily decision-making can be improved by actively leveraging the power of data, against instinct or snap decisions.

3. Utilizing Backward Design Model to Design Data-Literate Mindset Training

Despite importance of professional training to prepare higher education workers for the evolving responsibilities of their positions (Holzweizz, Walker & Conrey, 2018), it is poorly implemented when related to data (Laskovsky & O'Donnell, 2017). In HEIs, data is often an assumed skillset (Laskovsky & O'Donnell, 2017), and not considered professional development priorities (Knight, 2014). Professional needs in data, thus, are not clearly defined, causing lack of relevance in training materials and programs (Florian & Hegarty, 2004). It is proposed that Backward Design Model be employed to design data-literate mindset training. The model is originally developed for course design, but applies equally to professional training. Instructors typically develop a course by first designing activities through which the content is taught, then aligning assessments with the activities, and finally drawing connections to learning goals (Bowen, 2017). Backward Design Model, however, reverses the process, featuring three sequential stages, "identify desired results", "determine acceptable evidence", and "plan learning experiences and instruction" (Wiggins & McTighe, 1998). At the first stage, expected learning outcomes are defined with different levels of priority. It is to answer the question "what is expected of the learners when they complete the course". Wiggins and McTighe (1998) provided guidance to establish curricular priorities, where expected outcomes are positioned along the spectrum of worthiness. For example, within the available time and resources, enduring understandings, as compared to knowledge only "worth being familiar with", is given higher priority. The second stage is where instructors determine what assessment evidence is acceptable to demonstrate accomplishment of the expected learning outcomes. Assessment can be conducted in various forms, but need to provide direct evidence on whether the learning goals are met. Otherwise it becomes nothing but an additional burden for both instructors and learners. At the final stage, instructors design teaching as a means to an end, in terms of what to be taught and how. Instructional resources and teaching strategies are designed at this stage against the expected learning outcomes and assessment methods. The benefit of this model is obvious. It focuses limited resources on a clear pathway towards expected outcomes of learners, and eliminates learning activities that are purposeless and thus useless in achieving this goal. Given its natural emphasis on learning outcomes, employment of this model guarantees the relevance of data-literate mindset training, which is lacking in most data training programs for higher education workers.

3.1 Breaking Common Misconceptions of Non-Specialists Around Data

The "fuzzy concept" of data-literate mindset, however, makes it hard to identify a specific and concrete list of learning outcomes without knowing the context where this concept is to be applied. On the other hand, practical knowledge indicates that there are common misconceptions around data among higher education workers, especially non-specialists. These misconceptions prevent them from producing quality data and actively utilizing data in their daily work. The efforts to develop a conceptual framework of data-literate mindset training also entail debunking and breaking these misconceptions. Since it is hard to elaborately characterize a data-literate mindset, this working paper explores and challenges the boundary of this concept with its application in higher education administration, by demystifying misconceptions that are outside of this domain. In order to streamline the efforts of identifying these misconceptions that are of a practical nature, they are mapped against Bloom's taxonomy (Bloom & Krathwohl, 1956), which provides an actionable roadmap for HEIs to operationalize this concept in their own specific contexts.

Bloom's taxonomy is a model for instructors to identify educational learning objectives along a spectrum of cognitive complexity (Adams, 2015). It was devised by a group of educators in 1950s, and since then has wide usage and significant influence on teaching and learning practice (Adams, 2015). The taxonomy consists of six categories of cognitive learning objectives, “knowledge, comprehension, application, analysis, synthesis, and evaluation” (Bloom & Krathwohl, 1956). A learner that masters the all six levels is expected to be able to memorize learned material, grasp the meaning, transpose it in new contexts, deconstruct and reconstruct it, and eventually perform value judgement on it. The six categories differentiate between levels of complexity and specificity, respectively requiring skills ranging from lower order that needs less cognitive processing to higher orders (Adams, 2015). In accord with the logic behind backward design, Bloom’s taxonomy is also beneficial by calling attention to developing learning objectives across foundational and advanced level skills. Following discussion outlines the six stages that higher education workers are to proceed through in breaking data-related misconceptions. Higher education workers with a data-literate mindset are expected to be able to understand the concept of data (3.1.1), show evidence of comprehension through paraphrasing it in the context of their own work (3.1.2), apply the concept to redefine quality work (3.1.3), break down institutional data flow into elements to define data quality (3.1.4), reunion data elements to form data reporting (3.1.5), and finally critically judge the value brought by data (3.1.6).

3.1.1 Knowledge of Data Concept

The foundation of the data-literate mindset is an inclusive and adaptive understanding of the concept of data. A typical impression of data is numeric, or sometimes aggregated information that is stored in data warehouse, passed through equations, and presented in dashboards or tabular reports, naturally leading to a sense of detachment if one’s work does not seem to involve numbers beyond elementary calculations. However, technology advancement has allowed us to access and treat “non-typical data” in the same way as quantitative information (McEvoy, 2018). Qualitative data is a primary source of “non-typical data”, and the datasets they form make up a significant portion of the environment around us and broaden the range of insights that can be gained. Thus, data are not necessarily numbers. They can also be words, which could communicate even more information than pure numbers. Equally important is a type of data called metadata. It is also known as the data of data, which helps explain and interpret the attributes of each piece of data. Nonetheless, this is not the entire big picture to be seen with a data-literate mindset.

Among the definitions of data, though nonconsensual, one factor remains consistent that a macro understanding of data subsumes the action performed on data (McEvoy, 2018). The value seen in data that necessitates “data” literacy is not inherent, but achieved through collecting, measuring, reporting, analyzing and visualizing data (McKenna, 2018; Wikipedia, 2020). A data-literate mindset does not define data by rigid rules or restrict the concept to a fixed scope. It rather conceptualizes data according to the reason why it is gathered, the way how it is processed, and the intension of how it is to be utilized. This mindset recognizes data by answering the question, “is this as useful as what is purposed for data”, not a simple “is this data”. Classroom capacity is data to traditional HEIs, just as bandwidth to institutions only offering online programs. Data is a contextualized concept. It actually depends on the functional and analytical needs of the institution, to determine what is data, and what needs to be treated with the same standards under the umbrella of institutional data management. Building such a resilient but targeted understanding of data concept underlies the higher level of a data-literate mindset, which is to paraphrase it in the context of one's own work to provide evidence of comprehension.

3.1.2 Comprehension towards Data Relevance

A narrow perception of data leads to a misconception that data is irrelevant in one's work. Non-specialists usually feel that they do not work with data, but that is a total myth. While it is true that not everyone codes, or uses statistical software such as SPSS, they enter information in spreadsheets or systems, design surveys or webpages to collect input, and make key decisions about which data gets digitized or disposed (Tozzi, 2017). Higher education workers that manage and follow students along their registration pipeline are key to tracking student data divergences. Those that collect paper documents, such as passports, in support of faculty and staff hold the opportunities to enrich institutional digital datasets. Those that utilize operational systems to automate business processes are the designers of institutional data structures, though in most case they are not conscious of this.

A data-literate mindset brings those lightbulb moments when trainees realize that their work contains "data" as well, succeeded by a further question "to what extent and in what ways". Each administrator's work contains diverse data, while same data is of relevance to different administrative roles in diverse ways. Similar to data concept, the effort to identify data relevance but make no reference to the context is bootless. For example, to assist with visa applications, the workers on migration should gather geographic origin data of students. Wrong information would result in a rejection that compromises student's academic progress. However, the same piece of wrong information may mean nothing when kept in a spreadsheet managed by course planning and enrollment. The story gets reversed in an online university, where this same error of failing to accommodate time differences during course scheduling affects student learning just as much.

Thus, the mindset that is proposed in this framework is one that is able to identify what exists in one's work that is valuable to understanding and improving the way how one works and how the institution functions. This is to be further applied by the trainees to appraise their work through the lens of data, at the third level of Bloom's taxonomy.

3.1.3 Application in Data Quality

Modern higher education administration is greatly shaped by the increasing role of evaluation in higher education, resulted from growing competition for resources and thus demand for HEIs to demonstrate effectiveness (Heck, Johnsrud & Rosser, 2000). An essential object of the measurement towards institutional functioning is the evaluation of worker's performance, which has long been based on the worker's role and what is expected out of it (Heck, Johnsrud & Rosser, 2000). In addition to professional expertise required in specific functions, general administrative competence, such as whether rules are understood and followed, has been at the core of worker evaluation (Bider, 2008), while "data" is only treated as a by-product of administrative processes, and has little or no presence in the accountability schema of higher education administration. However, the increasing relevance of data leads to changing expectations and roles of higher education workers, and subsequently a review of evaluation standards (Flaniken, 2009). A data-literate mindset accepts that one's performance is evaluated not only through traditional measurements of administrative effectiveness, but also on the quality of data produced from or utilized in one's work. As a major force for behavioral changes, the changing mindset about the evaluation mandates of one's work underlies the changes in ways one's work is carried out.

Similar to data relevance that varies from role to role, requirements of data quality are specified in context. HEIs collect data through a variety of forms. A major source is declarative data, which is gathered through active participation of data subjects, e.g. via surveys (Hagan, 2017), such as students self-reporting their nationalities, or in an extended form, workers recording their nationalities according to information on their passports. At all events, major quality concerns of this type of data are completeness and correctness, which are to be bore by the workers handling the process. Those that work with behavioral data, however, are subject to different expectations, which involve higher levels of data collection skills and greater vigilance for misuse. Developing the accountability for data quality is a key step, but trainees are yet to be equipped with the ability to benchmark quality data, until they approach the fourth level of the data-literate mindset.

3.1.4 Analysis of Data Flow

The previous three levels focus on "why" non-specialists should care. The analysis level, however, where what is commonly thought of as critical thinking enters (Adams, 2015), starts to discuss "how" to work with data in their current capacity. A data-literate mindset is able to discern relevant data and benchmark quality of data by positioning oneself in the institutional data flow. In other words, trainees progress from knowing that data exists and data quality matters, to knowing how to find the data that exists and to improve the quality of that data. To cultivate a way of thinking leading to this goal, a major roadblock is a siloed culture in which workers operate. Higher education is a field that is no stranger to operational silos, from independently-functioning academic departments (Friedman, 2018), to traditional process-based division of functional areas (Commondore, etc., 2018), which is one of the major reasons why data is commonly viewed as a specialist work siloed in a centralized department or team. However, among the entire data life cycle, usage (or simplified to analytics) is only one of the stages (Chisholm, 2015). All stakeholders, who play a role in data life cycle, participate in provision of quality data programs, from data capturers, custodians, to analysts. These roles and responsibilities, which used to appear to be silos, are connected to each other along the institutional data flow.

Contextualization of data quality requirements in one's work, as mentioned in the third level of taxonomy, does not happen in isolation, because consequences of poor quality data may reveal themselves in latter phases of the data flow within the institution. Thus, quality data is not only data that fulfills the purpose of one's work, but also that meets institutional goals. Given prevalence of spreadsheet work within higher education administration (Laskovsky & O'Donnell, 2017), it is expected that a higher education worker who prepares a spreadsheet to be shared should have good understanding of the purpose of this spreadsheet to all intended users, and make sure it fulfills the purpose for which it is intended. One that makes modifications or additions to the spreadsheet should be aware of the risk of creating a duplicate dataset that contains outdated information, which could mislead future users. One that uses the spreadsheet should never forget to check metadata instructions and risk misinterpretation, just as “staff member” can be defined in ways distinct from intuition (Laskovsky & O’Donnell, 2017).

3.1.5 Synthesis for Data Reporting

If there is one piece of data work that is not uncommon to higher education workers, even data non-specialists, that is data reporting. Common practice of data reporting in higher education has been operated under the silo mentality at both the front and back ends. Reporting requirements overlap and sometimes duplicate across multiple levels or departments of government and agencies, but the platforms utilized by reporting entities vary (Whistle, 2017). This needless burden on HEIs to tailor and duplicate their work for each report exhausts institutional resources and operational agility for better proactive planning. Thus, instead of a streamlined process that attempts to present the institution in accurate and optimal ways, data reporting has been conducted through rough combination of data from different functional areas. The comprehensive nature of data reports that HEIs are required to complete today adds to the difficulty. However, the original intension of data reporting was to benefit both the institutions that report data and entities that use data (Whistle, 2017). During the wave of accountability in higher education, institutions are demanded to report institutional data to demonstrate effectiveness (Brown, 2017), which later becomes ways for public to know the institution, and leverage for policymakers to execute control (Heck, Johnsrud & Rosser, 2000). Though long viewed as a burden that is additional to regular administrative work, data reporting, in fact and as intended, is one of the ways data gets utilized in HEIs. A data-literate mindset recognizes data reports as obligations and opportunities to transparently present one’s work for evaluation.

Conducting data reporting in an optimal way takes a data-literate mindset that is able to clarify how data flows through the institution, and accordingly deconstruct the components of institutional data, to reconstruct the large puzzle by shuffling and rearranging these data pieces. Higher education data has been criticized for lack of adequacy and actionability (Whistle, 2017), largely due to the roughness of the way data gets prepared within HEIs. Instead of simple aggregation or calculation, higher education workers should understand meaningful ways of performing these actions. In a data report, questions need to be asked, such as “does the current classification of faculty appointment types accord with their teaching records”, or “is classroom capacity or student enrollment a better metric for this purpose”.

3.1.6 Evaluation of Data Analytics

In higher education, similar to other industries, there seems to be a natural divide between strategic planning and tactical implementation (Frølich, Stensaker & Huisman, 2017). While the significance of data on informing decision-making at strategic levels cannot be over-emphasized, it is often neglected in day-to-day decision-making that supports ground-level implementation and administration (Laskovsky & O’Donnell, 2017). It is true that typical data analytics projects in HEIs, which occupy majority of institutional data resources and capacity, aim at strategic-level topics such as admissions research, learning analytics, or program effectiveness

(Delaney 2008). However, the power of data applies equally to day-to-day decision-making at the ground level.

A data-literate mindset demystifies the perceived barriers to access, analyze and use data among non-specialists. The real truth is that workers have first-hand data about their own work that is needed to drive decisions. Most day-to-day decision-making can be conducted in a more informed way with minimum requirements of specialized data skills. Change management is prone to higher rates of success with usage of data that enables evidence-based decision-making. Managers and workers are to answer these questions with more solid roots in data, such as "where is team spending time", and "how may the business process change affect administrative efficiency". The popularized mindset towards utilization of data analytics is more important in an era of data democratization, where data becomes more accessible to non-specialist users (Laskovsky & O'Donnell, 2017).

In addition to the data democracy dividend, it is just as important for workers to mind the risks accompanying the growing awareness of the potential of data, necessitating more cautious and responsible approaches to data. A bevy of research has been done on the way users get deceived by data (Huff & Geis, 1993). Beyond ineffectiveness of conclusions, recent work has also revealed ethical concerns of inappropriate processing or use of data despite its aim that is quite the opposite (Berens, Mans & Verhulst, 2016). For example, if student financial aid is to be affected by decisions made upon data, misinterpretation of the situation leads to inequity. Attention should also be paid to rising advocacy for data privacy, such as GDPR or FERPA, which is changing the landscape of higher education data efforts.

Below is a quick reference table comparing the misconceptions commonly seen among higher education administration, and shift of those under a data-literate mindset.

Traditional Mindset	**Data-literate Mindset**
Data is statistics.	Data is an inclusive and contextualized concept.
I do not work with data.	Everyone works with data, in different ways.
Workers are evaluated by administrative effectiveness.	Worker appraisal involves evaluation of the quality of data, produced from and used in one's work, which happens in specific contexts.
Data that fulfills my needs is good quality data.	Consequences of poor quality data may be reflected in other stages of the institutional data flow.
Data reporting means providing the data I have.	Data reporting incorporates integration of data pieces into the larger institutional puzzle, and is an opportunity and obligation to display for evaluation.
Data analytics is the solution for high-level decision making.	Data analytics is a tool that is applied to all levels of decision-making, and needs to be used with caution.

Table 1: Comparison of Perceptions under Traditional and Data-literate Mindsets

3.2 Evidence and Experiences of Training a Data-Literate Mindset

Higher education institutions share similarities in functionality that allows alignment of training goals. However, ways of evaluating learning outcomes and designing learning experiences need to be catered for each specific training program. Thus, following sections focus on a general direction to developing a solid data-literate mindset training program.

Administrative skills are often taken as a given, especially in higher education settings (Bider, 2008; Laskovsky & O'Donnell, 2017). Organized professional training for workers, therefore, is usually general, optional, and flexible. Seldom does professional training incorporate systematic evaluation of learning outcomes (Klenowski, Askew & Carnell, 2006). Survey results provided by participants are commonly used as a substitute. However, this declarative data suffers from potential biases. Lack of appropriate evaluation of learning outcomes not only prevents the institution from assessing and adjusting its training programs, but also reduces motivation of participants. Among impediments to professional training in higher education administration, lack of relevance, along with imbalance between perceived gain and occupied working hours, is key to be addressed for an evaluation system to do more good than harm (Facteau et al., 1995).

Project-based assessment, which is a method to assess performance through projects, is well-suited for overcoming these barriers (Wong & Siu, 2018). It requires deployment of multiple levels of cognitive learning objectives to contribute to a cumulative project (teAchnology, 2020), and presents evidence of learning outcomes in the process of solving problems and making decisions. Flexibility in determining the topics of the projects to work on is a powerful stimulus for participants, which is an area where they usually do not have as much freedom from instructors' interference. It also allows participants to integrate the assessment projects with functional needs in day-to-day work to reduce opportunity costs. Participants' completed work that addresses instant business priorities can be put into production, which is helpful especially to institutions at early stages of building data-informed models.

A transformation of evaluation is linked to a transformation of the specific learning experiences of participants. Common practice of professional development in HEIs is still conducted in outdates ways that lack engagement of participants (Brown et al., 2015). Research has shown that employing active learning significantly improves learning outcomes (Freeman et. al, 2014; Michael, 2006). Although the discussion mainly concerns traditional students, it applies equally, or even more preferably to professional training, where participant motivation plays a larger role. Multidimensionality of active learning requires careful selection of approaches to facilitate engagement in different cognitive processes (Markant et al., 2016; Menekse et al., 2013). Further work has been done to provided vast and varied resources for training developers to design specific instructional and learning experiences. Active learning activities are compiled that instructors can refer to as techniques to engage participants and perform formative assessment, (Yee, 2019). Dimensions of learners' engagement are identified to frame appropriate selection of active learning techniques (Fredricks et al., 2004).

4. Implications and Future Work

Emergence of tools that facilitate easy access to institutional data leads HEIs farther into the unprecedented and irreversible trend of digitalization and connectivity. Nevertheless, there is wide agreement that capacity shortfalls in centralized data offices such as Institutional Research are preventing HEIs from full commitment (Swing, 2016). Derived from this conflict between rapidly increasing awareness of the power of data and slowly growing capacity of data specialists, data democratization becomes an inevitable choice of HEIs. Wider and easier access to data for everyone has established a solid foundation for data democratization. However, these efforts are greatly undermined if a data-literate mindset is not established among non-specialists, who are key players in safeguarding institutional data quality.

Drawn on previous research and practical knowledge, this working paper proposes a conceptual framework for nurturing data-literate mindset among higher education administrative workers, especially data non-specialists. The framework is aimed as a reference for higher education practitioners in their efforts to enhance institutional data through engagement of non-specialists. The paper then discusses the relevance of Backward Design Model as methodology to develop such a framework. Common misconceptions of non-specialists around data identified in practice are mapped against Bloom's taxonomy of learning objectives, as a roadmap to operationalize the framework in HEIs (Bloom & Krathwohl, 1956). Prevalence of online learning and rapidly upgrading skill requirements of data are shaping the landscapes of both data education and professional training, and undermining the effectiveness of traditional training in data literacy. The working paper also aims to advocate that higher education institutions provide context-specific enrichments to the proliferated field of data education, by exploring the uniqueness and complexities of data in higher education settings. It is as essential for higher education to become an active participant in the dialogue about data, as to be a strong data user.

Future work is to be conducted to operationalize the framework in specific contexts in HEIs, which in return would enrich the preliminary results in this working paper. In implementing the training programs, empirical evidence is to be collected and analyzed to understand whether and to what extent is such training effective to enhance institutional data efforts. It is expected that further theorization formalized around the findings may develop solid guidance for HEIs to facilitate their transformation towards data-informed models.

5. Acknowledgement

This training framework was originally conceptualized and developed at the Course Design Studio hosted by Dr. Jace Hargis, Director of Center for Teaching and Learning at NYU Shanghai

References

Abraham, I. (2016, March 30). *Digital transformation: People are the irreplaceable changemakers*. Retrieved from https://www.accenture-insights.nl/en-us/articles/digital-transformation-disruption-automation.

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association*, 103(3): 152-153.

Australian Bureau of Statistics. (2011, November 22). *Types of data sources*. Retrieved from https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4533.0Chapter52011.

Berens, J., Mans, U., Verhulst, S. (2016). *Mapping and comparing responsible data approaches.* Retrieved from http://thegovlab.org/just-released-mapping-and-comparing-responsible-data-approaches/.

Bhargava, R., & D'Ignazio, C. (2015). Designing Tools and Activities for Data Literacy Learners. In *Wed Science: Data Literacy Workshop*. Oxford, UK.

Bider, I. (2008). *Evaluating and improving quality of administration*. Retrieved from http://www.ibissoft.se/whitepapers/admquality.pdf.

Bloom, B. S., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, Handbook I: Cognitive domain*. New York, NY: Addison-Wesley Longman Ltd.

Bonikowska, A., Sanmartin, C., Frenette, M. (2019). *Data literacy: What it is and how to measure it in the public service*. Canada: Statistics Canada.

Bowen, R. S., (2017). *Understanding by Design*. Vanderbilt University Center for Teaching. Retrieved from https://cft.vanderbilt.edu/understanding-by-design/.

Brown, D., Chheng, S., Melian, V., Parker, K., Solow, M. (). *Global human capital trends 2015: Leading in the new world of work*. New York, NY: Deloitte University Press.

Brown, J. T. (2017). The seven silos of accountability in higher education: Systematizing multiple logics and fields. Research and Practice in Assessment, 7: 41-58.

Cavanagh, D. (2019, February 22). *Academic work-out: Bridging silos in student successwork*. Retrieve from https://www.utsystem.edu/sites/academy-distinguished-teachers/blog/academic-work-out-bridging-silos-student-success-work-2019-02-22.

Chisholm, M. (2015, July 14). *Seven phases of a data life cycle*. Retrieved from https://www.bloomberg.com/professional/blog/7-phases-of-a-data-life-cycle/.

Commodore, F., Gasman, M., Conrad, C., Nguyen, T-H. (2018). Coming together: A case study of collaboration between student affairs and faculty at Norfolk State University. *Frontiers in Education*, 3(39). doi: 10.3389/feduc.2018.00039

Cook, C. (2017, March 31). *Documenting DH: Brianna Marshall*. Retrieved from https://researchdata.wisc.edu/author/ccook/page/4/.

Cunha, J. M., Miller, T. (2014). Measuring value-added in higher education: Possibilities and limitations in the use of administrative data. *Economics of Education Review*, 42: 64-77.

Klenowski, V., Askew, S., Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment and Evaluation in Higher Education*, 31(3), 267-286.

"data." *Wikipedia.org*. Retrieved February 27, 2020, from https://en.wikipedia.org/wiki/Data.

"Data vs. Information." *Diffen.com*. Retrieved February 26, 2020, from https://www.diffen.com/difference/Data_vs_Information.

Delaney, A. M. (2008). Typical institutional research studies on students: Perspective and examples. *New Directions for Higher Education*, 141: 57-67.

Facteau, J.D., Dobbins, G.H., Russell, J.E., Ladd, R.T., Kudisch, J.D. (1995). The influence of general perceptions of the training environment on pretraining motivation and perceived training transfer. *Journal of Management*, 21(1): 1–25.

Flaniken, F. (2009). Performance appraisal systems in higher education: An exploration of Christian institutions. *Electronic Theses and Dissertations*, 3857.

Florian, L., Hegarty, J. (2004). *ICT and special educational needs: A tool for inclusion*. Berkshire, England: Open University Press.

Fredricks, J. A., Blumenfeld, P. C., Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59-109. doi:10.3102/00346543074001059

Freemana, S., Eddya, S. L., McDonougha, M., Smithb, M. K., Okoroafora, N., Jordta, H., Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 111(23): 8410-8415.

Friedman, H. H. (2018). *How the creation of too many academic departments stifles creativity, encourages a silo mentality, and increases administrative bloat* (Working Paper). Retrieved from https://www.researchgate.net/publication/322200199.

Frølich N., Stensaker B., Huisman J. (2017). understanding strategy practices in universities. In: Bleiklie I., Enders J., Lepori B. (Eds), *Managing universities*. Palgrave Studies in Global Higher Education. London: Palgrave Macmillan.

Grajek, S. (2019, January 28). *Top 10 IT issues, 2019: The Student Genome Project.* Retrieved from https://er.educause.edu/articles/2019/1/top-10-it-issues-2019-the-student-genome-project#data.

Hagan, K. (2017, August 22). *Declarative data versus behavioral data: Understanding the difference*. Retrieved from https://www.netquest.com/blog/en/declarative-data-versus-behavioral-data.

Harfield, T. (2017, June 13). *What is data democratization*. Retrieved from https://blog.blackboard.com/what-is-data-democratization/.

Hayhurst, C. (2019, June 20). *Breaking down data governance: Data quality*. Retrieved from https://edtechmagazine.com/higher/article/2019/06/breaking-down-data-governance-data-quality-perfcon.

Heck, R., Johnsrud, L., Rosser, V. (2000). Administrative Effectiveness in Higher Education: Improving Assessment Procedures. *Research in Higher Education*, 41(6): 663-684.

Holzweizz, P. C., Walker, D. W., Conrey, M. (2018). Preparing new professionals for administrative leadership in higher education: identifying specific skills for training, *Perspectives: Policy and Practice in Higher Education*, 23:2-3, 54-60. doi: 10.1080/13603108.2018.1543217

Huff, D. (1954). *How to Lie with Statistics*. New York, NY: W. W. Norton & Company.

Huron, American Council on Education (ACE), GeorgiaTech (GIT). (2019). *The transformation-ready higher education institution: How leaders can prepare for and promote change*. Chicago, IL: Author.

James Jacob, W., Xiong, W. & Ye, H. (2015). Professional development programmes at world-class universities. *Palgrave Communications,* 1, Article number: 15002. doi:10.1057/palcomms.2015.2

Jim, C. K., Chang H-C. (2018). The current state of data governance in higher education. *81st ASIS&T Annual Meeting Paper*. 55: 209-217.

Kim, J. (2018). *2019 Trends to watch: Higher education*. Retrieved from https://www.ellucian.com/assets/en/white-paper/2019-trends-watch-higher-education-ovum.pdf.

Knight, J. (2014, December 1). *Professional Development for Faculty and Staff in Ras Al Khaimah's Higher Education Institutions*. Retrieved from http://www.alqasimifoundation.com/en/publication/20/professional-development-for-faculty-and-staff-in-ras-al-khaimahs-higher-education-institutions.

Kubilus, N. J. (2016). Avoiding failure with higher education technology projects. *EDUCAUSE Review*, 51(4): 8-9.

Laskovsky J., O'Donnell J. (2018). Professional and support staff in higher education: Data and decisions. In F. F. Padró, C. Bossu, N. Brown (Eds.), *Professional and Support Staff in Higher Education*. University Development and Administration. Singapore: Springer.

Mameli, C., Passini, S. (2017). Measuring four-dimensional engagement in school: A validation of the student engagement scale and of the agentic engagement scale. *Testing, Psychometrics, Methodology in Applied Psychology*, 24(4): 527-541.

Markant, D. B., Ruggeri, A., Gureckis, T. M., Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, 10(3), 142–152.

McEvoy, D. M. (2018). *A guide to business statistics, First edition*. Hoboken, NJ: John Wiley & Sons, Inc.

McKenna, S. P. (2018, February 13). *Five misconceptions about data science – knowing what you don't know*. Retrieved from https://www.datasciencecentral.com/profiles/blogs/five-misconceptions-about-data-science.

Meerman, K. (2020, January 17). *How do you create a data culture in your company*. Retrieved from https://www.techzine.eu/blogs/data/44507/how-do-you-create-a-data-culture-in-your-company/.

Menekse, M., Stump, G. S., Krause, S., Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, 102(3), 346–374.

Michael, J. (2006). Where's the evidence that active learning works. *Advances in Physiology Education*, 30: 159-167. doi:10.1152/advan.00053.2006

Morrow, J. (2018). *Developing a data literate workforce: A strategy and framework for the enterprise*. Radnor, PA: QlikTech International AB.

Organization for Economic Co-operation and Development (OECD). (2017). Core skills for public sector innovation. In *Skills for a High Performing Civil Service*. Paris: Author.

Petrides, L. (2004, February 1). *The democratization of data in higher education: A case study of the challenges that institutions face as they seek to improve student success*. Retrieved from https://www.iskme.org/content/democratization-data-higher-education-0.

"Project Based Assessment." *Teach-nology.com*. Retrieved Feb 29, 2020 from https://www.teach-nology.com/currenttrends/alternative_assessment/projectbased.html.

Santos, G. D., Takaoka, H., de Souza, C. A. (2010). An empirical investigation of the relationship between information quality and individual impact in organizations. *AMCIS 2010 Proceedings*. Paper 192.

Sheng, Y., Mykytyn, P. (2002): Information technology investment and firm performance: A perspective of data quality. *Proceedings of the 7th ICIQ*, pp. 132–141.

Sternkopf, H., Mueller., R. M. (2018). Doing good with data: Development of a maturity model for data literacy in non-governmental organizations. In *Proceedings of the 51st Hawaii International Conference on System Sciences*. http://hdl.handle.net/10125/50519

Swing, R. L. (2016). *Institutional research capacity: Foundations of federal data quality*. Retrieved from http://www.ihep.org/postsecdata/resources/institutional-research-capacity-foundations-federal-data-quality.

Tozzi, C. (2017, December 11). *Five best practices for improving data quality*. Retrieved from https://blog.syncsort.com/2017/12/big-data/data-quality-best-practices/.

Whistle, W. (2017, November 1). *How higher education data reporting is both burdensome and inadequate*. Retrieved from https://www.thirdway.org/report/how-higher-education-data-reporting-is-both-burdensome-and-inadequate.

Wiggins, G., and McTighe, J. (1998). Backward Design. In *Understanding by Design* (pp. 13-34). Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. London, UK: Higher Education Funding Council for England (HEFCE). doi: 10.13140/RG.2.1.4929.1363

Wolff, A., Gooch, D., Montaner, J. J.C., Rashid, U., Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, 12(3): 9-26.

Wong, Y. L., Siu, K. W. M. (2018). The development and evolution of design education from secondary school to tertiary education. In: Wang V. X. (Ed), *Handbook of Research on Positive Scholarship for Global K-20 Education*. Pennsylvania, PA: IGI Global.

Wu, J-H., Wang, Y-M. (2006). Measuring KMS success: a respecification of the DeLone and McLean's model. *Information and Management*, 43, 728-739.

Contact email: jane.hu@nyu.edu