Plan Overview

A Data Management Plan created using DMPonline

Title: DEVELOPMENT OF A PROJECT RISK MANAGEMENT FRAMEWORK A STUDY OF AFRICAN MAJOR ECONOMIES

Creator:BABATUNDE DOSUMU

Principal Investigator: BABATUNDE DOSUMU

Data Manager: BABATUNDE DOSUMU

Affiliation: University of Manchester

Template: University of Manchester Generic Template

ORCID iD: 0000-0001-7285-3297

Project abstract:

ABSTRACT This research is investigating study project risk management within African major economies and then attempt to develop a project risk management framework. Existing literature of similar and relevant works will be examined that will address the aim and the objectives, and evolve to the development of well-structured questionnaire combined with interview that will aid collection of data on the topic which will now be analysed using appropriate tool(s) that will assist in arriving at the conclusion(s) and recommendation(s). The overall results will lead to the development of a framework that will assist in the development of the construction industry within major African economies, and should pave way for further re-orientation of practitioners within the built environment within the study area. Thus, the principal research question of this investigation is: How are political risk, security risk, corruption and other risk impacting on cost variability on construction projects within African major economies, and how can they be mitigated? Major construction under the current scenario, Projects in major African economies (outside the United Kingdom and the EU zone) executed at variable cost, which justifies the need to collect data. This region of the construction industry. As needed data for this task Research needs to be gathered from outside the EEA for a complete analysis of the UREC. The obtaining of an ethical approval is justified. In view of this peculiarity, this research aims to carry out an in-depth analysis of how the threat of the project has been identified, estimated and mitigated. The data collection is therefore aligned with instrumental approach to case study and qualitative in nature, which is face-to-face semi-structured interviews (audio-recorded and transcribed), observation notes and collection of related documents. Such triangulation/corroboration in data collection will ensure the reliability of collected data by enabling fact checking, along with in depth understanding of the issues. The exact questions to be asked during the interview process can be found on the Interview Protocol. Similarly, the location and objects of observation have been detailed in the Observation Protocol. Both protocols have been submitted with the application for ethical approval, which this data management plan is a part of. It is expected that the interviews will be conducted, transcribed and analyzed in English language.

ID: 59865

Last modified: 16-11-2020

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

DEVELOPMENT OF A PROJECT RISK MANAGEMENT FRAMEWORK A STUDY OF AFRICAN MAJOR ECONOMIES

Manchester Data Management Outline

1. Will this project be reviewed by any of the following bodies (please select all that apply)?

- Ethics

2. Is The University of Manchester collaborating with other institutions on this project?

- No - only institution involved

3. What data will you use in this project (please select all that apply)?

- Re-use existing data (please list below)
- Generate textual supporting information only
- Acquire new data

- Acquire new data
- Re-use existing data (please list below)

This research is investigating study project risk management within African major economies and then attempt to develop a project risk management framework. Existing literature of similar and relevant works will be examined that will address the aim and the objectives, and evolve to the development of well-structured questionnaire combined with interview that will aid collection of data on the topic which will now be analysed using appropriate tool(s) that will assist in arriving at the conclusion(s) and recommendation(s). The overall results will lead to the development of a framework that will assist in the development of the construction industry within major African economies, and should pave way for further re-orientation of practitioners within the built environment within the study area. Thus, the principal research question of this investigation is: How are political risk, security risk, corruption and other risk impacting on cost variability on construction projects within African major economies, and how can they be mitigated?

Major construction under the current scenario, Projects in major African economies (outside the United Kingdom and the EU zone) executed at variable cost, which justifies the need to collect data. This region of the construction industry. As needed data for this task Research needs to be gathered from outside the EEA for a complete analysis of the UREC.

The obtaining of an ethical approval is justified.

In view of this peculiarity, this research aims to carry out an in-depth analysis of how the threat of the project has been identified, estimated and mitigated. The data collection is therefore aligned with instrumental approach to case study and qualitative in nature, which is

face-to-face semi-structured interviews (audio-recorded and transcribed), observation notes and collection of related documents. Such triangulation/corroboration in data collection will ensure the reliability of collected data by enabling fact checking, along with in depth understanding of the issues. The exact questions to be asked during the interview process can be found on the Interview Protocol. Similarly, the location and objects of observation have been detailed in the Observation Protocol. Both protocols have been submitted with the application for ethical approval, which this data management plan is a part of. It is expected that the interviews will be conducted, transcribed and analyzed in English language.

.Acquiring New Data:

Professionals related to and/or working in the construction industry within major African economies will be interviewed - especially, the participants who are in direct exposure to decision making and running of construction works (e.g. project managers, architects, quantity surveyors, engineers, contractors etc.). Observation notes will accompany each interview (if the participatory organization and participant consent to provide guidance in facilitating observations), which will record observations regarding the information collected during the interviews. The recorded (through notes) observations then will be checked to be in support or contradiction to the information collected through the interview and publicly available/collected non-public documents. The interviewee will act as the guide to ensure deep understanding from the observational data. The audio recordings and transcripts of these interviews and the observation notes will be the 'newly acquired data'.

Pre-Existing Data:

The pre-existing data in this research consists of two main types.

- Firstly, the publicly available documents to be collected and analyzed are: international and local legislations, government and reports as well as grey literature (websites of related companies and news articles).

- Secondly, access to non-public documents will also be requested from construction companies and relevant agencies. Such documents will comprise of but not limited to - contract documents, final account and strategy documents, organizational and project structures, case studies of specific projects, documents explaining procedures of the key activities, and documents related to the management of construction contracts. Sharing of such documents in optional for an organization or participant's participation in the study (permission from proper authority in the organization for participant's sharing of such documents and participant's own consent will be managed through consent forms). Also, no document which the participant/participatory organization deems to be sensitive or confidential will be collected, and the participants themselves will judge the sensitivity/confidentiality of requested documents. The collected documents will only be used for research purposes and will not be made public.

4. Where will the data be stored and backed-up during the project lifetime?

- P Drive (postgraduate researchers and students only)
- University of Manchester Data SafeHaven
- University of Manchester Research Data Storage Service (Isilon)
- University of Manchester Research Data Storage Service

The research data will be stored securely in the University of Manchester's Research Data Storage service (RDS). This is university provided data storage, which is registered and linked with the Principal Investigator's University of Manchester email address, shared from the PhD supervisor's (and data custodian) storage space of 8TB. Intended to use only 1 TB maximum, this will be the primary storage, working and backup space for all the research data. The data custodian (the main PhD supervisor) will be provided access to the data from this space - through their own University of Manchester email addresses.

5. If you will be using Research Data Storage, how much storage will you require?

- < 1 TB

The interview recordings [audio will be in two formats - a) FLAC, and b) MP3, the industry standard for compressed audio files, for regular usage and/or sharing purpose] will require the bulk of the storage space. Interviews with a maximum of 60 individuals are expected (more likely 35-40). One hour of MP3 audio (44.1 kHz, 128 kbps) is estimated at 54.84 MB; and one hour of FLAC audio (44.1 kHz, 24 bits) is estimated at 334.8 MB (see dsd-guide.com/size-comparison-chart-various-formats-dsd-wav-flac-mp3). Considering a maximum of 3 hours of audio data per interview (more likely 90-100 minutes), the estimation of necessary storage for audio is about 72 GB.

In addition, the metadata file, transcripts of the interviews, consent form, interview protocol, observation notes and documents collected from the participating organizations will also be preserved. While PDF format for storage is being prioritized, all collected documents will be stored in original file format - TXT, DOC/DOCX, XLS/XLSX and JPG/JPEG formats are expected. However, the file size for these document files is likely to be comparatively smaller than the audio files, estimated at a maximum of 8 GB total.

Thus, a maximum of 80 GB of storage (less than 1 Terabyte) will be necessary.

6. Are you going to be working with a 3rd party data provider?

- No

7. How long do you intend to keep your data for after the end of your project (in years)?

- 5 - 10 years

Questions about personal information

Personal information, also known as personal data, relates to identifiable living individuals. Special category personal data is more sensitive information such as medical records, ethnic background, religious beliefs, political opinions, sexual orientation and criminal convictions or offences information. If you are not using personal data then you can skip the rest of this section.

Please note that in line with data protection law (the General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it should be pseudonymised (partially de-identified) and/or anonymised (completely de—identified) as soon as

practically possible. You must obtain the appropriate ethical approval in order to use identifiable personal data.

8. What type of personal information will you be processing (please select all that apply)?

- Personal information, including signed consent forms
- Pseudonymised personal data
- No sensitive or personal data
- Audio and/or video recordings

● Audio and/or video recordings
● Pseudonymized personal data

The only person identifying data collected in this research project will be the person conducting the face-to-face interviews with the participants, which will be audio recorded and transcribed. The personal information will include details about projects in terms of initial cost, final cost, types of risk causing variation, implications of such risk and how the risk can be controlled. without mentioning the names of the suppliers of such information because they will be coded To understand the viewpoint of the participants regarding threats to construction projects (Risk), they will be requested to explain their background (nativity, education, professional career etc.). Response to such queries may contain information related to the participant's current and previous professional positions (name of Employer Company, the participant's role and responsibilities etc.), gender etc. Such information can essentially become person identifying, despite the interview transcripts being pseudonymized. [Voice alteration of the recorded audio step has been removed.]

9. Please briefly outline how you plan to store, protect and ensure confidentiality of the participants' information.

To be able to ensure the confidentiality of participants information,having effective processes and methodologies in place to maintain data integrity is essential.

- UK Data Protection Act will be complied with.
- Data will be collected using code that is known to the principal investigator alone so as to protect the person giving out the information
- The data will be adequately and securely protected.
- Good data protection management having effective processes and methodologies will be put in place to maintain data integrity.
- Collected data will be securely stored at the stipulated university domain
- Data will be wiped off at the stipulated period of safe keeping.

To store, protect and ensure the confidentiality of the participants' information, the following steps are being undertaken:

1. Consent will be taken from the participants, they will be informed in details about the research design through Participant Information Sheet (PIS),which will explain clearly that their personal information (other than their gender and role in the company) will not appear in any published material the project may produce. PIS will also inform the participants that, they have the right to withdraw from the project at any time if they want, in which case data collected from them will be erased securely.
2. Collected data will not be created at a personal non-encrypted device. As soon as data backup is completed, the recordings will be deleted from the device. Any paper form data will be kept securely locked in a University provided cabinet. During fieldwork within African major economies, a university provided portable document scanner (Fujitsu ScanSnap S1300i) will be used to directly store collected documents in the Research Data Storage through an encrypted laptop; no hard/paper copy of the documents will be collected/stored during fieldwork i.e. will be returned to the participant.

The interview transcriptions will be pseudonymized.

In addition, any person/organization recognizable information in the collected documents will be 'blacked out' (utilizing Adobe Acrobat). Such measures will be taken at the earliest possible stage to decrease risk.

[Voice alteration of the recorded audio step has been removed. 'Sensitive' replaced by 'person/organization recognizable', as no documents will be collected containing information which are sensitive/confidential to the organization, as judged by the participants themselves.]

1. Data is being stored only in secure storage - University of Manchester Research Data Storage (RDS). The storage period will be the minimum possible time to adhere to the university policy (of preserving data for at least 5 years after the latest publication from this PhD research project). Only the Principal Investigator and Data Custodian will have access to the complete data and will be capable of linking data with the source (i.e. access to pseudonymization information). Even, in the RDS storage, data will be kept encrypted (utilizing the 7-Zip software, further described below).

For complete detail of the data collection, processing and storage plan, please see the 'Data Collection' section of this Data Management Plan.

10. If you are storing personal information (including contact details) will you need to keep it beyond the end of the project?

- Yes – Other

11. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester?

- No

12. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?

- Not applicable

13. Are you planning to use the personal information for future purposes such as research?

- Yes

14. Who will act as the data custodian for this study, and so be responsible for the information involved?

Dr Obuks Ejohwomu and Dr Akilu Yunusa-Kaltungo [Main supervisor and co supervisor of the PhD Research Project]

15. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).

2020-11-14

Project details

What is the purpose of your research project?

This research is investigating the types, causes, effects and of construction risk resulting into cost variability within African major economies, such as to develop a framework to mitigate the effects and be able to prepare a more realistic cost estimate. This issue of investigation incorporates research gap from both theoretical and contextual perspectives. Thus, this study aims - to assess the risk impacts on the variability between contract sum and final account with a view to developing a framework that will assist the construction industry professionals to have a better prediction of the final account (actual project cost) from the contract sum, Within African major economies, and how accurate construction cost estimate can be generated at the very early stage of construction. The data collection is for this research project is intended:

1. to identify the key risk factors
2. to examine how such risk factors affects cost variability
3. to develop a framework that will assist in the preparation of accurate cost estimates.

What policies and guidelines on data management, data sharing, and data security are relevant to your research project?

The data collection and management will not contravene the data and information protection act,data on cost variability in the construction industry within African major economies will be considered, and should comply with:

- Research Data Management Policy
- Records Management Policy
- Data protection policy

- Intellectual property policy
- IT policies and guidelines
- All should comply with existing procedures and guidelines of the university complying with stated formal standards.
- **And the following**:
- University's Code of Good Research Conduct. Link to code: documents.manchester.ac.uk/display.aspx?DocID=2804
- The University of Manchester Research Data Management Policy. Link to the policy: documents.manchester.ac.uk/DocuInfo.aspx?DocID=33802
- Standard operating procedure 'Taking recordings of participants for research projects'. Link to the SOP: documents.manchester.ac.uk/display.aspx?DocID=38446
- Standard operating procedure 'Bring your own technology'. Link to the SOP: documents.manchester.ac.uk/DocuInfo.aspx?DocID=31417
- UoM List of High-Risk Countries (Last Reviewed February 2020). Link to the list: documents.manchester.ac.uk/display.aspx?DocID=42983
- University Health & Safety Arrangements - Chapter 24 - Health & Safety in Fieldwork (including field trips). Link to the procedure: documents.manchester.ac.uk/display.aspx?DocID=15496
- UoM Safety Services Guidance - The Role of Generic and Dynamic Risk Assessments. Link to the guidance: documents.manchester.ac.uk/DocuInfo.aspx?DocID=10129
- University Health & Safety Arrangements: Chapter 10 - Lone working. Link to the procedure: documents.manchester.ac.uk/display.aspx?DocID=13891
- UoM Safety Services Guidance - Guidance on lone working. Link to the guidance: documents.manchester.ac.uk/display.aspx?DocID=13644
- The University of Manchester Data Protection Policy. Link to the policy: documents.manchester.ac.uk/display.aspx?DocID=14914
- The University of Manchester Records Management Policy. Link to the policy: documents.manchester.ac.uk/display.aspx?DocID=14916
- The University of Manchester Information Governance Office Records Retention Schedule. Link: documents.manchester.ac.uk/display.aspx?DocID=6514
- The University of Manchester Publications Policy. Link to the policy: documents.manchester.ac.uk/DocuInfo.aspx?DocID=28526
- The University of Manchester Intellectual Property Policy. Link to the policy: documents.manchester.ac.uk/display.aspx?DocID=24420
- UKAN's Anonymization Decision-Making Framework. Link to guideline: net/wp-content/uploads/2015/05/The-Anonymisation-Decision-making-Framework.pdf

Responsibilities and Resources

Who will be responsible for data management?

- The principal investigator, DOSUMU Babatunde, will be primarily responsible for collecting, processing, storing and updating the data as necessary.
- The main PhD supervisor, Dr Obuks Ejohwomu and co supervisor Dr Akilu Yunusa-Kaltungo will also be responsible in this regard.

What resources will you require to deliver your plan?

The primary investigator will himself collect, process and store the data as a part of his PhD research project.
Acceptable and recognised resources which includes the following will be required:

1. Storage: The University has already provided with the storage options – University of Manchester Research Data Storage – with adequate storage space (>1tb) to execute this data management plan.
2. Software: For the purpose of data management the software which will be used are – Microsoft Outlook, Microsoft Word, Microsoft Excel, NVivo, NCH Wavepad Audio Editor, 7-Zip and Adobe Acrobat. All these software are either already provided access to or freeware for non-commercial usage/academic purpose.
3. Devices:

- The University will provide a recording device with ample storages space (2 x 32 GB Micro-SD cards) to collect at least 10 3-hours long interviews continuously [also cable, charger and rechargeable batteries for the device].
- The university has also provided a password-protected Laptop with enough computation capability and storage space to undertake the research work.
- The researcher will use a university provided portable document scanner (Fujitsu ScanSnap S1300i) for scanning collected documents in field.

Finally, a personal earphone will be used to listen to audio recordings and transcription purpose.

Data Collection

What data will you collect or create?

Data on causes and effects on construction cost variability caused by risk in construction within African major economies will be collected, while a framework for accurate cost estimation evolving from data analysis will be created.

- This research intends to conduct an in-depth examination on how major construction risk factors not limited to Security, corruption political among others contributes to cost variability and estimating inconsistencies, and to proffer solution to the effects occasioned by the factors. In this mixed study, seeking generalizations/reproducibility is not the main objective - rather, practices, real life visible issues will be considered. Thus, the data collection is aligned with an instrumental case study approach and qualitative in nature, which are: face-to-face semi-structured interviews (audio-recorded and transcribed), observation notes and collection of related documents. Such triangulation of data collection will ensure the quality of collected data by enabling fact checking, also enabling in depth understanding of the issues.

Acquiring New Data: For the interviews, the preferred method is face-to-face; if this is not feasible conduction through Microsoft Themes will be attempted. [Telephone removed as virtual data collection method.] The face-to-face interviews will occur in the participant's workplace and the time will be inside normal working hours. In either method, face-to-face or virtual, the interviews are expected to be 1-2 hrs. Long. The interviews will be conducted, transcribed and analyzed in English language to ensure no loss of meanings, and only the pseudonymized segments presented as evidence in publications. The Interview Protocol comprises the broad-questions and probes to be used during the interview. Observation notes will accompany each interview (if the participatory organization permits observations in the organization premise and participants consent to provide guidance in facilitating observations), which will record observations regarding the information collected during the interviews. The recorded observations (through notes) then will be checked to be in support or contradiction to the information collected through the interview and publicly available/collected non-public documents. The audio recordings and transcripts of these interviews and the observation notes will be the 'newly acquired data'.

Pre-Existing Data: The pre-existing data in this research consists of two main types. Firstly, the publicly available documents to be collected and analyzed are: international and local legislations, government reports, as well as grey literature (websites of related companies and news articles). Secondly, documents will also be collected from the Identified and major construction outfits and consultants on such major projects. Such documents will comprise of but not limited to – Contract documents, general policy and strategy documents, organizational and project structures, case studies of specific projects, documents explaining procedures of the key activities, Final account of projects

How will the data be collected or created?

The data will be collected using ;

- Case study
- well structured questionnaire
- interviews and focused group discussion which may include audio recording with the consent of those to be interviewed asking open ended questions.
- The interview will be transcribed into word text
- Relevant peer reviewed data will also be used when available and appropriate.
- Data will be coded in a way that providers of the data will not be identified by everyone except the principal investigator for confidentiality purpose.

For easement of data collection, the researcher scoped the construction industry within African major economies by attending an international conference where stakeholders within the construction industry converged at Abuja the capital of Nigeria in February 2020. The visit has been fruitful in creating necessary connections. Through a series of meetings with reputable contractors and consultants within the construction industry who had worked or still working on mega projects. The researcher has been successful in drawing up a map of stakeholders within the However, it is being clarified that, the stakeholders are free to decline participation. This aspect has been made clear in the 'Participant Information Sheet', which will be provided to the participant at least 24 hours before the interview; to ensure that no coercion takes place. Through the scoping visit, the researcher has been able to identify 43 preliminary participants, which is being expected to only increase through snowball sampling. Such participants could be grouped inside three major categories:

- Consultants who had worked or still working on projects of high magnitudes – 20 participants
- participants working in organizations who are clients or contractors on mega projects-15
- participants working in professional regulatory agencies – 8 participants

Thus, the initial recruitment of participants is premised around this map of recognized potential participants. For this purpose, the researcher is acquiring a letter of request from the Head of PGR, Dept. of MACE, UoM to the identified participants. Also, the researcher has 22 years of work experience in the construction industry in Nigeria and his own connections are also working as another source of participant recruitment. It is being expected that from such contacts, new possible participants will be recognized,

through snowball sampling method. In such cases, first contact will be through an ‘Introductory Email’ (text of the email attached with the application for ethical approval).

The data collection, processing and storage method for this study has been designed to ensure complete adherence to the University of Manchester Policies regarding Data Recording (especially audio data), IT Policies and Guidelines, Data Management, Record Management, Data Protection, Publication and Intellectual Properties.

The University of Manchester Research Data Storage (UoM-RDS is being utilized as the primary secure data storage option for this research project. This is university provided data storage, registered and linked with the Principal Investigator’s University of Manchester email address, shared from the PhD supervisor’s storage of 8TB. Intended to use only 1 TB maximum, this storage will store, process and backup all the research data. The data custodian and the PhD project supervisors will be provided access to the data collected from this space - only through their own University of Manchester email addresses.

For the purpose of data collection, management and analysis, the software which will be used are – Microsoft Outlook, Microsoft Word, Microsoft Excel, NVivo, NCH Wavepad Audio Editor, 7-Zip and Adobe Acrobat. All these software’s are either already provided access to by the university or freeware for non-commercial usage/academic purpose. As the data collection will be outside the UK, the researcher will encrypt his personal laptop (through the University's IT support service), which will be used for research work, data management and uploading data to secure storage purpose.

The devices which will be utilized for the data collection purpose are:

- A personal laptop (MacBook Pro), encrypted by the university
- A sound recorder (Zoom H5) and two 32 GB memory cards provided by the university [also cable, charger and rechargeable batteries for the device]
- A university provided portable document scanner (Fujitsu ScanSnap S1300i) for scanning collected documents in field.
- An earphone (to listen to audio recordings and transcription purpose).

While organizing data, all files and folder names will contain only characters in the alphabet, the number, and the underscore (_) symbol. This convention is to ensure the maximization of understandability and compatibility.

The planned folder and files structure inside UoM-RDS for this stage of data collection is:

Level 1:

[Research_Data]
[Digitized_Research_Diary]
Metadata.TXT

Level 2: Inside [Research_Data] Folder

[A_Processed_Dataset]
[B_Field_Audio_Recordings]

Level 3: Inside [A_Processed_Dataset] folder

[Company_X]
(X = 1-n)

Level 4: Inside each [Company_X] folder

COMPANY_X1_Participant_YY_CF.PDF
COMPANY_X2_Participant_YY_IP.PDF
COMPANY_X3_Participant_YY.FLAC
COMPANY_X4_Participant_YY.MP3
COMPANY_X5_Participant_YY_ON.DOCX
COMPANY _X6_Participant_YY_TR.DOCX

(YY = 01, 02…; separate for each participant; CF is Consent Form, IP is the Interview Protocol used for the specific interview, ON is Observation Notes; TR is Transcript of the interview.)

COMPANY_X_D.PDF

(D is Document Details; ex. Waste_Report, Work_Schedule etc.)

The complete data collection, storage, processing and backup includes the following steps and measures:

1. The data collection process starts with communication with the participants. All the communications through emails regarding the research will be organized using the ‘folders’ feature in Microsoft Outlook – university’s preferred email client. The name of the overall folder will be ‘Research_Data_Collection’. Communication with participants from different companies will be kept in separate sub-folders.
2. Any communications with the participants prior to and after the interviews out of the emails will be noted down in a separate research diary. On each of such notes, the date and time of communication will be recorded. This diary will always be kept securely locked in the university provided cabinet when not in the presence of the principal investigator. New information added to this diary will be digitized as frequently possible (at least once a week) during the ongoing data collection phase and stored in the UoM-RDS in the [Digitized_Research_Diary] folder in PDF format.

Company names and Participant IDs will be recorded as pseudonyms in this diary and throughout the research data [Company Names as A, B, C… etc. and Participant IDs as 01, 02… etc.] to ensure that loss of this diary or breaching of digital data does not result in disclosure of participants personal/person-identifying information. Participant IDs will be continuous throughout the data collection (will not restart for a different company).

All pseudonymization data (to link interview ID with the participant’s information) and will be recorded in a separate password-protected TXT file (named Pseudonymization_Information.TXT) with password-access to the Principal Investigator and Data Custodian only. This file will be separately stored in the University P Drive only. This file will have the following information:

- The company indicated by X [A, B, C… etc.]

- The participant indicated by YY [01, 02, 03… etc.]
- Participant details [Name, Company working for, Working Position, Gender and Contact Information)

1. A specific time, method (face to face or through Microsoft theme) and/or place of interview will be agreed between the participant and the researcher. The interview schedule will be maintained by Microsoft Outlook's calendar feature. In case of face-to-face interview, the interview will occur in the participant's workplace and the interview time will be inside normal working hours. At the start of the interview, before recording commences, a 'Participant Information Sheet' will be provided to the participants; which will contain information related to the research, research design and usage of the collected data. Then, the consent form will be signed by the participant providing consent to start the interview and recording. The name of the interviewee will not be recorded in audio. To ensure a good quality of data collection it is intended that, no more than 02 interviews will be conducted per day – each lasting a maximum of 3 hours.

Only the Pseudo-IDs (X - company ID – A, B… and YY – Participant ID – 01, 02…) will be recorded in the Interview Protocol used by the researcher (no name/designation of the participant will be recorded) to take notes and follow predesigned 'queries, probes and prompts' during the interview.

1. The interviews will be audio recorded. The device will always be kept with the interviewer. The recording device is the university provided Zoom H5 audio recorder. Audio data will be recorded in WAV format in this device. Considering recording in 44.1 kHz and 16 bits format – each hour or audio data in WAV requires an estimated 604 MB of space, thus a memory card with 4536 MB (4.42 GB) of space or greater will suffice. The university will provide two memory cards of 32 GB of capacity for recording purpose in case of failure during data collection in field. The card will be formatted in the FAT32 file system as this ensures interoperability in all Windows, MacOS and Linux operating systems. Though, the FAT32 file system has a limitation of a maximum of 4 GB file size - this is unlikely to be reached by any single interview recording.

During data collection on the field, the researcher will carry a university encrypted personal laptop. After each interview is over, the recorded audio data (in WAV format) will be uploaded to the [B_Field_Audio_Recordings] folder in UoM-RDS in the earliest scope, through the encrypted laptop. After ensuring data is properly uploaded and checked for playing normally, the memory card will be 'Formatted'. After the data collection period ends, the voice recorder will be returned to the university. These audio recording files will be renamed as COMPANY_X_Participant_YY.

1. The additional documents requested from the participant in support of different topics discussed during the interview will be collected in whichever form provided by the participant - digital (through email) and/or hard copy (paper). In the case of documents collected through an email, the communications will be sorted in the respective pre-created 'Microsoft Outlook sub-folders' for each company.
2. The consent form, interview protocol and other documents collected from the participant in paper format will be kept in possession of the researchers in a locked cabinet placed in the University. At the earliest possible time, these documents will be digitized and uploaded to the UoM-RDS through the university encrypted laptop computer. In case of fieldwork in Africa, where such secure cabinet is not available, a University provided portable scanner (Fujitsu ScanSnap S1300i) will be used to scan the paper documents in field and stored directly in RDS through the encrypted laptop. No paper format document will be stored i.e. will be returned to the participant. The files will be scanned at Color, 600 DPI resolution without any automatic image correction provided by the scanning machine to ensure accurate data preservation. Immediately after scanning, these files will be passed through an OCR (Optical Character Recognition) process using Adobe Acrobat to create searchable PDF files with an invisible text layer, again, ensuring accurate data preservation and sustainability.

The consent form and interview protocol for each interview will be digitized in PDF format, renamed COMPANY_X_Participant_YY_CF and COMPANY_X_Participant_YY_IP respectively and uploaded to UoM-RDS in the earliest possible opportunity in the respective Company_X folder.
For the other documents collected in digital form, the original files format will be retained as long as they are commonly readable i.e. PDF, DOC/DOCX, XLS/XLSX, PPT/PPTX, JPG/JPEG/GIF/TIFF. In case of any other format, the files will be converted to PDF to maximize usability and sustainability of data. Any document collected as hard copies will be digitized in PDF format. These files will be renamed as COMPANY_X_D (where D is Document Details; ex. Work_Breakdown_Structure, Waste_Report, Work_Schedule etc.) and saved in the associated Company_X folder.

1. At this stage, the field audio recordings, preferably from the WAV files (recorded through the Zoom H5, the primary recording device) will be processed in Wavepad Audio Editing Software, which is freeware for non-commercial usage (nch.com.au/wavepad). In this software, any segment of recordings before and after the actual interview will be deleted, de-noised to improve audible clarity, the volume first amplified and then normalized. These files will be saved and uploaded in two formats: FLAC (44.1 kHz 24 bits) and MP3 (44.1 kHz 128 kbps) in Research_Data/A_Processed_Dataset/Company_X directory, renamed as COMPANY_X_Participant_YY. As FLAC is a lossless format, it is suitable for long-term storage, however, the industry standard and compressed MP3 format will also be a part of the dataset, in case the comparatively large FLAC files and format support become problematic in future. This measure maximizes the sustainability of the data.
2. In case of using virtual method for data collection (i.e Microsoft theme), the snowball sampling method will still apply to obtain additional participants. The initial participants will be requested to provide contact number or email address of further referred contacts, and/or this information will be searched for in the related organization's websites and in LinkedIn. In this case, the introductory email (attached with the application) will be used to request for participation. If the response is positive and consent obtained, the PIS and Consent form will be sent through email; the consent form will be requested to be digitally signed and returned. The Virtual interview will be conducted in the Encrypted Laptop; and thus, audio recordings will directly be saved in the RDS storage (if the participant consent to recording interview audio).
3. It is preferable that the observations are recorded right after the face-to-face interview takes place. This will create a flow from the discussions during the interview to the observations, as well as ensure availability of the participant. The participants will also be encouraged to communicate the researcher about any additional information they think of after the interview which might be useful in the study. Such communication can be both face to face, or virtual (through Skype or email).

The observation notes will be recorded in a DOCX file named COMPANY_X_Participant_YY_ON. Data inside this document will also be pseudonymized and uploaded in the relevant Company_X folder.

The audio recordings in MP3 format will be used to transcribe the interviews. An earphone will be used for transcription purpose in a secure environment (university workstation) so that they cannot be heard by any person out of the research team. The transcription files will be in DOCX format, renamed as Company_X_Participant_YY_TR and stored in the respective company's folder. As the transcription of all the audio recordings are completed, all the audio recordings in WAV, MP3 and FLAC format will be erased. [Voice alteration of the recorded audio step has been removed as all audio recording will be erased after transcription.]

As the data analysis are completed, any person/organization recognizable information in the collected documents from different companies will be hidden using the 'Black-out' feature of Adobe Acrobat software. Also, the Research_Data and Digitized_Research_Diary folders will be encrypted to AES 256 standard password protected ZIP file using the 7-Zip software (open source). ['Sensitive data' replaced by 'person/organization recognizable information'.]

1. Finally, a Metadata.TXT file will be added to the dataset, which will include:

- description of research context and design
- data origin and contextual information
- information on data collection instrument and methods utilized
- descriptions of data preparation and anonymization methods
- description of data, data files (format) and conventions of the file names
- description of various terms used (and full form of abbreviations)
- summary of findings from the data
- information on data availability

Participants will have the right to withdraw from the project if they want (informed through the Participant Information Sheet, Consent Form and verbally during interview). In this case, their interview transcript and recordings will be removed from the dataset, and this information will be added in the Metadata file. (To clarify, the data will only be pseudonymized, thus the identity of the participant will be replaced by a participant specific ID. This information - which ID represents which participant - will be separately preserved in a password-protected TXT file accessible by members of the research team only. Through this information, data related to a participant requesting withdrawal, will be discarded from collected dataset.)

Documentation and Metadata

What documentation and metadata will accompany the data?

To clarify the meaning of the data and ensure the transparency of data collection method - the metadata documentation will be extensive and will include the following:

- description of research context and design
- data origin and contextual information
- information on data collection instrument and methods utilized
- descriptions of data preparation and anonymization methods
- description of data, data files (format) and conventions of the file names
- description of various terms used (and full form of abbreviations)
- summary of findings from the data
- information on data availability

A separate file named Metadata.txt will accompany the data with this information.

Ethics and Legal Compliance

How will you manage any ethical issues?

- All issues relating to data collection and storage will be complied with ethically
- Data may be referred to departmental/Institution ethics committees
- Data storage to be in accordance with established regulations and norms
- All issues not conforming with ethics to be critically looked into
- All ethical concerns which includes and not limited coding of data will be handled according to the departmental or institutions ethics committees.

The main ethical issue regarding the collected data is the 'person-identifying information' aspect which is related to the confidentiality of the participant. The only person identifying data might be the principal investigator in this research who will be conducting interviews with the participants, which will be audio recorded and transcribed. To clarify, for in-depth understanding of

the viewpoint of the participants regarding various activities related to cost variability caused by risk in the construction industry, they will be requested to explain their background (nativity, education, professional career etc.). Response to such queries may contain information related to the participant's current and previous professional positions (name of Employer Company, the participant's role and responsibilities etc.), gender etc. Such information can essentially become person identifying, especially in case of rarity (e.g. female workers in construction companies).

To manage the ethical issues related to the 'person-identifying information' aspect, the following steps have been taken.

- Participant's name will not be recorded in the audio recording or the interview protocol. Rather for pseudonymization purpose the identity will be replaced by a participant specific code.
- The interview transcripts will also be pseudonymized. The pseudonymization data and contact details of the participants will be recorded in a password protected files with the access of the researcher and data custodian only.
- The data management plan ensures data security during data collection, processing, analyzing, storage and sharing.

In addition, the following steps have been taken to manage ethical issues in general.

- An application for ethical approval has been submitted for Full University Research Ethics Committee (UREC) Review (as Nigeria and other major African economies, selected country for data collection are outside the EEA zone, requiring full UREC approval). Data collection will only commence after the ethical approval has been obtained.
- The participants will be provided with a Participant Information Sheet (PIS) before the interview. The PIS explains to the participants what data will be collected, how their data will be used and shared and their freedom of withdrawing from the study at any point.
- Formal written consent from the participants will be gained using a consent form. Separate consents for collection of three separate forms of data (interview, providing guidance in observation and providing access to non-public documentation) is being obtained through separate points in the consent form.
- The consent form is also collecting permission from proper authority inside the participatory organization for allowing the participants to take part in interview in workplace during work hours, providing guidance in observation and providing access to documents with no sensitive/confidential information to the organization as judged by the participants themselves.
- The principal investigator has obtained trainings ethical issues like research integrity, data protection etc. as suggested by the University of Manchester.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Research data have a long-term value for *research* and academia, as *well* as taking into account other *research data* management physically present at the institution, *Intellectual property rights (IPR) which* are subject to the *University's Intellectual Property* Policy.

The copyright and intellectual property rights are the exclusive rights of the Principal Investigator of the research project, newly acquired data (interview audio files, transcripts and observation notes). CC BY 4.0 license will be applied to the data to enforce copyright (see the description of CC-BY 4.0 here: creativecommons.org/licenses/by/4.0). This will be enforced by adding the "©DOSUMU Babatunde" notice at the footer of every page of the interview transcripts. If such data are to be re used without going against the regulation, consent will be sought from appropriate authorities.

Storage and backup

How will the data be stored and backed up?

The research data will be stored securely in the University of Manchester's Research Data Storage (research-RDS), which automatically updates (i.e. uploads) any changes in the data. This is university provided data storage, registered and linked with the Principal Investigator's University of Manchester email address, shared from the PhD supervisor's storage space of 8TB. Intended to use only 1 TB maximum, this will be the primary storage, working and backup space for all the research data. The data custodian (the main PhD supervisor) will be provided access to the data collected from this space - only by providing them access through their own University of Manchester email addresses. The data will be deleted from this storage 5 years (following UoM's policy of data retention schedule) after publication.

Please refer to Data collection segment of this plan for elaborate data storage and management procedure.

How will you manage access and security?

To manage access and security of the data, the following steps will be undertaken:

1. Data will be uploaded to the University of Manchester provided secure UoM-RDS directly from the field without making a copy in any personal non-encrypted device. In field, this will be achieved by using a university encrypted personal laptop.
2. Paper data like consent forms, interview protocol other supporting documents collected from participating organizations will be kept locked in a secure cabinet provided by the university. Also, they will be converted into high-quality digital copies as soon

as possible after collection (without uploading to any cloud service or keeping a copy in any personal device).
3. The workstation provided for analysis and processing of data is password protected (associated with the researcher's university email address) and automatically backed-up (all files in the P Drive).
4. The interview transcripts will be pseudonymized. Also, the sensitive information in the documents collected (from the participants which are not publicly available) will be blacked-out.
5. An earphone will be used for listening to the audio recordings for the transcription purpose, so that none other than the researcher can hear.
6. The pseudonymization information will be kept in an encrypted password-protected .TXT file, with access of the primary investigator and data custodian only. This file will be separately stored in the University P Drive only.
7. Collaborators (only the PhD supervisors in this case) will be provided access to the data only through secure UoM-RDS, through their organizational email address.

Please see the 'Data Collection' section of this data management plan for further elaborate response to this query.

Selection and Preservation

Which data should be retained, shared, and/or preserved?

- Collected and analysed data to be retained according to the universities regulations on records retention schedules
- While the outcome that will lead to the development of frame work should be retained and preserved

All collected data is intended to be analyzed. However, for possible future comparison studies with other construction industry of other countries (ex. USA, China, India, Pakistan) and/or longitudinal studies - it might be beneficial to preserve the complete collected data (interview audio, transcript, observation notes and other collected documents). The data will be prepared in a way to be shared for the PhD project itself by the principal investigator and supplemented with in-depth metadata, thus no additional effort is necessary for data preparation. The size of the complete dataset is estimated to be maximum 80 GB even in the highest quality format, requiring no additional expenditure, thus financially viable.

The personal data of the participants will be kept for re-contacting purpose only with the participants for future studies, separately stored in a secure encrypted password-protected file, with access of the primary investigator and data custodian only and will be preserved for 5 years after publication (as per The University of Manchester's Information Governance Office Records Retention policy requirement).

What is the long-term preservation plan for the dataset?

- Anonymised data or questionnaires created for the research are to be preserved for a bit reasonable longer period.
- There may be the need to review and update such data sets
- Information that are to be used after the end of this project such as methodologies,questionnaires will be handled according to the records retention schedules of the institution.

The UK Data Service (ukdataservice.ac.uk/media/622368/costingtool.pdf) suggests that long-term preservation of dataset should be of low cost and sustainable. To conform to this policy and ensure high quality of collected data, the principal investigator will be gathering the consent, conducting the interviews, transcribing those and create the metadata to understand the dataset himself. In addition, the data storage planned for long term preservation of data (UoM-RDS) will be provided by the University of Manchester itself which ensures the longevity of storage as well as security of the dataset. The PhD supervisor is the data custodian and as a member of the research team he will remain at the University and have continued access to the data storage service. The data will be preserved for a period of 5 years after publication, the minimum preservation time required to conform to University of Manchester's Information Governance Office Records Retention Schedule. All of these factors are considered 'low or no additional cost' in the costing tool provided by the UK Data Service.

Please see the 'Data Collection' section of this data management plan for an outline of the plans for preparing and documenting data for sharing and archiving.

Data Sharing

How will you share the data?

The main-supervisor and co-supervisors of the PhD project will need to have access to the data for supervision and publication purpose. The data will be shared with them through UoM-RDS accessed through their University of Manchester email address.

Additionally, the University of Manchester also needs to view the data for audit purposes, the method of sharing will be assured upon

request for audit.
Except for this purpose, there is no intention to share the data with anyone else at this point in time, as the data might be further useful for the research team to conduct a longitudinal study in future on relevant topic and/or industry.

Are any restrictions on data sharing required?

As the data is planned to be pseudonymized, participant consent gained for data sharing (without personal information) and only shared among the research team (the principal investigator and PhD supervisors) – no additional restriction is necessary for data sharing.