Establishment of data centre at Mzuzu University: A Survey of anticipations and aspirations of key project stakeholders

Felix Majawa¹ and Ralph P. Hall²

Abstract
Mzuzu University lost its Library as a result of a fire that took place on December 18, 2015. In response, the university established two processes to ensure the library services were not interrupted. The first process was to restore information services within six months by creating an interim Library. The second was to design a new library in collaboration with Virginia Tech’s School of Architecture and Design in the United States. A total of three conceptual designs were developed, from which Mzuzu University selected a final design. One key aspect of each conceptual design was a dedicated space for a data centre. The initial concept was that the data centre would support research activities at the University, within Malawi, and with international partners outside Malawi, such as Virginia Tech. This paper captures the anticipations and aspirations of the key stakeholders involved with the library design project at Mzuzu University in Malawi and Virginia Tech in the USA. Data were captured by a survey that was shared via email with 29 stakeholders. A total of 10 responded at Mzuzu University, and 12 responded at Virginia Tech. A key finding from the survey was the need to create clear plans for each aspect of the project to ensure the effective implementation of the data centre. Critical aspects to the project include staffing, equipment procurement, the management of the data centre, data literacy programming, and the long-term sustainability of the data centre. Developing a policy/process to guide the operations of the data centre was also found to be critical. The library construction began in February 2021 and is expected to end in February 2023. Having a clear plan for how the data centre could be operationalized will be essential to ensuring the centre is successful. The data centre will be a new facility for the university and this paper is a first step towards shaping the requirements of, and potential for, this new facility.

Keyword
Data Centre, Data Literacy, Data Management, Mzuzu University, Malawi, Virginia Tech

1. Introduction
An increase in research by universities and other institutions associated with research has resulted in the production of large amounts of data (Fox, 2013), advancing the need for mechanisms to generate, store, process, and use such data. One mechanism has been the development of data centres, which can be independent, providing data services to organisations on a commercial basis, or attached to specific institutions such as universities, providing data-related support functions to promote research endeavours, learning, etc.

Data has been defined as the representation of facts, figures, and concepts in a manner suitable for processing, interpretation or communication by human, or automated means while the term “data centre” has been defined “as a centralized repository, either physical or virtual, for the storage, management, and dissemination of data and information organized around a particular body of knowledge or pertaining to a particular business” (TechTarget, 2012). However, Telecity Group (2011) defines “data centres as buildings that have within them electrical and mechanical infrastructure that creates an environment in which computing and telecommunications equipment can run without interruption.” The former definition emphasizes the use of a data centre while the later emphasizes the technical equipment that enables a data centre to perform its functions. The former definition is more suitable for the Mzuzu University data centre as it focuses more on service provision and the need for data literacy.
Data Literacy means “the ability to understand and use data effectively to inform decisions” (Mandinach and Gummer, 2013). Further, Calzada et al. (2013) state that data literacy enables individuals to access, interpret, critically assess, manage, handle, and ethically use data. Also, Carlson et al. (2011) and Rin (2011) outline the necessary data literacy competencies as follows: discovery and acquisition of data; data management; data conversion and interoperability (dealing with the risks and potential loss or corruption of information caused by changing data formats); metadata; data curation and re-use; data preservation; data analysis; data visualization; and ethics, including the citation of data.

Mzuzu University (Mzuni) was established by an Act of Parliament in 1997. It has the following six faculties: Faculty of Education; Faculty of Environmental Sciences; Faculty of Health Sciences; Faculty of Humanities and Social Sciences; and Faculty of Science, Technology, and Innovation. The University has about ten thousand students. The University Library is mandated by the University Act of 1997 as an integral part of the University’s teaching, learning, and research. The library was established with the Mission “to provide up to date and relevant information resources; promote effective utilization of those resources; and facilitate rapid access to information held within and in remote places through convectional and electronic means.” Progress in pursuing goals for achieving improved library services were hampered by the fire that destroyed the library on December 18, 2015, resulting in the loss of all its collections and ICT equipment. Specifically, 53,000 books, 68 desktop computers, 403 reading chairs, 62 reading tables, 111 shelves, three heavy duty photocopiers, eight printers, and other countless valuable items. The total value of items damaged was MK 5,891,214,532 (approximately $7,854,952) (Chawinga and Majawa, 2018). Upon learning of the fire, Virginia Tech, a long-standing partner of Mzuni, stepped forward to assist with the design of a new library. The School of Architecture and Design provided support by developing three conceptual designs, from which the executive leadership at Mzuni selected one design for further development. The provision of a data centre is a critical component of the final design.

2. Literature Review

In recent years, data centres have become an essential, if almost a surreptitious, element of business and social life throughout the modern world (Jones et al., 2013). The critical functions associated with data management such as processing, storage, management, and exchange are performed in data centres, hence data centres have become the driving hub of the economy, and in some ways, of society (ibid.).

For a data centre to operate, it needs specialized equipment. Kumar (2020) outlines some of the equipment needed for a data centre to function, such as “Rack server, SAN (Storage Area Networking) & NAS (Network Attached Storage) devices, network, security, backup devices, additional services of equipment which help to support customers through NOC (Network Operation Center) & SOC (Security Operation Center) Operations.” Similarly, Jones et al. (2013) identified rack cabinets, servers, routers, switches, and air-conditioners as necessary equipment for a data centre to function effectively.

Data centres require people with specialised skills and knowledge to manage and operate them. Snipes (2018) stated that “On a broader scale, the concepts of Big Data, data-driven decision-making and data literacy have become an important part of life of librarians.” Similarly, Cox and Pinfield (2014) argue that academic library services are well positioned to play an important role in Research Data Management (RDM). Koltay and Hungary (2015) reasoned that in the same way as libraries have traditionally facilitated access to documents, they now need to facilitate access to data. However, for this to happen, new expertise is needed from experts in ICT, statistics/data analytics, data visualization, etc., expertise that is not normally available among library scientists.
Data Literacy also plays a critical role in ensuring usability for a data centre. The concept of data literacy refers to “the ability to transform information into actionable instructional knowledge and practices by collecting, analyzing, and interpreting all types of data” (Gummer and Mandinach, 2015). This definition reveals that data literacy is central to the provision of data centre services. Data literacy incorporates aspects of statistical literacy, assessment literacy, pedagogical knowledge, and data-driven decision making (Henderson and Corry, 2020). Carlson et al. (2011) and RIN (2011) outline the following range of data literacy competencies: discovery and acquisition of data; data management; data conversion and interoperability (dealing with the risks and potential loss or corruption of information caused by changing data formats); metadata; data curation and re-use; data preservation; data analysis; data visualization; and ethics, including the citation of data. These aspects are considered to be critical for the effective operation of a data centre.

3. Problem Statement
The increasing use of IT services has been accompanied by an increase in the need to process, manage, preserve, and use data in support of decision making. Data centres came into being when a large number of companies required rapid and constant internet connectivity and dedicated data storage facilities (Jones et al., 2013). The idea of establishing a data centre at Mzuzu University developed through discussions between Mzuzu University and Virginia Tech. The type of data centre being considered for Mzuzu University is unique in Malawi. Given the potential and scope of a new data centre, it was important to identify the expectations of the key stakeholders at Mzuzu University and Virginia Tech. The findings from this paper are intended to support the development of a proposal for a data centre that aligns with the ambitions of the stakeholders and provides guidance on the equipment, staffing, and manage structures needed for Mzuzu University to run a state-of-the-art data centre.

4. Objectives of the Study
The aim of this study was to investigate stakeholders’ expectations for the establishment of a data centre at Mzuzu university.

The specific objectives of the study were as follows:

(a) To identify the staffing and equipment needed to run the data centre;
(b) To establish the functions that the data centre would perform;
(c) To determine the need for data literacy in carrying out the functions of the data centre; and
(d) To identify the beneficiaries of the services provided by the data centre.

5. Methodology
Data for this study were collected using a semi-structured questionnaire that was distributed through email to Assistant Librarians, Senior Assistant Librarians, Deans, Faculty, and Directors of Research and Studies at Mzuzu University and Virginia Tech. In total, 15 stakeholders at Mzuzu University and 14 at Virginia Tech were sent the survey. A total of 10 responded at Mzuzu University and 12 responded at Virginia Tech, implying a response rate of 67% for Mzuzu university and 86% for Virginia Tech, and an overall response rate of 76%. Data analysis was undertaken using SPSS.

6. Findings and Discussion
6.1 Definition of data centre
The first area of interest was to understand how respondents understood the concept of data centre. A definition by TechTarget (2012) – who defined a data centre as “a centralized repository, either physical or virtual, for the storage, management, and dissemination of data and information organized around a particular body of knowledge or pertaining to a particular business” – was presented in the questionnaire and respondents were asked whether they agreed or not with this description. Of the
22 respondents, 15 agreed with this definition while 8 disagreed (Figure 1). One respondent was non-committal and neither agreed nor disagreed with the definition, stating “The promise of the Data Centre as a physical place should provide access to data and information relative to a broad base of knowledge as well as targeting the knowledge clusters of the partnering institutions. It should be wide-reaching and possess the capability to change, re-orient, and grow. The Data Centre’s Mission should specify the desired extent of this.” Another respondent, despite agreeing with this definition, said “I agree because it covers the different aspects of data (storage, management and dissemination). However, it does not cover data generation.” One of the respondents who disagreed with the definition stated: “The data center needs to be conceived in a much broader sense. First, the center should do much more than simply curate and share data. It needs to process, organize, and add value to the information. I also think the center should not limit itself to a specific body of knowledge. If this happens, it will greatly reduce the potential for future research/funding. I believe the center needs to be designed around a purpose/mission that is focused on unlocking the potential of data to advance positive social and environmental change.” Similarly, another responded pointed out that “It would be best not to limit the data center to a single focus. Though data on environmental science might be one area of concentration, the facility should accommodate new alternative research as it develops along with the university curriculum.” Finally, one responded observed that the issue of information security should be part of the definition.

The definition by TechTarget (2012) proved to be a useful way for stakeholders (respondents) to reflect on the concept of a data centre. However, the definition was limited in the sense that it emphasised the collection and management of data in a particular body of knowledge. Another limitation of the definition raised by a respondent is that it “does not capture the transformation/computation of data that is so important in the modern world. I would find it hard to classify Google or Microsoft’s data centers as data centers under the above definition.” This observation, combined with the multidisciplinary nature of Mzuzu University and Virginia Tech, indicates that the data centre should be able to accommodate a broad array of research and adapt to new research opportunities as they arise. Another respondent remarked, “This definition is limited to a data-centric perspective, and a more forward thinking and sustainable definition would include the organizational and policy aspects of a data center. If these critical components of a data center are not included, then the perspective may become data-centric, [losing …] the human component, which is then at risk [of becoming …] an afterthought.” In this regard, one of the respondents suggests that the definition should include the following statement: The data centre should be “organized around one or more particular bodies of knowledge or pertaining to one or more businesses.”

![Fig 1: Agreement with the provided definition of a data centre](image)
6.2 Staffing the data centre

The survey asked stakeholders about the type of staff needed to run/manage a data centre. As indicated in Figure 2, most respondents agreed that librarians should be part of the team managing the data centre. One rationale for this is that librarians are specialised in acquisition, processing, and dissemination of information and these skills could certainly be applied to the management of data. This view is supported by Snipes (2018) who states that “On a broader scale, the concepts of Big Data, data-driven decision-making and data literacy have become an important part of life of librarians. As information professionals, they need to understand data issues and how to integrate them into their work.” ICT experts were also identified as critical for managing the technical aspects of the data centre.

Researchers are both generators and consumers of data and are therefore key stakeholders of the project. However, one question raised by a respondent is whether it will be necessary to hire a researcher for the data centre. However, another respondent stated that researchers should not be employed as members of staff, but that the data centre staff should be able to work with researchers and support their data/analysis needs. Statistical experts could play an important role in supporting the analysis of data. One respondent referred to statisticians as “Experts in qualitative and quantitative data analysis.” Another respondent highlighted the need to have a Data Center Project Manager, but did not specify the skills or qualifications needed for such a position.

The need for “Marketing/communications for effective dissemination” of research, was also highlighted by one respondent. This role could as well be performed by data literacy experts. A similar role could also be played by librarians by expanding their information literacy activities. Another respondent pointed out that “the center should develop a system where advanced graduate students can work in the center to develop and use their skills and support services provided to other researchers/students.” While another respondent indicated that there was need to have computational scientists, who could also be system administrators and engineers. Another skillset identified was the need for “Data visualization experts, artists, designers [...] to help get the ‘data’ out into usable formats.” Domain/Subject experts and transdisciplinary knowledge management experts were also considered as important.
What this feedback reveals is the broad range of skills that are needed to support a successful data centre.

**Fig 2: Recommended staff for the data centre**

### 6.3 Equipment for the data centre

The survey asked respondents about the types of equipment needed to operate the data centre. All respondents agreed that equipment such as public server (available in the public cloud), computers, routers, switches (1-Gbps ethernet switch for local connectivity and connectivity to the campus network), wireless access points for local user access, rack cabinets, air conditions, storage systems, and water and smoke detectors were all considered to be important (Figure 3). Jones et al. (2013) highlight rack cabinets, servers, routers, switches, and air-conditions as necessary equipment for a data centre to function. Similarly, Lam et al. (2020) indicated that the computing resources for data centres include servers, storage and access devices, communications equipment (routers and switches), databases, and software applications.

However, some respondents added that there was need to have printers, genset for power back up, large smart screens/projectors, networking equipment, security cameras, a load balancer (to direct incoming traffic to various application nodes/servers), a telecommunication platform, a data center environmental monitoring system (e.g., sensaphone), a fire containment system that will not destroy the hardware during a fire event, and scanners for the conversion of nondigital data (pictures, notes, maps, etc.) to a digital form. One of the respondents observed that “Storage, possibly multiple levels: some amount of high speed on-line primary storage, slower but much higher capacity near-line backup storage, off-line storage (generally tape) and off-site backup storage. A means for off-site backup of critical data and a plan for disaster recovery/continuity of operations is very important. If the data is important enough to collect, it is important enough to protect.” As Kumar (2020) highlights the importance of IT equipment power utilization: “It includes power load utilization of IT equipment, i.e., Rack server, SAN (Storage Area Networking) & NAS (Network Attached Storage) devices, network, security, backup devices, additional services of equipment which help to support customers through NOC (Network Operation Center) & SOC (Security Operation Center) Operations.”
Finally, one respondent commented that “All of this equipment will be necessary. The degree of investment will be dependent on the mission. If the goal is to serve university students the investment will be modest. However, if the data center is to be a regional, national, or multi-country resource, outside support will be necessary.” This comment highlights the critical need to develop a mission, vision, and goals for the data centre that clearly outline the scope of work that the centre will support.

**Fig 3: Equipment for the data centre**

### 6.4 Functions of the data centre

The development of a clear set functions (or capabilities) of the data centre will be essential to securing funding for research and building the reputation of the centre both nationally and internationally.

As shown in Figure 4, respondents agreed that the data centre should perform a wide range of functions associated with the acquisition, processing, and management of data. Similarly, Jones et al. (2013) observed that critical functions associated with data management such as processing, storage, management, and exchange are performed in data centres, hence they have become the driving hub of the economy, and in some ways, of society. Similarly, Laughton (2012) describes the Open Archive Information System (OAIS) functionalities of a data centre such as ingest, archival storage, data management, preservation planning, and provision of access. Another function mentioned by the respondents was the importance of data ‘synthesis,’ for new horizons that develop from the intersection of seemingly disparate or different pieces of information.

However, some respondents added that functions such as data generation, ingestion, and retrieval should be shaped by the organizational structure and mission of the centre.
6.5 Data literacy topics necessary for the beneficiaries of the data centre

All respondents agreed that the data literacy topics presented in Figure 5 are important. In addition, one respondent indicated that data security is also an important topic for users of the data centre. Jones et al. (2013) also argue that the constantly evolving concern of data security should be included in data literacy programmes offered by data centres. They further state that “There are many potential threats to data centres ranging from intruders physically trying to gain access to the facility and ultimately to the servers to cyber-attacks and hackers trying to gain access to the network and to the data stored on the servers” (ibid.). Also, creation of data and metadata, long term backup and storage, offsite mirror, intellectual property and licensing, and ethics, including the citation of data, is very important since data carries risk to be used in misinformation and unethical applications. In addition, Mandinach and Gummer (2013) enumerate data literacy skills that include knowing how to identify, collect, organize, analyze, summarize, and prioritize data. In this case Mandinach and Gummer are emphasizing what people should be able to do after going through a data literacy programme.
6.6 Beneficiaries of the data centre

The survey also sought to identify the potential beneficiaries of the data centre. The most commonly cited beneficiaries were researchers, lecturers, students, and communities surrounding Mzuzu University, and institutions in and outside Malawi. In addition, the respondents also mentioned NGO’s, private companies, government institutions, and research and grant partners. One stakeholder remarked that “All of these could have access but with ‘tracking’ to make sure information is not misused.” Similarly, Henderson et al. (2014) stated that primarily, the beneficiaries of data services are the undergraduate and graduate students as well as researchers.

In relation to providing services to communities surrounding Mzuzu University, one respondent observed that such action would need to be based on “a policy decision. Access to the [data centre …] services will give the community greater opportunities and increase the general well-being of the population, besides providing talent for working at Mzuzu University.” This statement reveals that there is need to develop a data centre policy to guide the operations of the centre. Another respondent alluded to the fact that the communities should be able to use the centre when they collaborate with staff/faculty at the University. One respondent indicated that “private companies represent a special case regarding revenue and expenses, though any group outside the university will need an agreed contract.” However, the respondent cautioned that there is need to make sure that all intellectual property created by the center is carefully managed. Another respondent further observed that extending services to institutions in and outside Malawi will be essential for generating financial resources for the center.

The above feedback reveals the need to identify the costs of providing services and the need for a clear plan to manage the revenue generated by the data centre. While it is anticipated that the data centre will not be focused on profit making, it will need to generate sufficient funds to offset the costs of providing services. Jones et al. (2013) highlight that “in the face of rising costs and seemingly ever more sophisticated technological innovation, many companies have not been in a position to develop and manage their own data centres. There is need therefore to find ways of sustaining the functionalities of the data centre.”

Fig 5: Data literacy topics
The final section of the survey focused on understanding potential institutional arrangements between Mzuzu University and the data centre. A majority of respondents (13 out of 22) indicated the data centre should be located under the leadership of the library. This position is supported by Snipes (2018) who states that “On a broader scale, the concepts of Big Data, data-driven decision-making and data literacy have become an important part of life [...] librarians.” Similarly, Cox and Pinfield (2014) stated that several commentators have proposed that academic library services are well positioned to play an important role in Research Data Management (RDM). However, survey respondents also mentioned that the data center should serve the university and be advised by a board or panel of key actors/stakeholders. Such a board/panel could act as an important policy decision making body for the data centre, while allowing the data centre to be managed by library leadership. However, six of the respondents opted for the centre to be under the Directorate of Research, while three opted for the centre to be independent reporting to the Vice Chancellor. However, one respondent observed that “Being independent directly to VC is risky unless monetization is [a] primary goal.”

Similarly, other respondents had different opinions regarding the position of the centre in the University. “Why not a panel with representatives of all the entities? This will keep it from becoming political. Rather than centralized authority, perhaps an ‘oversight committee’ or ‘work group’ comprised of important stakeholders, [...] Librarian, Director of Research, VC, members of key academic departments, technical professionals, etc.”
7. Conclusion and Recommendations

This exploratory study reveals a number of important issues that need further consideration with regards to establishing a data centre at Mzuzu University. These issues range from defining what is meant by a data center, to considerations of the expertise needed to manage the centre, what the centre’s core functions should be and what this means for equipment, what data literacy programs might need to accompany the launch of the centre, how the centre will be positioned institutionally within the university, and who the target beneficiaries will be.

The key observations from this study are as follows:

- **Defining the data centre**: While there are different perceptions of the concept of a data centre, adopting a broader view of the scope of the data centre would position it to engage in a wide range of research and educational opportunities. These opportunities would also enable the centre to diversify its potential revenue, creating a more robust financial model.

- **Data centre expertise**: The operation of the data centre would require people with different types of expertise, ranging from ICT, statistics, data management, etc. Library staff with skills relating to the curation of data and communication were also identified as important. In addition, the role of post graduate students in supporting the services offered by the data centre was highlighted as way to enable them to use and develop their skills.

- **Equipment**: A broad range of equipment needs were identified that included servers, computers, routers, switches, rack cabinets, air conditioners, and water and smoke detectors. However, equipment for off-site backup of critical data and a plan for disaster recovery were also deemed important tools for the centre.

- **Data centre services/functions**: Several functions of the data centre were revealed through the study in relation to the generation, processing/analysis, visualization, storage, and dissemination of data. The role of a data literacy programme was viewed as essential for enabling data centre users to fully benefit from its services. One recommended subject to include in such a programme is data security.

- **Data centre beneficiaries**: The stakeholders revealed different groups of people who could benefit from the centre. However, there is need to first establish a policy that clearly outlines who can use the centre and what the related costs may be, so revenue generated can offset some of the costs of running the data centre.
• **Institutional location of the data centre**: Given the physical location of the data centre and the alignment between the services it could provide and those offered by Mzuzu University’s library, it is recommended that the data centre be positioned under the library.

The above points reveal a broad range of issues that need serious consideration if Mzuzu University and its partners are to realise the full potential of the data centre. While construction begins on the new library, Mzuzu University has a unique window of opportunity to develop a clear plan for creating a state-of-the-art data centre. It is recommended that a team of administrators, faculty, and staff at Mzuzu University and Virginia Tech is convened with the charge of developing this plan. An important aspect of this effort would be to identify the initial areas of expertise at Mzuzu University and Virginia Tech, where faculty could come together around joint research projects that leverage the data centre’s services/capabilities. The identification of these synergistic research opportunities should also help shape the type of services provided by the data centre.

On the other hand, Mzuzu University needs to consider collaborating with other Universities in this project as the data centre could also support research data services for other institutions. One such institution is the Kamuzu University of Health Sciences, which could benefit from sharing the services provided by the data centre. However, currently the project is being undertaken in close collaboration with Virginia Tech, thereby providing an opportunity to share data in the future when the project is realised.

**References**


---

**Endnotes**

1 Felix Majawa, Librarian, Mzuzu University, Malawi. Email address: majawa.f@mzuni.ac.mw

2 Ralph P. Hall, Associate Professor and Associate Director of the School of Public and International Affairs (SPIA) at Virginia Tech, USA. Email address: rphall@vt.edu