

HISTORY OF CHANGES		
Version	Publication date	Changes
1.0	26/09/2024	Initial version

Project¹ Number: 101104921

Project Acronym: LAGOSTECH

Project title: Africa's Silicon Landscapes: negotiating Information and Communication Technology in Lagos

Call: HORIZON-MSCA-2022-PF-01
(MSCA Postdoctoral Fellowships 2022)

Topic: HORIZON-MSCA-2022-PF-01-01

Type of Action: HORIZON-TMA-MSCA-PF-GF

Name of fellow: Davide Casciano

Department: Department of Social and Cultural Anthropology

Name of Supervisor: Katrien Pype

DATA MANAGEMENT PLAN

¹ The term 'project' used in this template equates to an 'action' in certain other Horizon 2020 documentation

1. Data Summary

Purpose of data collection/generation and relation to project objectives:

It is the purpose of the data collection to examine the experiences of technology and digital creativity in Lagos, focusing on the social negotiation dynamics that determine whether Information and Communication Technologies (ICTs) are accepted, modified, or rejected. This is directly related to the project's objectives of:

- Identifying how tech experts, entrepreneurs, users, and consumers contribute to the development of local digital products in Lagos.
- Analysing the extent to which the emergence, success, or failure of silicon places is influenced by sociocultural factors in addition to technical factors.
- Examining the ways in which these African silicon places produce future preconfigurations, social spaces, and power relationships.

Types and formats of data:

- Qualitative data: Transcripts from open interviews, life stories, and focus groups (text files)
- Field notes and diaries from participant observation (text files)
- Audio recordings of interviews (audio files, temporary, to be deleted after transcriptions for better pseudo-anonymisation as agreed with the Social and Societal Ethics Committee (SMEC) KU Leuven, approved file G-2024-7977-R2(AMD))
- Video recordings (video files, for those who consent to be featured on the LAGOSTECH website)
- Survey data (possibly in spreadsheet format)
- Digital ethnography data from social media platforms (text, images)
- Quantitative data: Public statistics on internet usage and device availability (likely spreadsheets or databases)
- Archival and documentary sources (text files, possibly images)

Re-use of existing data:

The project will re-use existing public statistics (e.g., those collected by Yaba College of Technology) and documentary sources (e.g., from the Nigerian Communications Commission or University of Lagos archives).

Origin of the data:

- Primary data will be collected directly from Lagos participants through interviews, observations, and surveys.
- Secondary data will come from public sources, archives, and existing statistics.

Expected data size:

While an exact size is not specified, given the nature of the project (around 140 respondents, 9-12 months of fieldwork data collection), we can estimate:

- Text data (transcripts, field notes): Several GB

- Audio recordings (temporary): 10-20 GB
- Video recordings (if applicable): 10-60 GB
- Survey data and statistics: A few GB

Total estimate: 30-100 GB

Data utility:

This data could be useful for:

- Researchers in anthropology, sociology, African studies, and Science and Technology Studies (STS)
- Policymakers interested in ICT development in sub-Saharan Africa
- Entrepreneurs and investors interested in learning more about the Lagos technology ecosystem
- Organizations (such as the EU) interested in digitalization processes in Africa
- Stakeholders in Lagos and Nigeria engaged in technology and development initiatives

2. FAIR data

2. 1. Making data findable, including provisions for metadata

2. Discoverability and identifiers:

- Data produced in the project, while not restricted, will be made discoverable through metadata.
- For data uploaded to KU Leuven RDR or Zenodo repository, Digital Object Identifiers (DOIs) will be assigned, providing a persistent and unique identifier.
- For data stored on KU Leuven network drives, a local identification system will be implemented to ensure data is identifiable and locatable within the project's storage structure.

3. Naming conventions:

- I will implement a consistent naming convention for all files:
LAGOSTECH_[DataType][YYYYMMDD][ParticipantID]_[VersionNumber]
Example: LAGOSTECH_Interview_20240315_P001_v1.0

4. Search keywords:

- Keywords will be related to the project themes, for example:
"Lagos", "ICT", "Silicon landscape", "Tech entrepreneurship", "Digital innovation", "African technology", "Tech hub", "Yaba", "Itana", "Computer Village"
- Include specific technical terms relevant to each dataset

5. Version numbers:

- I will use clear version numbers for all documents and datasets:
v1.0 for the first complete version
v1.1, v1.2, etc. for minor updates
v2.0 for major revisions

6. Metadata creation:

There may not be a specific metadata standard for this interdisciplinary field, but we can create rich metadata including:

- Project identifier: LAGOSTECH
- Creator: Davide Casciano
- Title of dataset
- Date of creation
- Date of last modification
- Description of the dataset
- Methodology used for data collection
- Geographic location: Lagos, Nigeria
- Time period of data collection
- Data type (e.g., interview transcript, field notes, survey data)
- File format
- Software used (if applicable)
- Related publications or outputs
- Funding information: Marie-Curie Postdoc Global Fellowship, Project number 101104921
- Access rights information
- Terms of use / License

However, for datasets uploaded to repositories like KU Leuven RDR or Zenodo, we will use their built-in metadata schemas, which are typically based on established standards like Dublin Core. For KU Leuven RDR, see: <https://www.kuleuven.be/rdm/en/rdr/full-metadata-model>.

2.2. Making data openly accessible

1. The following data will be made publicly available:

- The study's aggregated and pseudo-anonymized results will be made publicly available.
- Public statistics and documentary sources that are already publicly available.
- Any publications resulting from the project will be made open access.

2. Data with restricted access:

- Raw interview transcripts, field notes, and any data containing personal or sensitive information will not be made openly available due to privacy concerns and the sensitive nature of the data.
- This restriction is in line with the FAIR principle of "as open as possible, as closed as necessary" and complies with GDPR requirements.

3. Data accessibility:

- Open data will be deposited in the KU Leuven RDR (Research Data Repository) or Zenodo.
- Restricted data will be stored securely on KU Leuven network drives and on a personal encrypted (VeraCrypt, AES 256bit) USB removable device.

4. Methods or software tools for data access:

- Open data will be accessible through standard web browsers and common software (e.g., PDF readers, spreadsheet applications).
- Any specific software needed to access certain data types will be documented in the metadata.

5. Documentation:

- Comprehensive documentation about data collection methods, data processing, and any software used will be included with the datasets.
- 6. Software inclusions:**
 - I will refer to or include open-source software used in data analysis whenever possible.
- 7. Data repository:**
 - KU Leuven RDR and Zenodo are the preferred repositories for this project, as they are certified and support open access where appropriate.
- 8. Repository arrangements:**
 - As per the communication with KU Leuven RDR, arrangements have been explored and confirmed for both open and restricted data storage.
- 9. Access to restricted data:**
 - KU Leuven RDR has an established procedure for access requests to restricted datasets.
 - Most access requests will be handled by the RDR team and the legal department.
 - For approval of data sharing with requestors, the principal investigator (Davide Casciano) will be consulted.
- 10. Data Access Committee:**
 - A formal data access committee is not needed, as the RDR team and legal department will manage access requests.
- 11. Conditions for access:**
 - The metadata and associated documentation will specify the conditions for access.
 - Machine-readable licenses (e.g., Creative Commons licenses) will be applied where appropriate.
- 12. Identity verification:**
 - For restricted data, the identity of persons requesting access will be verified through the established procedures of KU Leuven RDR and the legal department.
 - It may include checking institutional affiliations and signing agreements for the use of data.

2.3. Making data interoperable

- 1. Data interoperability:**

The project will strive to make data as interoperable as possible, allowing for data exchange and re-use between researchers, institutions, and countries. This will be achieved by:

 - Using standard file formats for data storage:
 - Text files: UTF-8 encoded plain text (.txt) or widely compatible document formats (.docx, .pdf)
 - Spreadsheets: CSV or .xlsx formats
 - Audio (temporary): WAV, MP3, or FLAC
 - Video: MP4 with H.264 encoding
 - Images: JPG or PNG formats
 - Ensuring all data is accompanied by comprehensive metadata

2. Standard vocabularies for interdisciplinary interoperability:

The project spans multiple disciplines, including anthropology, sociology, African studies, and Science and Technology Studies (STS). To ensure interdisciplinary interoperability:

- Where possible, use controlled vocabularies from established sources, such as:
 - Thesaurus of social and human sciences terms maintained by the UNESCO
 - Place names can be found in the Getty Thesaurus of Geographic Names
- For tech-related terms, align with standard glossaries like those provided by IEEE or ACM

3. Project-specific ontologies or terms:

Given the unique focus on Lagos' tech ecosystem, some project-specific terms may be unavoidable. In such cases, we will create a clear glossary of project-specific terms and concepts.

2.4. Increase data re-use (through clarifying licences)

1. Data licensing for re-use:

- Open data (aggregated results, public statistics, publications) will be licensed under Creative Commons Attribution 4.0 International (CC BY 4.0), allowing for the widest re-use possible while ensuring proper attribution.
- For video content on the LAGOSTECH website, as mentioned in the SMEC Ethics Assessment, a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International license will be used.

2. Data availability timing:

- Upon completion of data analysis, quality checks, and publications, open data will be made available as soon as possible.
- An embargo period of up to 12 months after the end of the project may be applied to some data in order to allow for the publication of research findings. The embargo period on data still to be used for scientific publications is justified so that the researcher can fully analyse and publish their findings while still making the data available within a reasonable period of time.
- Some scientific publishers may impose a 12-month restricted (embargo) access period on publications.

3. Data usability by third parties:

- During and after the project, open data will be available to third parties.
- Due to privacy concerns and ethical considerations, restricted data (raw interviews, field notes containing personal information) will not be made publicly available. It is, however, possible for researchers to request access to these data through a formal procedure involving the RDR team and the legal department at KU Leuven.

4. Duration of data re-usability:

- In accordance with the SMEC Ethics Assessment, personal data will be retained for 10 years after the end of the project, as specified by KU Leuven's RDM guidelines.
- In the KU Leuven RDR or Zenodo repository, open data will remain re-usable indefinitely.

5. Data quality assurance processes:

- All data will be backed up regularly and version controlled.
- Use of data validation techniques for survey data.
- Clear documentation of data collection and processing methods.

3. Allocation of resources

1. Costs for making data FAIR:

The costs associated with making data FAIR in the LAGOSTECH project include:

- Time spent on data organization, documentation, and metadata creation
- Storage costs for secure data management during the project
- Costs associated with depositing data in the KU Leuven RDR or Zenodo repository
- Potential costs for open access publication of research findings

Estimated total cost: from €0 to €7,000 over the course of the project

2. Coverage of costs:

- These costs are eligible as part of the HORIZON-MSCA-2022-PF-01-01 grant, as they comply with the Grant Agreement conditions.
- The project budget includes allocation for open access publication fees.
- KU Leuven provides institutional support for data management, including access to the RDR, which may offset some costs.

3. Responsibility for data management:

- The principal investigator, Dr. Davide Casciano, will be responsible for data management throughout the project.
- Dr. Casciano will oversee the implementation of the Data Management Plan, ensuring compliance with FAIR principles and ethical guidelines.
- The project supervisor, Prof. Katrien Pype, will provide guidance and oversight on data management practices.
- Support will be available from KU Leuven's Research Data Management team for technical aspects of data storage and sharing.

4. Resources for long-term preservation:

- KU Leuven's institutional infrastructure covers long-term preservation, specifically the Research Data Repository (RDR).
- The data's potential value lies in its unique insights into Lagos' tech ecosystem and its contribution to understanding ICT development in sub-Saharan Africa. This value justifies long-term preservation resources.
- Decisions about what data will be kept and for how long will be made collaboratively by:
 - Dr. Davide Casciano (Principal Investigator)
 - Prof. Katrien Pype (Project Supervisor)
 - KU Leuven RDR team (for technical feasibility and in compliance with institutional policies)
- As per the SMEC Ethics Assessment, personal data will be retained for 10 years after the end of the project, following KU Leuven's RDM guidelines.
- Open data and research outputs will be preserved indefinitely in the KU Leuven RDR or Zenodo repository.

4. Data security

1. Provisions for data security:

a. Secure storage:

- As mentioned in the SMEC Ethics Assessment, all digital data will be encrypted using AES-128 encryption with a 256-bit key.
- Open-source software such as GPG or VeraCrypt will be used for encryption.

Data will be stored on:

- Removable devices (USB Flash Drive shock resistant and IP69 certified)
- KU Leuven secure network drives
- The researcher's personal laptop system partition is encrypted using the same algorithm.

Non-digital data will be digitized and physically destroyed, except for consent forms.

b. Secure transfer:

- Data transfer will occur only through secure, encrypted channels.
- When transferring data from the field (Nigeria) to KU Leuven, encrypted removable devices will be used.
- For any necessary online transfers, secure file transfer protocols (SFTP) or encrypted cloud services approved by KU Leuven will be employed.

c. Data recovery:

- Regular backups will be made to the KU Leuven network drives, which are automatically backed up by the university's IT services.
- Multiple copies of encrypted data will be maintained on separate secure devices to prevent data loss.

d. Access control:

- Access to the data will be limited to authorized KU Leuven researchers involved in the project.
- Strong, unique passwords will be used for all devices and accounts containing research data.
- Two-factor authentication will be implemented where possible.

2. Long-term preservation and curation:

a. Data will be stored in certified repositories for long-term preservation and curation:

- KU Leuven Research Data Repository (RDR): This is the primary repository for the project's data. It is a certified and secure institutional repository that ensures long-term preservation and proper curation of research data.
- Zenodo: As an alternative, Zenodo may be used for some open datasets. Zenodo is a general-purpose open-access repository developed under the European OpenAIRE program and operated by CERN, ensuring high standards of data security and long-term preservation.

b. These repositories provide:

- Regular security audits and updates
- Redundant storage to prevent data loss
- Version control and data integrity checks
- Compliance with international digital preservation standards

c. For restricted data:

- Access will be managed through the established procedures of KU Leuven RDR, involving the RDR team and the legal department.
- This ensures that sensitive data remains secure while still being accessible for verified research purposes.

d. Data retention:

- As stated in the SMEC Ethics Assessment, personal data will be retained for 10 years after the completion of the project, in line with KU Leuven's RDM guidelines.
- After this period, personal data will be securely deleted according to institutional protocols.

5. Ethical aspects

1. Ethical and legal issues impacting data sharing:

a. Personal and sensitive data:

- The project involves collecting personal data, including potentially sensitive information about political views and religious beliefs.
- This data is subject to GDPR regulations and the Nigerian Data Protection Regulation (2019).
- Impact: Raw data containing personal information cannot be openly shared and must be protected.

b. Pseudonymization:

- All personal data will be pseudonymized to protect participants' identities.
- Impact: While this allows data analysis, it limits raw personal data sharing.

c. Informed consent:

- Participants will provide informed consent for data collection and use.
- Some participants may explicitly request identification in research outputs.
- Impact: Data sharing must respect the terms of the informed consent provided by each participant.

d. Cultural sensitivity:

- The research deals with local technology ecosystems in Lagos, which may involve culturally sensitive information.
- Impact: Care must be taken in sharing data to avoid misrepresentation or potential harm to local communities.

e. Intellectual property:

- Information about technological innovations and business strategies may be collected.
- Impact: Sharing such data must consider potential intellectual property rights and competitive sensitivities.

2. References to ethics deliverables and DoA:

- The project has been approved by the Social and Societal Ethics Committee (SMEC) of KU Leuven (file G-2024-7977-R2(AMD)).
- The research complies with the Charter of Fundamental Rights of the European Union, particularly Article 8 concerning "Protection of Personal Data".
- It also adheres to the European Convention on Human Rights and its Supplementary Protocols, specifically Section I, "Rights and Freedoms".

3. Informed consent for data sharing and long-term preservation:

Informed consent for data sharing and long-term preservation is included in personal data consent forms. Specifically:

- The consent form and the information sheet explain that participants' personal information will be securely stored in digital form for a limited time (10 years after project completion) and for scientific purposes only.
- It specifies that data used for publications will be uploaded to KU Leuven RDR or Zenodo repository, and access to restricted data will be provided for verification purposes.
- Participants are informed that their name and surname will not be included in research materials, and that instead a pseudonym will be used, unless they specifically request otherwise.
- For video interviews to be featured on the LAGOSTECH website, consent is obtained, including agreement to a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International license.
- The consent form informs participants of their right to withdraw consent and request deletion of their data.

4. Additional ethical considerations:

- A Distress Protocol is in place to manage any potential emotional distress during interviews.
- Contact information for local support services in Lagos is provided for participants.
- Through workshops and talks to share the results with local communities, the project will benefit Nigerian society and research participants directly or indirectly.

6. Other issues

1. KU Leuven (Host Institution) procedures:

- The project adheres to KU Leuven's Research Data Management (RDM) guidelines.
- It utilizes KU Leuven's Research Data Repository (RDR) for data storage and sharing.
- The project follows KU Leuven's policy of retaining research data for 10 years after project completion.

2. Social and Societal Ethics Committee (SMEC) KU Leuven:

- The fieldwork data collection has been approved by this committee (file G-2024-7977-R2(AMD)) and follows its guidelines for ethical research conduct.

3. Belgian national procedures:

- The project complies with Belgian laws on data protection and privacy, which implement GDPR.

4. Nigerian procedures:

- The research is conducted in Nigeria and follows the ethical procedures employed by the Lagos Studies Association (LSA) at the University of Lagos, of which the researcher is an active member.

5. Interdisciplinary considerations:

- As the project spans multiple disciplines (Anthropology, Sociology, African Studies, STS), it considers best practices for fieldwork from each of these fields.