Designing Trustworthy Data Institutions

Scanning the Local Data Ecosystem in Climate-Induced Migration in Lake Chad Basin - Pilot Study in Cameroon

October 2023

GPAI | THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

This report was developed by Experts and Specialists involved in the Global Partnership on Artificial Intelligence's project on Data Institutions. The report reflects the personal opinions of the GPAI Experts and External Experts involved and does not necessarily reflect the views of the Experts' organisations, GPAI, or GPAI Members. GPAI is a separate entity from the OECD and accordingly, the opinions expressed and arguments employed therein do not reflect the views of the OECD or its Members.

Acknowledgements

This report was developed in the context of the Data Institutions Project, with the steering of the Project Co-Leads and the guidance of the Project Advisory Group, supported by the GPAI Data Governance Working Group. The GPAI Data Governance Working Group agreed to declassify this report and make it publicly available.

Co-Leads:

- **Teki Akuetteh Falconer***, Africa Digital Rights Hub

The report was written by:

- **Thomas Hervé Mboa Nkoudou**, CEIMIA
- **Lama Saouma**

GPAI recognizes the meaningful contribution of experts who supported the project in different phases: **Alison Gillwald**†**,** Research ICT Africa; **Lee Tiedrich***, Duke University, **Saiph Savage***, Northeastern University.

GPAI also recognizes the dedication of contributors of Ecole Supérieure des Sciences et Techniques de l'Information et de la Communication (ESSTIC) and local organizations who played a pivotal role through the whole project: **Nga Minkala Alice**‡, ESSTIC; **Ulrich Talla Wamba**‡, ESSTIC; **David Alex Eto Mengom**‡**,** ESSTIC; **Samuel Ngombi Oum**‡, ESSTIC; **Christabella Ateh**‡, ESSTIC; **Fomekong Félicien**‡, National Institute of Statistics (Cameroon); **Tirlamo Norbert Wirnkar**‡, Minrex (Cameroon); **Bekono Nkoudou Hugues,** Minrex (Cameroon) **Ngie Pierre Joel,** Minrex (Cameroon); **Aristide Donald BILOUNGA**‡, IPREVA (Cameroon); **Jafsia Élisée**‡, SDI, (Cameroon); **Kengne Tagne Yannick Ghislain**‡, AIDER (Cameroon); **MAHAMAT Allamine**‡, CADEPI (Cameroon).

GPAI would like to acknowledge the tireless efforts of colleagues at the International Centre of Expertise in Montréal on Artificial Intelligence (CEIMIA) and GPAI's Data Governance Working Group. We are grateful, in particular, for the support **Stephanie King**, **Gwenaelle Le Peuch,** and **Noemie Gervais** from CEIMIA, and for the dedication of the Working Group Co-Chairs **Jeni Tennison*** and **Maja Bogataj Jančič*.**

* Expert
** Observer
† Invited Specialist
‡ Contracted Parties by the CofEs to contribute to projects

Citation

GPAI 2023. Designing Trustworthy Data Institutions: Scanning the Local Data Ecosystem in Climate-Induced Migration in Lake Chad Basin - Pilot Study in Cameroon, Report, October 2023, Global Partnership on AI.

Table of Contents

Preface **2**
Introduction **6**
Our focus in Lake Chad basin - Pilot study in Cameroon **7**
Preparatory phase 7
Fieldwork in Cameroon 10
The data ecosystem **11**
Stakeholders 11
Dynamics in data collection 13
Gaps and challenges **17**
Barriers to local participation 18
Lack of data on climate change-induced migration 20
Sustainability of data collection activities 21
Opportunities **22**
Solutions proposed by local organizations 22
Implementation strategies 23
Conclusion and next steps **29**
Bibliography **30**
ANNEX **31**
Interview guide 31
Consent form 34

Preface

Launched in June 2020, The Global Partnership on AI (GPAI) is a multi-stakeholder initiative bringing together leading experts from science, industry, civil society, international organizations, and governments that share values to bridge the gap between theory and practice on AI by supporting cutting-edge research and applied activities on AI-related priorities. GPAI aims to provide a mechanism for sharing multidisciplinary research and identifying key issues among AI practitioners, to facilitate international collaboration, reduce duplication, act as a global reference point for specific AI issues, and ultimately promote trust in, and the adoption of, trustworthy AI. GPAI has a Council and a Steering Committee, supported by a Secretariat hosted at the OECD, and two Centres of Support: one in Montreal (CEIMIA, the International Centre of Expertise in Montreal for Artificial Intelligence) and one in Paris (at INRIA, the French National Institute for Research in Digital Science and Technology).

As one of the two Centers of Support of GPAI, the CEIMIA supports the Responsible AI and Data Governance Expert Working Groups. The mission of the Responsible use of AI working group is to “foster and contribute to the responsible development, use and governance of human-centred AI systems, in congruence with the UN Sustainable Development Goals, ensuring diversity and inclusivity to promote a resilient society, in particular, in the interest of vulnerable and marginalised groups.” While the Data governance working group aims to “collate evidence, shape research, undertake applied AI projects and provide expertise on data governance, to promote data for AI being collected, used, shared, archived and deleted in ways that are consistent with human rights, inclusion, diversity, innovation, economic growth, and societal benefit, while seeking to address the UN Sustainable Development Goals.” In consideration of the missions of these working groups, GPAI developed a roadmap for the responsible use of AI for the environment, and commissioned a significant body of research on Data Trusts and Data Justice with input from partner organizations from across Africa, Asia, Latin America, and Oceania.

A relevant body of work on data trusts has emerged over the past few years. The Mozilla Foundation Data Futures Lab published a report (Geuns & Brandusescu, 2020) examining several approaches to data stewardship and alternative data governance, pointing to respective real-world examples. The Ada Lovelace Institute explored legal mechanisms for data stewardship and advocated for participatory models (Ada Lovelace Institute, 2021). The Data Economy Lab by Aapti Institute developed a repository of data intermediaries or stewards initiatives (Soni, 2020a). More recently, ODI attempted to measure the impact of data institutions on the ecosystems in which they operate(Open Data Institute, 2022). IMDA Singapore released the Trusted Data Sharing Framework (IMDA Singapore, 2019) as part of the Data Collaborative Programme, establishing the “common data sharing language”, a baseline, to help companies overcome challenges in addressing trust between data providers. The Data Economy Lab looked at theoretical bases and policy instruments for data sharing for Public Good (Soni, 2020b). In particular, GPAI commissioned the Open Data Institute and Aapti Institute to explore real-world use cases and operationalisation strategies where data trusts could offer societal benefit with a focus on AI and climate action. One of the use cases considered was climate migration, where important challenges, as well as opportunities to improve how data is collected, shared and used, were highlighted. CEIMIA has carried forward this recommendation to move towards practical interventions and to explore how data institutions - with a wider lens than data trusts - could make a difference in climate-induced migration. The goal is to support novel data institutions and data governance practices that promote the safe, fair and equitable sharing of data and empower individuals and communities to enact their data

rights, enabling an environment of trust for the development of responsible AI applications, in service of the UN Sustainable Development Goals. With this objective in mind, this report presents the first steps in Designing Trustworthy Data Institutions for climate-induced migration on the African continent, specifically in the Lake Chad Basin region.

Sophie Fallaha
Directrice générale | Executive Director

In 2020, GPAI commissioned the Open Data Institute and Aapti Institute to explore real-world use cases and operationalisation strategies where data trusts could offer societal benefit with a focus on AI and climate action. From this foundational work by the ODI and Aapti teams, emerged the first international definition of a data trust, a synthesis of the 'state of the art' in the design and implementation of data trusts, and a review of the legal and legislative frameworks that are currently in place to support the operationalisation of data trusts (GPAI, 2021). However, the 'state of the art', revealed that data trusts remain a novel, experimental approach within the wider family of data institutions. The Data Governance Working Group then recommended considering wider, bottom-up **data institutions and trustworthy practices where communities are empowered around their data.** For all these reasons, CEIMIA has carried forward this recommendation moving towards practical interventions and exploring how data institutions - with a wider lens than data trusts - could make a difference.

In the interests of continuity and respect of the GPAI's priorities, CEIMIA has followed up this recommendation by focusing on climate-induced migration. To do so, our use case was Lake Chad Basin (Nigeria, Niger, Chad and Cameroon); a region grappling with a complex humanitarian crisis with over 3.2 million people displaced due to floods, droughts and farmers-herders conflicts. After an initial literature review on climate-induced migration in the Lake Chad Basin, we selected Cameroon to conduct pilot research during which we interviewed some informants from organizations and migrant communities).

From this research, we succeeded to map the local data ecosystem by identifying the key stakeholders and describing the dynamics and interactions in data exchanges on climate-induced migrations. We have also identified two main factors responsible for gaps and challenges the data exchange process is facing for climate-induced migration:

- Endogenous factors (lack of data literacy; local data institutions lack funding, infrastructure and highly trained experts on data collection for AI; etc.);
- Exogenous factors (databases on climate change are owned by international organizations; funding comes from external sources; etc.)

In order to overcome these challenges, organizations stewarding data, better, data institutions are recommended to take some key actions listed below;

Table 1: Key actions

Key Actions	Description
Building data literacy and participation of local communities	- Enhancing data literacy for climate change and the co-production of data on climate-induced migration to support informed decisions as to how to adapt, whether and when to move, and where to settle. - Empowering women with climate information and skills to ensure that their voices and knowledge are considered during decision processes on climate adaptation and in climate mobility. - Developing community-based governance models to enable stakeholders with the ability to use the data; - Improving inclusiveness and ownership within data steward organizations in order to foster greater representation of marginalized groups.
Strengthening local data infrastructures and experts	- Involving local experts and stakeholders in international research efforts through data collection, model development and scenario-building; - Encouraging national conversation between local actors working on climate migration and international organizations. Building on and enhancing coordination across existing national structures; - Increasing dedicated data collection and stewardship of climate migration data and strengthening data exchange by building a network of

	data-sharing (preferably digital) hubs to make existing data publicly available; - Encouraging community-led solutions for climate mobility in Lake Chad Basin, and ensuring their sustainability, even in the absence of international organizations.
Updating data collection tools	- Allowing consecutive survey waves and consecutive qualitative data gatherings to monitor displacements and conflicts over time and make more accurate inferences about implications for the future; - Supporting local organizations involved in the data ecosystem (statistical offices, meteorological agencies, ministries, etc) in standardized data collection to facilitate timely exchanges; - Considering the large penetration of mobile devices as new opportunities for real-time data collection and sharing and allowing for the inclusion of perspectives from marginalized communities; - Using a Human-centric approach in data collection by taking into consideration native languages, storytelling, indigenous knowledge, comics, etc.

Taken into account by data institutions, these actions provide a path to improve trustworthiness in data governance. But this is only possible within an integrated data governance system that places the values of sharing and openness at the center of data exchange. This is why, as a next step, we propose to develop a framework for data institutions.

Introduction

Climate change is driving displacement, causing conflicts and making life harder for refugees in settlements. Several organizations are working to better understand and document key drivers of migration as well as the impact of Climate Change. This is part of the justification as to why, considering the potential of Artificial Intelligence to be harnessed responsibly to accelerate positive environmental action, GPAI countries have put climate action at the top of their agenda[1].

Indeed, data and artificial intelligence have a vital role to play in helping us understand and tackle this climate crisis, from predicting climate-induced migrations to extreme weather events. However, as with data systems at large, individuals and communities tend to have little say in how data is collected, used and shared for climate action. Data trusts and other forms of 'bottom-up' data stewardship have emerged to reverse this trend and empower people to take part in the data economy (GPAI, 2021). This is the sense in which the GPAI has committed the Open Data Institute and Aapti Institute to exploring real-world use cases and operationalisation strategies where data trusts would offer societal benefits with a focus on AI and climate action.

The ODI and Aapti Institute report showed optimism for the potential of data trusts, but concluded that, in some cases, they are not the right solution. Indeed, assessments of small shareholder farming in India and climate migration in Peru revealed difficulties for the broader implementation of data trusts (especially regarding the data rights necessary for data trusts being present in local jurisdictions), whilst also highlighting important challenges in the Global South, with clear openings to improve how data is collected, used and shared. Alternatively, data could be stewarded and shared responsibly, without the need for trustees with explicit fiduciary obligations. There is therefore an opportunity to explore other existing or new institutional approaches to data stewardship and data sharing that would be more fitting to those contexts.

Following the ODI and Aapti Institute report's conclusion, the GPAI recommended that wider, bottom-up data institutions and trustworthy practices where communities are empowered around their data without the need for trustees with fiduciary obligations be considered. Data institutions are defined here as organizations that steward data on behalf of others, often towards public, educational or charitable aims (Open Data Institute, 2021). According to that same ODI and Aapti Institute report, climate-induced migration seemed to be a good use case to apply this recommendation.

This project aims to explore how data institutions and AI applications could make a difference on climate migration and empowering local organizations & communities. More specifically, we would like to:

- better integrate Global South perspectives and bring in the most affected communities to contribute to building evidence and developing solutions by participating in problem identification, data collection and analysis.
- co-design a framework for trustworthy data exchanges within the climate migration data ecosystem, that improves how data is being collected, stewarded, shared, and used, to better serve the needs of communities and empower them to play an active role in the data value chain.

However, we all agree that we will not be able to achieve these goals if we do not fully understand the local data ecosystem. Hence the pertinence of this report, which presents the first steps of our study: a mapping of the data ecosystem (stakeholders and their interactions), an overview of the gaps and challenges that the data collection process is facing and, finally, identified opportunities. Before sharing with you these key findings, it is important to present the context of the Lake Chad Basin where we chose to conduct our research as well as the methodology that we used during the fieldwork in Cameroon.

[1] https://gpai.ai/projects/responsible-ai/environment/

Our focus in Lake Chad basin - Pilot study in Cameroon

The pilot phase of this research was conducted by Thomas Hervé Mboa Nkoudou, researcher in residence at CEIMIA and Lama Saouma, under the guidance of GPAI expert Teki Akuetteh. In the field, we also received scientific support from local researchers, following an agreement signed between CEIMIA and the Ecole Supérieure des Sciences et Techniques de l'Information et de la Communication (ESSTIC). Through this agreement, Professor Alice Nga Minkala, Director of ESSTIC, has made available to us a team of junior researchers from the Master's programme in Digital Humanities of ESSTIC, to cooperate with us in the various phases of our pilot research. Our methodology is summarized in the following picture:

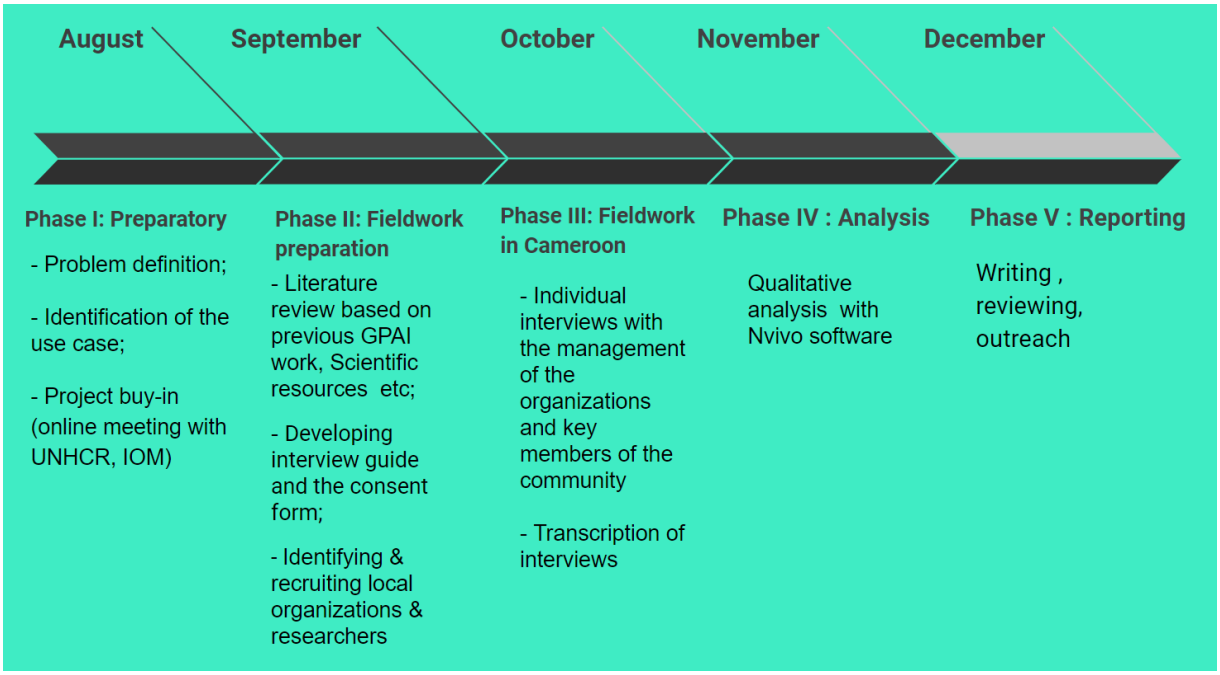

Figure 1: *Summary of the methodology applied during the fieldwork in Cameroon (Aug - Dec 2022)*

Preparatory phase

The objectives of this phase were to select a use case from previous GPAI work on climate-induced migration and identify relevant organizations and local communities we could engage with.

Selection of the use case: Lake Chad Basin

The ODI interim report entitled *"Enabling Data Sharing for Social Benefit Through Data Trusts: Data Trusts in Climate"* serves to gain a first understanding of the distribution of activities across the world, related to climate migration and to select specific organizations mentioned in the Peru use case. We used *Kumu.io* to build an interactive map of those organizations[2].

[2] Those in blue are handling data or are specialized on climate migration. You can access the interactive map here: https://embed.kumu.io/d1d25a9023b92133257d91a4ca821a91.

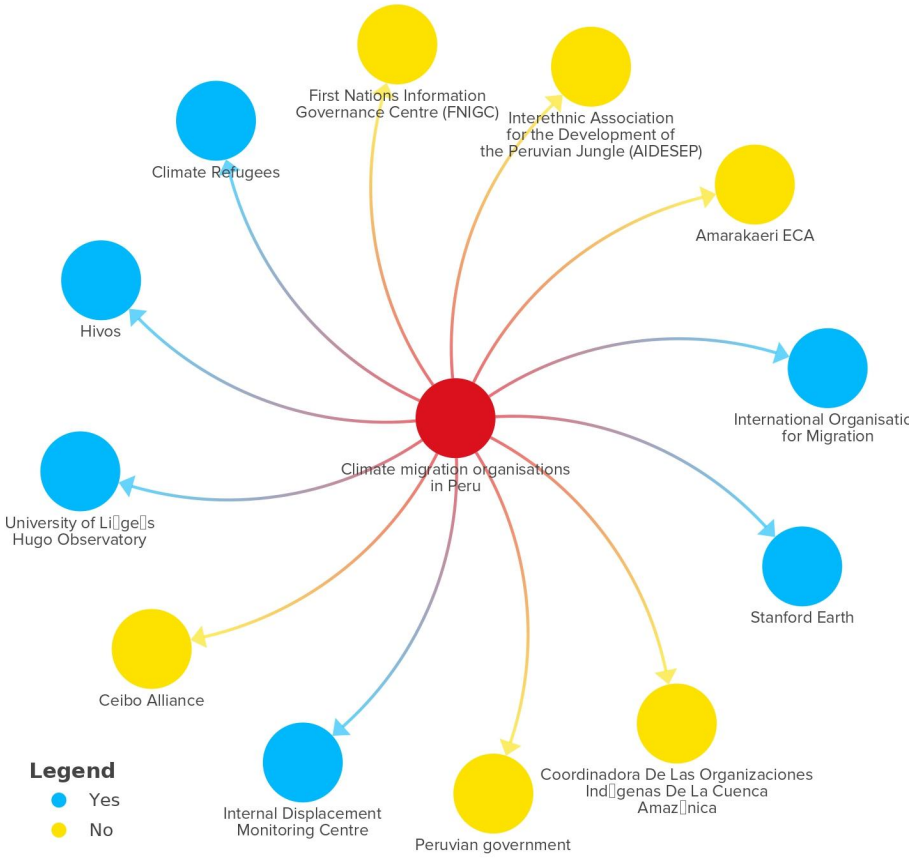

Figure 2: *Organizations mentioned in the Peru use case on climate migration*

Amongst organizations mentioned in the Peru use case, we decided to select those working on climate change; they are featured in blue in the figure. We then explored in depth their websites and the various reports they have published. This work was undertaken to identify organizations that have an interest in climate change and the geographic areas in which they work. During this quick analysis, we found that the Lake Chad Basin was mentioned in several reports on climate migration. This strong resurgence of this region motivated our choice for further study.

The Lake Chad Basin is located in the eastern part of the Sahel region at the southern edge of the Sahara Desert (IAEA Project, 2017). Lake Chad borders four countries — Chad, Nigeria, Niger and Cameroon. However, the Lake Chad "Basin" covers almost 8% of the continent and spreads over seven countries: Algeria, Cameroon, Central African Republic, Chad, Libya, Niger and Nigeria(Leon Usigbe, 2019).

Figure 3: *Localisation of Lake Chad Basin*

Unfortunately, the Lake Chad Basin is grappling with a complex humanitarian crisis with over 3.2 million people displaced due to floods, droughts and farmers-herders conflicts. But, several organizations including UN organizations are collecting, stewarding and using data to better understand the impact of Climate Change on migration patterns and migrants' condition to help develop mitigation and adaptation strategies. Therefore, the Lake Chad Basin was found to be the context par excellence for studying climate-induced migration in the Global South.

Identification of relevant organizations and local communities

The idea here was to make sure that the organizations selected in the previous figure (those in blue), are really stewarding data related to climate migration in Lake Chad Basin. For this purpose, we went through their websites and we found that the United Nations High Commissioner for Refugees (UNHCR), Internal Displacement Monitoring Centre (IDMC) and International Organization for Migration (IOM) are managing interesting and relevant data platforms.

Table 2: Relevant data platforms on climate change in Lake Chad Basin

Organizations	**Initiatives**	**Website**
UNHCR	Microdata Library	https://microdata.unhcr.org/index.php/home
IDMC	Global internal displacement database	https://www.internal-displacement.org/database/methodology
IOM	Migration Data portal	https://www.migrationdataportal.org/
	Displacement Tracking Matrix	https://dtm.iom.int/

This first work was highly appreciated by the Project Advisory Group of the GPAI Data Governance Working Group. However, an important comment was made: We already know that these UN organizations (IOM, UNHCR, IMDC) have in their possession a lot of data related to climate migration. However, we do not have any information on local organizations working in the field on this issue and their relationship with the UN organizations.

To address this comment, we conducted a literature review to try to identify local organizations working on the issue of climate migration in the Lake Chad Basin. For this purpose, we used Google Scholar to find relevant publications (scientific papers, books, reports, etc.) during the period 2000 - 2022; with keywords such as: *Lake Chad Basin, Climate refugees, Environmental migrants, Climate migrants, Environmentally displaced persons, Climate migration, Climate mobilities, Induced climate migration,* and *Climate displacement.* The selective bibliography obtained from this research is reported in Annex 1.

After the research on Google Scholar, we did a document analysis to find organizations working in the field in Lake Chad Basin and map them according to the ODI Data Ecosystem Mapping tools[3]. We identified three main categories of organizations:

1. **Data Stewards**: organizations responsible for collecting, managing or ensuring access to a database. It appeared that in the case of climatic migrations in the Lake Chad basin, this category consists primarily of the UN organizations.

2. **Data Intermediaries:** organizations which aggregate data in the ecosystem. These organizations are mainly International non-governmental organizations receiving funding from the Data Stewards, to operationalize programs and projects in Lake Chad Basin.

[3] The ODI Data Ecosystem Mapping Tools https://theodi2022.wpengine.com/wp-content/uploads/2019/06/ODI-Data-Ecosystem-Mapping-%E2%80%93-print-at-home-guide-A4-%E2%80%93-2019-06-26.pdf

Amongst them, we identified: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), Red Cross, Plan International , CARE International, Netherlands Development Organisation (SNV), Catholic Relief Services (CRS).

3. **Data Contributors or local organizations**: organizations who contribute to the dataset(s), either knowingly or unknowingly. These organizations are local non-governmental organizations and government institutions receiving funding directly from the Data Stewards or from Data Intermediaries to collect data in the country. The most relevant local organizations per country of the Lake Chad Basin are listed in Annex 2.

Given the large number of organizations and the exploratory nature of the research, we found it appropriate to go step by step with a pilot study in Cameroon.

Fieldwork in Cameroon

We conducted pilot research with Cameroon-based organizations to better understand their role in the climate-induced migration data ecosystem, their data governance practices and the challenges they face, the relationship and data exchanges between different players and the degree to which the most affected communities are included in the process.

- **Engaging with key organizations**

Before beginning our research fieldwork in Cameroon, it was crucial to make contact with the key organizations involved in the data collection process. To this end, we had a series of meetings with the IOM and UNHCR, which are the organizations collecting the most data on climate migration in the Lake Chad Basin region. The idea was to inform them about the work we wanted to do and to see how the results obtained could contribute to solving common problems.These meetings proved to be fruitful, as we were able to secure the buy-in of IOM, which asked us to provide a letter of interest to collaborate with them. However, it was difficult to secure collaboration due to the nature of the activities and programming structure of these organizations. Similarly, local organizations were difficult to get in touch with. Websites were often not functional and where they were, contact information was often not up to date. That is why, once in Cameroon, we started with informal meetings with our local contacts, to get more information on organizations whose actions we read about in our previous document analysis.

- **Interviews**

As a prelude to the research fieldwork to be conducted in Cameroon, we prepared an interview guide (see Annex 3) and a consent form (see Annex 4) to be read and understood before agreeing to participate in this research project. The consent form explains the purpose of this research project, its procedures, benefits, risks and drawbacks.

Participation consisted of taking part in a one-on-one interview with the responsible researcher, lasting between 30 minutes and 1 hour, covering the following topics: awareness about climate migration, data collection, data governance, and data literacy. In total, we conducted six interviews with six key informants from different local organizations. All interviews were recorded (audio only).

- **Processing of the interviews and qualitative analysis with Nvivo software[4]**

The processing of the recorded interviews began with the anonymization of all audio files, in order to ensure confidential and impartial processing. To do so, we used the following abbreviations: INF1 to INF6; with “INF” standing for “INFORMANT” followed by a serial number from 1 to 6, to designate the individual participants interviewed. Once the anonymization stage was completed, we proceeded to transcribe the audio files to text. The anonymized files from the transcription were

[4] NVivo is a qualitative data analysis software package produced by QSR International. It was designed for researchers working with very rich textual or multimedia information, where in-depth levels of analysis on small or large volumes of data are required. The most recent version of Nvivo was used in this study.
https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/about/nvivo/who-its-for/academia

then imported into the NVIVO software for thematic coding. The different results of this analysis are presented in the second part of this report.

The data ecosystem

Our interviews have been analyzed through the lens of the ODI's Data Ecosystem Mapping tools[5]; this has allowed us to illustrate the different stakeholders involved in the data ecosystem, as well as the dynamics that govern their interactions during the data collection process.

Stakeholders

By stakeholders, we mean organizations working in the Lake Chad Basin region with expertise on climate change and related topics like climate-induced migrations. They can be grouped into two categories: local organizations (non-governmental and governmental) and international organizations, who tend to act primarily as donors.

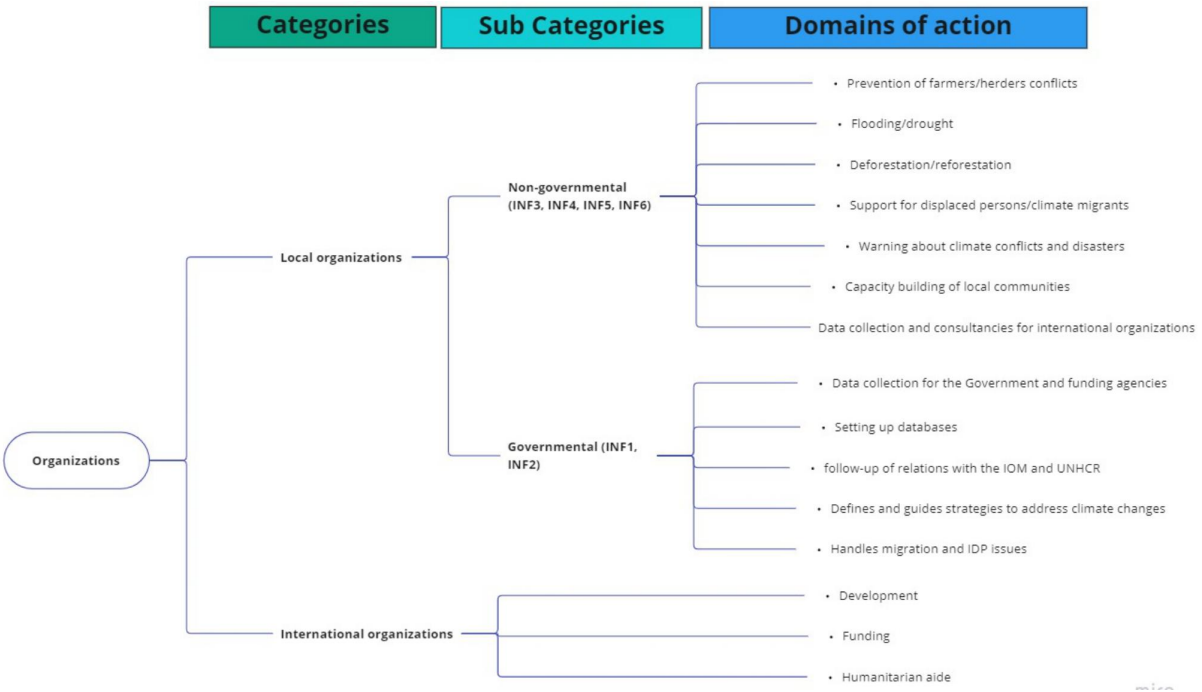

Figure 4: Overview of the actors of the data ecosystem in Cameroon, their different subdivisions, and actions they take.

Local non-governmental organizations

In the Cameroonian part of the Lake Chad Basin, the non-governmental organizations we met with are mostly associations, foundations, or groups of civil society organizations. Their areas of intervention are mainly related to: farmers/herders conflicts, floods/droughts, deforestation/reforestation, and prevention, support, and capacity building of affected local communities.

To illustrate how these local non-governmental organizations work with affected communities, INF6 explained that their work is done in districts where agriculture, fishing, and livestock are the main activities; specifically in the arrondissements of Zina, Logone Birni, and Waza in the Logone-et-Chari department. These districts are consistently exposed to threats such as drought

[5] The ODI Data Ecosystem Mapping Tools https://theodi2022.wpengine.com/wp-content/uploads/2019/06/ODI-Data-Ecosystem-Mapping-%E2%80%93-print-at-home-guide-A4-%E2%80%93-2019-06-26.pdf

and floods[6], thus giving rise to conflicts between farmers and herders, as INF5 explains: "In the region, sometimes there are internal problems because of the grazing areas. Others feel that an area where the neighbor is doing agriculture, is reserved for grazing for their animals, and lets them strut around. This can create tensions". This is the context in which local organizations intervene, usually starting with early warnings:

> I had to intervene in the plains of the Logone. I learned there that this part constitutes a grazing area par excellence; because it is the only base that serves as a pasture for the different herders of the area (and even those who come from Nigeria and Niger). But in the dry season, everyone comes there. Some come for the pasture, others for the cultivable land, etc. This is how conflicts emerge and it is this "Conflicts" dimension that concerns me the most. Indeed, the movements of these communities being limited in this zone, I saw herders settling all along the roads, suggesting future conflicts with farmers. This situation drew my attention and I had to alert the authorities (INF6).

This role of warning whistleblower in the sense of conflict prevention and climatic disasters such as floods and droughts, seems to be the primary mission of these local non-governmental organizations (INF3). However, it must be recognized that despite the many warnings, events such as floods often happen so quickly that the populations are subjected to them head-on. We met with people from communities affected by climate change who live this reality on a daily basis:

> I am a descendant of a family victim of climate migration. We migrated from the far north of the country to the north of Cameroon. When we left the Far North, it was because there was no rainfall. Today, we heard about floods. There is a kind of reversal of the situation (INF4).

Hence the focus of some local non-governmental organizations is accompanying displaced people and assisting in their integration into their new environment. Other organizations specialize in the implementation of resilience strategies through reforestation activities among others. For instance, the INF6 organization is promoting bamboo for land restoration in Pette, Mora and Waza districts.

Local governmental organizations

Within the Cameroonian government, we collaborated with key informants from the Ministry of External Relations (MINREX) and the National Institute of Statistics (INS).

The INS[7], among other things is responsible for: a) making available the statistical data and indicators needed for economic and social management and b) setting up databases and ensuring the conservation of census and survey files carried out by public administrations and bodies subsidized or controlled by the State. It has a department dedicated to migration and climate change data management at the national level. INF2 pointed out that the division dedicated to climate change has been recently created. According to INF2, INS does not work alone on climate change issues, but collaborates closely with other governmental organizations such as: the National Meteorological Office and the National Observatory on Climate Change (ONACC).

At the Ministry of External Relations (MINREX)[8], there is a Sub-Directorate for Relations with the International Organization for Migration (IOM) and the United Nations High Commissioner for

---

6 https://www.facebook.com/cosvilc/videos/inondations-dans-le-logone-et-chari-apres-zina-logone-birni-cest-waza/2436875213260412/

7 The INS was created by Presidential Decree N°2001/100 on 20 Avril 2001

8 For the organizational chart, see the Décret présidentiel N°2013/112 du 22 avril 2013 portant organisation du Ministère des Relations Extérieures.

Refugees (UNHCR): “We work much more with UN organizations, especially the UNHCR for forced migration; and IOM for other types of migration” (INF1). This Sub-Directorate includes the IOM/UNHCR Technical Secretariat and manages all aspects of migration in the broad sense (forced immigration, regular immigration, mixed immigration, etc.):

This sub-directorate is responsible for monitoring: (a) relations between Cameroon and the IOM; (b) relations between Cameroon and the UNHCR; (c) monitoring the activities of the IOM and the UNHCR; (d) the National Secretariat of the IOM; (e) the application of the commitments and obligations of the State in the framework of its cooperation with the IOM and the UNHCR; and (f) the periodic evaluation of cooperation with the IOM and the UNHCR.

However, INF1 pointed out that its organization is not working alone on this issue. They work closely with the Ministry of Territorial Administration, which is much more present on the ground and manages the consequences of these migrations (in the form of internal displacements) through the Sub-divisional Officer.

International organizations

International organizations that were regularly mentioned in the interviews were UN organizations, such as the IOM and UNHCR; international development organizations (GIZ), as well as funding agencies (World Bank). In the Cameroonian part of the Lake Chad Basin, GIZ is presented as the most active international development organization on climate change issues (INF3, INF4, INF5).

Dynamics in data collection

From our interviews with local organizations, three important aspects of the data collection process emerged: the choice and design of the data collection tools, the participatory approach in the field and data management.This section describes the dynamics within the data ecosystem.

Data collection tools

These tools are usually pre-designed by the sponsors of the surveys (international organizations, donors, etc.) and are organized around a matrix that provides information on the types of resources (natural, human, economic, etc.), the hazards that impact these resources (they may be climate-related), the adaptation and resilience strategies in the face of these hazards, etc. (INF3). In this ecosystem, we generally find traditional data collection tools (questionnaires, interviews, focus groups and storytelling) as well as geolocation and biometrics techniques.

Traditional methods

Questionnaire

Questionnaires are usually administered on paper, to collect socio-demographic data, as well as data on issues such as gender, vulnerability, hygiene, access to water and land, and environmental and climate issues (INF3). However, some donors or international organizations may provide tablets and telephones to collectors so that the sponsor automatically receives the data collected (INF6).

Focus groups

Another very common data collection tool is focus groups, also identified here as village assemblies (INF6). These are moments of group exchange, which give people the opportunity to freely discuss the issues that affect their daily lives, and to carry out a participatory diagnosis during which the various problems related to climate change could be identified. But INF5 warned us about how tricky it can be to conduct focus group discussions in this part of Cameroon; since in this context, “women cannot express themselves among men”. In order to avoid such socio-cultural influences, INF5

advised to set up different categories of people when running focus group discussions, such as young people, elderly people, men and women.

Interviews and storytelling

Although the notion of interviews was mentioned by some local organizations as a data collection technique, it must be noted that very few actually use them. Conversely, some organizations recognize the power of storytelling in these contexts, as INF3 explains: "There are common tools that allow the organization to have a lot of information coming from people who seem to be vulnerable, but it is in the life story that we will better understand other problems, which sometimes do not automatically emerge".

Integration of mapping and biometrics techniques

From the various interviews conducted with local organizations, mapping appears to have great potential to improve the living conditions of populations in areas affected by climate change. In this sense, INF5 explains that they have used mapping to list the different natural resources, infrastructures and to delimit grazing areas from agricultural areas. He sees great potential in mapping:

> Mapping allows us to know the different infrastructures present in the community. For example, let's take the case of water points, boreholes and others; when they are mentioned on the map, the populations can better appreciate the number of functional and non-functional water points (INF5).

In the same manner, INF2 argues that when they conduct household surveys, the data is geo-referenced; this allows them to subsequently perform other geospatial analysis.

With respect to biometrics technique, it is much more widely used to identify refugees. The UNHCR uses the Biometric Identity Management System (BIMS)[9] to identify people digitally and assign them a code that will allow monitoring their movements. Thus, "if a refugee is identified in Chad and comes back to Cameroon to ask for asylum, just with the iris, we can detect that you are already identified" (INF3).

The participatory approach: a prerequisite!

The main advantage of involving the local population in the data collection process is that they are more familiar with the realities they experience on a daily basis and have relevant information on the subjects covered (INF3, INF4, INF5, INF6). This approach is supported by INF4 when he says that if we do not work with local leaders, we will be working with a small number of people and we run the risk of using irrelevant data. That is why, before beginning the data collection process, it is important to go through a preparatory phase that will ensure the effective participation of local communities. This is a participatory process whose objective is to engage everyone (the government, the population, traditional leaders, etc.) in the data collection process so that everyone feels included. In order to do this, local non-governmental organizations have a whole protocol of participatory approaches (generally preconceived by the sponsors of the data collection) to get the affected communities on board. This protocol begins by meeting with traditional authorities, the selection of translators, local leaders and collectors, and the training of collectors.

Meetings with the traditional authorities

The process begins with a declaration to the sous-prefecture, in order to obtain administrative authorization. It is only after this authorization that local non-governmental organizations can meet with traditional authorities (chiefdoms) and local populations to explain the purpose of the data

[9] https://www.unhcr.org/protection/basic/550c304c9/biometric-identity-management-system.html

collection that will take place in their community (INF3, INF4, INF5, INF6). INF2 indicates that, depending on the field, a speech is prepared to arouse the interest of the population for the data collection that we wish to do with them.

> For all our interventions, we first ask for the agreement of the Sous-Prefect. As soon as the agreement is obtained, we proceed with negotiations with the traditional chiefs. They play a very important role. We must always go through them. For the commune, we can pay the mayor a courtesy visit, but the sous-prefet is the gateway. (INF5)

Recruitment of translators, local leaders and data collectors

It is important to be aware that French is not widely spoken by the local populations, while it is the language used in the data collection process. That is why INF4 points out that language is an important factor to take into account when meeting with local populations. Hence the almost obligatory presence of a translator whose role is described below:

> He would go to them (the local populations) and exchange with them in *Toupouri* and *Fulfulde*. He would tell them this: On such and such a day, experts from such an organization will arrive to discuss with you your needs, your expectations... They will exchange directly from your place. They will list a certain type of information in terms of infrastructure and equipment. (INF4)

Beyond the need for a translator, it is also important to have local leaders who serve as a bridge between local organizations and the communities. These leaders are usually designated by the local populations themselves; by doing so, they feel more reassured and open. The role of the local leaders is to work closely with the local non-governmental organizations, to accompany them directly to the homes and to raise awareness among the population (INF5).

Then comes the recruitment of data collectors in the affected communities:

> It was arranged that the collecting agents were people who were locally active. In this area, the populations do not always speak the official languages. You must therefore understand their languages well in order not to bias the data (INF5).

Training of data collectors

Regardless of the data collection tool, it is crucial to train those who will be going into the field. After being recruited from the affected communities, data collectors must be trained in the use of traditional data collection tools, digital tools and associated software, or GPS mapping data collection (INF6). Once equipped with these new skills, collectors are ready to go into the field to collect personal data (name, surname, age, gender, home, village, etc.), take notes, images, geospatial data and prepare reports for donors.

In addition to training in data collection tools, data collectors are also prepared to face the difficult working conditions that often await them in the field. In particular, walking under the sun, door-to-door, etc. :

> This work is done by walking and under the hot, dry sun of the far north. You must synchronize your program with that of the communities, since they will explain that they cannot leave their crops to participate in the surveys. Especially since the people who live there walk these distances. You must therefore enter into their daily lives to understand how these communities live and how they got there (INF4).

It is only after this process has been completed that good data collection can take place. However, it should be mentioned that a bonus should be prepared for the populations participating in the study to compensate them for their time: “Every time, we made sure that each inhabitant of the area got a minimum of food and 3000 XAF. Since they are often farmers and go to their fields, this helps to compensate a little bit; their loss of time” (INF4).

Data management, the triad: Processing - Evaluation - Presentation of results

Processing of the data

Once data are collected, they are anonymized to hide information that could reveal the respondents' personal information:

> When processing the data, the identity of the persons involved is removed. In this way, you will not be able to find your own data if you search in the databases. The database is therefore anonymous (INF2).

Depending on the terms of reference received by the local organization that was responsible for collecting the data, the raw data may be transferred to the donor databases or a report about the data is prepared by the local organization and transmitted to the donor). This is the point at which the effective involvement of local organizations ends, leaving room for evaluation and monitoring by international organizations. In fact, whatever form they take after processing (raw data or reports), the information collected will always be routed to a dedicated platform managed by the donors. From there, international organizations will verify the data received from the field.

Evaluation

For raw data, INF4 explains that: "This aspect is managed by the Project Lead. He was the one who regularly reviewed the quantities of data collected. But more than that, there was a check on the quality of the data collected in the field”.

For reports, INF5 says that "after writing the report, it first goes to the lady who is in charge of monitoring and evaluation. She examines our reports and checks if they were done according to the norms and after her approval, the reports can be sent to the hierarchy". However, INF4 would like to point out that this process is not always so linear. It could happen that: "every three weeks, the project lead returns to the field to conduct additional verifications with the population directly". It is at the end of this verification process that the final phase of presenting the results to the population takes place.

Presentation of results

In this regard, all local organizations interviewed are unanimous on the fact that it is always necessary to plan sessions for the restitution of survey results with those communities affected by climate change. Everything that is done with their data must be presented to them each time, in the presence of administrative and traditional authorities: “There are feedback meetings during which reports are presented in the presence of traditional authorities” (INF5).

At the end of the various feedback meetings, copies of the final survey report are always made available to the administrative and traditional authorities. This ensures that the content of the survey is not perceived as a secret by members of the local community who wish to consult its contents (INF4, INF5). Thus, reports are the preferred form of data feedback to the local populations and organizations from which they were collected. INF1, for example, emphasizes that its (governmental) organization works closely with the UNHCR and the IOM and receives regular reports from them.

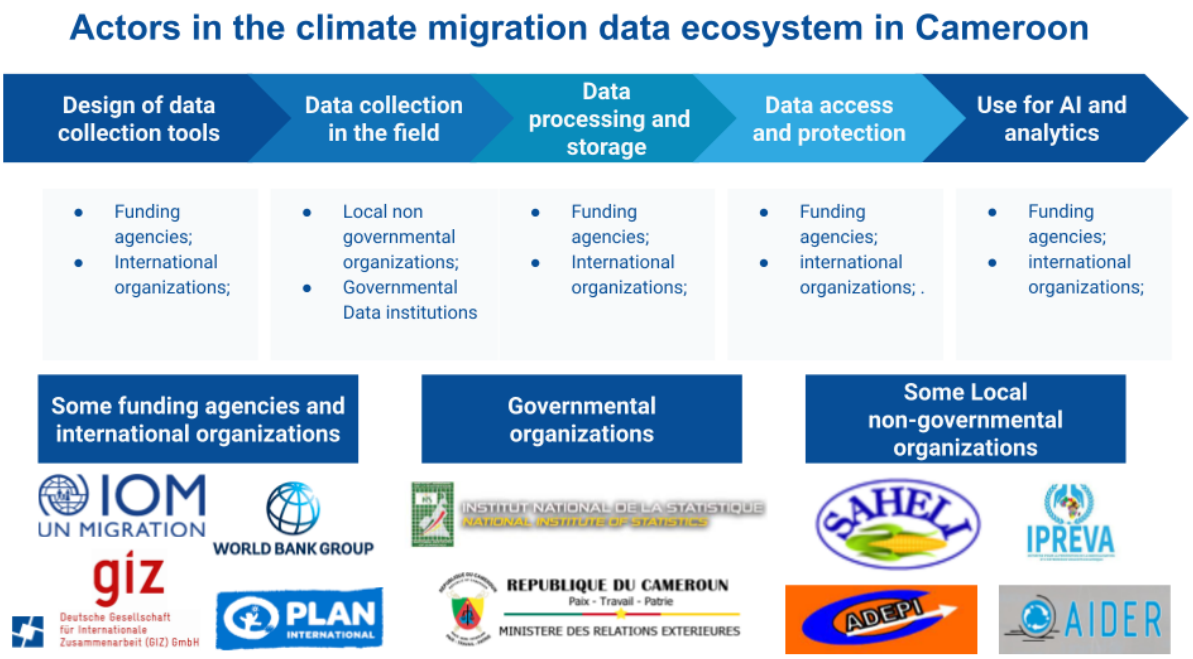

Figure 5 : *An overview of the data ecosystem*

Gaps and challenges

Describing the local data ecosystem has allowed us to better identify the challenges and issues facing the process of collecting climate-induced migration data in the Lake Chad basin. In this section, we describe these key issues and challenges. We identified two main categories of challenge: barriers to local participation, lack of data on climate-induced migration and sustainability of data collection activities.

Figure 6 : *Gaps and challenges*

Barriers to local participation

Restricted access to the data collected

Local organizations are unanimous on the fact that once the data is collected and transmitted to the international organizations, they (local organizations) are no longer involved in the rest of the process; the management of the data is done by the donors. This restriction of access to data is even more pronounced for local non-governmental organizations:

> As far as the handling of data is concerned, we really don't have access to the rest of the process; it's hard to tell you clearly how it is used. This information is generally internal to the international organizations. Once the data have been put on the digital support and sent, the follow-up does not involve the local organizations that were in the field to collect them (INF5).

Although they sometimes store a copy of the data on hard drives, local non-governmental organizations cannot use the data without the consent of the international organization for which they collected it.

In contrast, governmental organizations seem to have easy access to international organizations that hold this data. In this sense, INF1 states: "Given the fact that these international organizations are accredited by the Cameroonian government, If the state needs the data, the state can just send a correspondence. I don't think that should be a problem". Despite this possibility, the balance of power may vary depending on the sponsor involved. Indeed, INF2 (whose organization is in charge of data exploitation and archiving at the national level), indicates that when projects are financed by donors such as the World Bank, its organization is limited to data collection, while the database is managed by the donor.

However, INF6 did not fail to express its dissatisfaction with these international organizations and the lack of recognition of the contribution of local organizations.

> I think that the partners to whom we give our data should at least cite the sources and mention us. We are regularly victims of this type of problem. Another thing, at the level of the collection itself; they call themselves experts, they stay in their office and formulate things that have nothing to do with what is really happening in the field. This is a pity! You have to go to the source of the information. It is important.

Data protection: an unpleasant requirement!

When the database is owned and administered by a governmental organization, you can access the data, after making a request as INF2 explains: "If you want the data, you have to send a motivated request to the Director General, indicating what you want to do with it, as well as the variables you need". However, he points out that a certain number of precautions must be taken, avoiding giving the whole database to the user; moreover, the data must be protected.

> In the 2000s, for example, we practically gave the entire base when people asked. Unfortunately, sometimes we ended up with publications with indicators that were contradictory to what we produced in the official results. (INF2).

In terms of precautions, it should be noted that the Cameroonian government has recently adopted a new statistical law[10] which governs the confidentiality of data since its collection. However, the various organizations in the data ecosystem are not immune to the issues surrounding the protection of personal data or the use that could be made of it. In this respect, we could therefore agree with international organizations that decide to restrict access to their data, even to local organizations that have collected it. INF5 comments: "As far as the content of the reports is concerned, we have signed confidentiality clauses, for example, which we must respect and which prevent us from disclosing the information in any way". Similarly, INF1 explains that the Cameroonian government has memoranda of understanding with the UNHCR and the IOM on sharing data on refugees and asylum seekers: "We try to give figures without going into too much detail. This seal of confidentiality applies from the collection to the use of the data".

Nevertheless, there is discomfort between the Government and international organizations over the ownership of data collected in Cameroonian territory. Indeed, with the introduction of sophisticated technologies such as biometrics and mapping, debates about data ownership are being raised. These technologies are not always carried by the state, but by external companies. Note that it was only recently (April 2018) that Cameroon issued a decree on biometrics and identity card reform, including for refugees (INF3).

[10] Loi n° 2020/010 du 20 juillet 2020 régissant l'activité statistique au Cameroun

Lack of data literacy and reproduction of social inequalities

The idea of involving local populations in data governance is laudable. But with a little hindsight, we should ask ourselves if the population is aware of the importance of data? In this regard, information from the field suggests that people do not always know what data is, or how to use it to influence policies and decisions for their benefit. Therefore, people are not always interested in data as explained by INF5: “After collecting the data, the communities are generally not interested in the paperwork (reports). What the communities are looking for are the achievements. So even if you share that data, they are not interested”. This lack of data literacy extends beyond the population to the government and its departments, to the point that INF2 wonders: "Sometimes you wonder if these results are really used by the relevant people in the government. One could have the impression that these reports are done just because it is a requirement of the donors, or to obtain funding”.

Furthermore, culturally, we observe that the power dynamics that exist within African societies are reproduced in the data collection processes. Indeed, INF5 explains that during data collection in the Cameroonian part of the Lake Chad Basin, there are great disparities related to gender, resulting in difficulties for women to speak candidly in the presence of men. As a result, it would be difficult to collect data from women, despite the fact that they are in possession of a lot of relevant information.

Lack of data on climate change-induced migration

Climate-induced migration: an issue not considered in surveys

The issue of climate migration is not new in the Cameroonian context and the Government is aware of it as expressed by INF1: "We regularly manage cases of climate migrants. We are in the midst of climate change. Whether it is the absence of rain that forces people to move from one place to another to find water, we hear about it". Moreover, the climatic disturbances (drought, floods, etc.) at the origin of these climatic migrations are recurrent and well known by local populations, traditional and administrative authorities (INF2).

Despite all this evidence, there is very little or almost no academic work on climate migration in the Cameroonian context (INF3). Even worse, all the local organizations interviewed agree that they do not systematically collect data on climate-induced migration. The words of INF2 illustrate this state of affairs:

> Climate migration (for the moment) is not one of our priorities. It is true that we study migration in general with the different motivations of migration and others. But, specific migrations in relation to climate change... No, not yet (INF2).

This situation can be justified by the reasons mentioned in the previous section. On the one hand, even if they have the will to do so, the financial precariousness of local organizations could prevent them from collecting data on climate migration. On the other hand, international organizations have not yet sufficiently addressed the issue.

However, some of the interviewees recognize that by reading in detail the reports published by international organizations, one can find elements that deal with climate-induced migration, however, these data are not structured properly (INF3). This idea is supported by INF2, when he indicates some ways to find data on climate migration in different surveys:

> Currently, we are just completing another project on poverty among households. In this particular survey, there is a focus on mobility and migration. Perhaps, it is at the level of the reasons for migration that we can see whether people have migrated because of the climate, conflicts or the search for employment.

Difficulties in collecting this kind of data

According to some of the local organizations interviewed, we cannot really call this displacement climate-induced migration. This is because these communities do not migrate; they are only displaced internally—they are attached to their land, develop some means of adaptation, bypassing it, and other means of subsistence (INF4, INF5, INF6). For INF2, one could even associate this phenomenon with nomadism, which makes it difficult to have standardized data on climate migrants.

Indeed, they could be Cameroonians who move towards other Cameroonians. Sometimes they go directly to the homes of their relatives, living among their brothers and cousins (INF1, INF3). In this situation, it is complicated to identify who is a climate migrant and who is not; it would be easier to identify them if they were foreigners. This situation becomes even more difficult when financial or in-kind assistance is offered to these climate migrants; there is a risk of having distorted results, because everyone has the incentive to claim being climate migrants (INF1).

Another aspect that makes it difficult to collect data on climate migration is the fact that on the ground, the missions of international organizations are compartmentalized, a hindrance when faced with the need to act urgently. The example of INF6 on this subject is quite illustrative. Indeed, he wanted to alert the Red Cross to a risk of conflict between herders and farmers in the Logone et Chari area; but he was shocked by the response of the Red Cross, when one of the managers of this international organization told him that he could do nothing, because it was not their area of intervention. INF3 confirms this compartmentalization of missions and explains that it is the United Nations Office for the Coordination of Humanitarian Affairs (OCHA[11]), which assigns missions and delimits intervention zones to all international organizations involved in humanitarian action.

Sustainability of data collection activities

The financial vulnerability of local organizations

In the Lake Chad Basin, data collection on climate change and migration is largely funded by international organizations. They will offer consulting contracts to local organizations to collect data and return it to the databases they (international organizations) own.

> I was required to always upload the images to a platform managed directly by the funders. This data is important because it allows the supervisors to monitor in real time. These data (photos) are then integrated into the participatory reports (INF4).

In addition, it is important to mention that these data collection tools are always pre-designed by international organizations (INF3, INF4 INF5). This strong dependence on international organizations is one of the major weaknesses of local organizations, since they mostly respond to donor requests for data collection. INF2 admits that it is rare that his organization will conduct reflections on such an innovative theme as climate migration and make it a subject of research.

This financial dependence extends from data collection to the functioning of local organizations. Indeed, these organizations have based their economic model on providing services to international organizations. Thus, the more services or consultations a local organization provides, the better off it is and the more attractive it is to international partners.

> When a funder comes with its funding, some come directly to our organization. If we have work on one subject or another, they can help us with funding. They don't have to go looking everywhere anymore (INF6).

[11] https://unocha.org/

Based on this evidence, it is clear that the process of collecting data on a theme depends rather on the interest and financial means that international organizations make available to local organizations (government, civil society, local populations). In other words, if there is no special program financed by international partners, local organizations will not initiate data collection on an issue as important as climate migration.

A lack of frank collaboration at the national level

From the various interviews, it is clear that local governmental organizations do not really work with local non-governmental organizations, but much more with international organizations. The words of INF1 are quite evocative in this regard: "We don't really work with local associations. We work much more with anything external to the country. In general, it is these international organizations that work with the local associations".

This lack of collaboration can also be seen within the government apparatus itself. The words of INF2 perfectly illustrate this notion, when he explains that in his organization, there is a whole division that deals with climate issues, as well as another that deals with migration. However, they have never thought of bringing the two divisions together to deal with the climate migration component.

Speaking of non-governmental organizations, INF5 believes that collaboration is limited to the exchange of information on methodological and technical levels. As for collaboration on data, its answer is quite scathing: "everyone manages their own problems, so we remain focused on the work that has been assigned to us".

Thus, at the local level, governmental and non-governmental organizations operate in closed silos, which is not conducive to collaboration to find common solutions to problems affecting communities.

Opportunities

The obstacles to collecting data on climate-induced migration described in the previous section are not insoluble. Indeed, there are several opportunities that would allow us to overcome these limitations and implement our project on an even more global scale. Among the most promising opportunities we have, we can mention: solutions proposed by the local organizations interviewed and possible ways of implementing the proposed solutions through international initiatives on climate migration; networking opportunities; South-South collaboration, and connection with other projects of the GPAI's working groups.

Solutions proposed by local organizations

In order to address the gaps and challenges noted above, the local organizations interviewed made multi-faceted calls to build concrete solutions adapted to local realities. Among other solutions, they are:

- ***Improving the data literacy of local populations and communities would increase their effective participation in the climate migration data collection process.***

To achieve this, we need a two-pronged approach. Firstly, strengthen the capacities of individuals, by accompanying local populations with a special training program on data. Secondly, train local associations in data governance, and organize them into consortiums so that they can better defend their right to access and use data (INF3, INF4, INF5, INF6). Usually, local people are receptive to getting involved when you explain the merits of an initiative. Better yet, it empowers and gives them the feeling that they are playing a role in the community.

- ***Rethink data collection tools, so that they collect as much relevant information as possible to develop AI solutions to the climate migration problem.***

The idea is to add new categories of questions to the standard donor questionnaires, so that they can effectively capture climate change-induced migration issues. These new questions could address issues such as the integration of migrants in new places, endogenous strategies to mitigate the effects of climate change, etc. (INF3). And, to add other data collection methods such as mapping, to limit farmer/breeder conflicts and to consider the distribution of land by creating spaces dedicated to grazing, nurseries, etc. (INF6).

- ***Create a centralized national database that would consolidate all information on climate migration.***

All the organizations interviewed agree that there is a strong need for a digital platform supervised by the State of Cameroon, where each organization that collects data on climate migration can store it. Thus, local and international actors will be able to share this national database that will allow them to make decisions and predict climate migration (INF1, INF2, INF4). With this in mind, INF2 states that his IT department wants to develop systems that allow the user to make faster queries to obtain (or exploit) the data. However, INF1 stresses the importance of this database being state-controlled, with access limited to certain key players.

Implementation strategies

Global initiatives on climate-induced migration

Climate action is already underway under the leadership of the United Nations to address both climate change issues and its collateral consequences such as migration. This is an anchor point for our study, mainly because our results can be used and implemented on a much larger scale than Cameroon and the Lake Chad Basin. In the African context, two initiatives led by the United Nations in collaboration with the African Union caught our attention and present interesting avenues for future collaboration for GPAI/CEIMIA on the basis of this exercise. These initiatives are: The Africa Climate Mobility Initiative[12] and The Sahel Predictive Analytics (Sahel PA) Project[13].

- ***The Africa Climate Mobility Initiative***

The Africa Climate Mobility Initiative (ACMI) is a joint undertaking of the African Union Commission, the United Nations, and the World Bank to support the African continent in addressing climate-forced displacement and migration while harnessing climate mobility to advance its development agenda and further its economic & political integration[14]. Guided by three core tenets (Plan for Mobility, Empower People & Transform Development), the ACMI recommends eight key actions for the next eight years (2023-2030) and calls for partnerships to take these actions.

Table 3: ACMI key actions (2023 - 2030[15])

No	Actions	Description
1	**Anticipate and plan for climate mobility**	Anticipate and plan for climate-forced displacement and migration, including permanent relocation, to foster social cohesion in affected communities, prevent involuntary immobility, drive economic growth, sustain peace, and protect people on the move.
2	**Integrate climate mobility in climate**	Recognise and support mobility as a legitimate strategy for climate adaptation in local, national, regional, and international policies, and

[12] https://africa.climatemobility.org/

[13] https://unis-sahel.org/2022/11/02/sahel-predictive-analytics-report-moving-from-reaction-to-action-anticipating-vulnerability-hotspots-in-the-sahel-in-support-of-uniss/

[14] https://environmentalmigration.iom.int/sites/g/files/tmzbdl1411/files/event/file/acmi-executive-summary-final.pdf

[15] https://africa.climatemobility.org/agenda-for-action

No	Actions	Description
	action and finance commitments	build cross-sector partnerships to support people and communities in staying, moving, and receiving.
3	**Inform people of climate risks**	Enhance public understanding of climate risks and threats, including through building climate change literacy, the co-production of actionable climate information services and access to early warnings, to support informed decisions on how to adapt, whether and when to move, and where to settle.
4	**Amplify Women's agency**	Empower women with climate information, adaptive skills, social and legal protection to bolster their agency in decisions on climate adaptation and in climate mobility.
5	**Harness the ambitions of the Youth**	Foster and leverage the creativity and potential of Africa's already-mobile youth, to build resilience, economic prosperity, and advance the green transition.
6	**Build from the Local**	Pursue community-led solutions for climate-resilient development, disaster response and climate mobility across the continent, and invest in locally anchored climate adaptation and resilience pathways, including strong connections in border areas.
7	**Invest in resilient and connected cities**	Enable cities with the actionable data, financial and technical resources, and political agency needed to facilitate planned, resilient, and inclusive urban growth, social inclusion, and social protection, while building stronger ties across cities and between cities and rural areas and economies.
8	**Pursue Nature-positive development**	Manage land, water, and other shared natural resources cooperatively and sustainably to support agricultural and ecosystem-based livelihoods and boost productivity, while reducing environmental impact and harnessing ecosystems and biodiversity protection for economic development and climate resilience.

- ***The Sahel Predictive Analytics (Sahel PA) Project***

Funded by the German Federal Foreign Office and facilitated by UNHCR in support of the office of the UN Special Coordinator for Development in the Sahel. The Sahel PA project doubles down on the special challenges facing the Sahel across the triple nexus of humanitarian aid, peace-building, and development with the aim of guiding decision-makers by anticipating and quickly identifying where multiple risks overlap to allow for better preparedness and to support context analysis, planning, training, and capacity-building, while also outlining where additional data is needed[16].

Based on their report, the Sahel PA project has elaborated three key recommendations to improve anticipatory action and preparedness in the Sahel: data availability and quality, adaptation, and governance[17].

[16] https://viewsforecasting.org/about/sahel-pa/

[17] https://unis-sahel.org/wp-content/uploads/2022/11/SAHEL-PREDICTIVE-ANALYTICS-REPPORT-compresse.pdf

Table 4: Sahel PA key recommendations

Recommendations		Specific Actions
1	**Data availability and quality**	- Building a network of data-sharing (preferably digital) hubs to make existing data publicly available - Support the data landscape of the Sahel countries (statistical offices, meteorological agencies and ministries) in collecting data in a standardized manner and to facilitate timely exchanges; - Allow for consecutive survey waves and consecutive qualitative data gathering to monitor displacement and conflict over time and make more accurate inferences about implications for the future; - consider the large penetration of mobile devices as new opportunities for real-time data collection and sharing and allows for the inclusion of perspectives from marginalized communities; - Involve local experts and stakeholders in data collection, model development and scenario-building; - Joint international research with institutions in the Sahel.
2	**Adaptation**	- Using participatory approaches and indigenous knowledge is key; - Leveraging lessons learned and best practices while building on and enhancing coordination across existing regional and national structures; - Investing in national social protection systems; - Investment in disaster prevention and disaster risk reduction is needed; - Improving safe pathways for migrants and protecting the rights of people - mobilize funding for displaced and trapped populations; - Access to domestic energy is absolutely essential; - The establishment of a win-win partnership with the private sector to support both the supply and demand of gas is essential.
3	**Governance**	- Strengthening both the rule of law and the inclusivity of public institutions and social networks are important measures in addressing conflict risks; - improve inclusiveness and ownership within governmental institutions in order to foster greater representation of marginalized groups. - Strengthening collaboration with local elected officials through area-based approaches at all stages of defining and implementing strategies and interventions is crucial; - improving intercommunal relationships, particularly between displaced and local populations, can help reduce potential conflict. One suggested way to do this is through participatory land use planning and mapping; - Addressing pastoral grievances; - Community and vocational training centres can help improve communication and ease intercommunal tensions; - improve intraregional freedom of movement.

Networking and collaboration

The theme of climate-induced migration has led us to meet other organizations that have shown interest in our work in the Lake Chad Basin and have expressed their willingness to join our efforts. The Peer Learning Network and Data Portability are initiatives that we engaged with in a certain way, to get feedback on our work.

The Peer Learning Network

The Open Data Institute (ODI) and Microsoft are running a peer-learning network[18] that provides cohorts of organizations ready to improve data practice and increase data sharing with opportunities to connect with peers, receive guidance from data and sector experts, funding and other support. The purpose of the peer-learning network is to convene a cohort of organizations that are working to collect, share and/or use data for climate resilience and sustainability.

The Peer Learning Network provides participants with funding to enable them to learn from data experts, and from one another, in order to more effectively address the challenges they face through a better understanding of relevant data topics. The CEIMIA team built a consortium with local organizations in Cameroon, applied for this funding with the aim to develop a framework enabling trustworthy data exchanges.

By participating in this Peer Learning Network, we hope to learn from the experience of other organizations that successfully established data collaborations and to receive feedback on the framework we co-develop. In particular, the National Institute for Statistics in Cameroon would benefit from guidance to put mechanisms (or infrastructure) in place to collect and aggregate data on Climate Migration from different local organizations (possibly supported by IOM) in a sustainable, responsible and participatory manner. Local non-profit organizations can be empowered to realize their role in the data value chain. All participating organizations would improve their awareness of best practices in data governance. The table below presents the schedule and here are the organizations that will take part in this session:

- Women Income Network
- Global Partnership for AI/CEIMIA
- Gender and Environment Data Alliance
- DataKind
- South African Cities Network
- PlanAdapt

Table 5 : Schedule of the ODI Microsoft Peer Learning Network - Data for Climate Resilience

Dates (2023)	Session
Tue 10th Jan	Kick off session
Mon 6th Feb	WS1: Value of data sharing roundtable
Tue 28th Feb	WS2: Modeling data ecosystems
Tue 21st March	WS3: Sustainable, responsible and participatory data stewardship
Tue 11th April	WS4: Data infrastructure and maturity
Tue 25th April	WS5a: Data in built environment - WS5b: Data in water resilience
Tue 16th May	Wrap up session
Wed 24th May	PLN ends - participants publish final reflections

- ***The Data Portability initiative***

Data Portability[19] is the ability of a person to obtain and reuse the data they have provided to one organization (and IT system) for their own purposes across different services and organizations (and different IT systems). Data portability is different from interoperability in that interoperability

[18] https://www.theodi.org/article/open-call-data-for-climate-resilience-peer-learning-network-join-us/
[19] https://www.collaborativecash.org/dataportability

tends to focus on systems sharing data and working together, while portability puts the person at the centre of the sharing.

Funded by the European Union Humanitarian Aid, the Data Portability initiative is led by the Collaborative Cash Delivery Network and seeks to identify and lay the human capacity and governance groundwork for data portability and 'putting people at the centre of data management'. The project aims to address the issues of literacy, politics, and culture which are the real barriers of data sharing, not the technology. To this end, the project has two specific objectives:

- Increase the Digital Literacy of frontline project staff of CCD agencies and their local partners.
- Develop Digital Governance models and pilots to enable data stakeholders the ability to use the data we have about them in other areas of their lives.

South-South collaboration

Beyond all the existing international initiatives and new collaborations that we have been able to secure, it should be noted that GPAI and CEIMIA are very keen on South-South collaboration. In this sense, one of the most promising avenues is our collaboration with ZMQ Global[20] through their branch ZMQ Africa Ltd.

ZMQ Global is an international organization based in India. ZMQ Global creates and implements practical technology-based tools and solutions to empower communities by providing timely information and connecting them to vital services, thereby achieving sustainable development. ZMQ's model combines social mission and technology to maximize the impact of health, education, and livelihood interventions. ZMQ therefore works with communities, local NGOs, community workers, and government infrastructure to facilitate the exchange of knowledge and sharing of delivery and connection services on social innovation practices to improve the lives of rural communities.

Currently, ZMQ has expertise in the field in terms of the Human-Centric approach. Indeed, the programs underway at ZMQ are rich in lessons that could be transferred and applied to our project on data institutions for climate-induced migration. These programmes are:

Table 6: Ongoing programmes at ZMQ

Programmes	Description
MIRA Channel	Also called Women Mobile Lifeline Channel, is an integrated cell phone channel to provide health information to rural women and connect them to public health services using cell phones in low-resource settings. MIRA's goal is to enable women to improve health indicators through self-management of their health and to achieve a broader goal of empowering women through digital connection.
Freedom TB	Freedom TB is a ZMQ technology-based program aimed at improving overall Tuberculosis treatment, adherence, and compliance.The program uses its innovative approach called Active Patient Compliance System to improve TB treatment and adherence using a mobile phone-based observed video therapy (VOT) approach. It is a system that changes the way TB patients report adherence to treatment from their villages and homes.
Storytelling and comics	ZMQ uses innovative digital approaches based on social and behavioral change communication to raise awareness of critical health issues. To this end, ZMQ uses storytelling and creates "Talking Comics" for rural women with low literacy skills to instill healthy behaviors.

[20] https://www.zmqglobal.com/

Connection with other GPAI projects

Several projects are currently underway within the GPAI. However, considering the first results emerging from our fieldwork, it would be beneficial for us to work closely with two of them, namely: Data Justice and RAISE (Responsible AI Strategy for the Environment).

- **Building on the learnings of the Data justice Project**

This project was established to fill a gap in data justice research and practice and provides a framework to help policy makers, practitioners and users move beyond understanding data governance narrowly as a compliance matter of individualized privacy or ethical design. Rather, considerations of equity and justice specifically are considered as they relate to redressing the uneven distribution of opportunities, and harms, associated with AI and ML.

The data institutions project can learn from data justice principles and put them in practice, in order to identify risks related to data colonialism and data extractivism. Moreover, in collaboration with the data justice project we can set up strategies contextualized to Lake Chad Basin, in order to ensure more equitable access, greater visibility and fairer representation of those individuals and communities marginalized from data used in the development of AI/ML systems.

- **RAISE (Responsible AI Strategy for the Environment)**

Project RAISE aims to develop and operationalize a global responsible AI adoption strategy for climate action and biodiversity though:

- Implementing strategic roadmap and recommendations to inform policy and high-impact use cases;
- Developing an impact and risk assessment framework harnessing AI for climate action and biodiversity preservation responsibly;
- Work with institutional partners to anchor AI for environmental action at COP, member states, and other international bodies

RAISE's work can be very helpful in designing the framework for trustworthy data institutions.

Conclusion and next steps

This project aims to explore how data institutions and AI applications could make a difference by empowering local organizations & communities to contribute in building evidence and developing solutions by participating in problem identification, data collection and analysis. That is why in this pilot study, we mapped the data ecosystem by identifying the different stakeholders, described the dynamics in data collection, identified gaps and challenges and presented some opportunities to mitigate these challenges. But presented in that way, there is a risk of missing to take concrete actions to implement trustworthiness in data institutions. It is therefore crucial to develop a framework for trustworthy data exchanges that improves how data is being collected, stewarded, shared and used, to better serve the needs of communities and empower them to play an active role in the data value chain. As part of the ODI/Microsoft Peer Learning Network, our next step is to co-design a trustworthy data Institution framework that would be applicable to any situation (besides climate-induced migrations).

Bibliography

Ada Lovelace Institute. (2021). *Exploring legal mechanisms for data stewardship*. https://www.adalovelaceinstitute.org/blog/disambiguating-data-stewardship/

Geuns, J. van, & Brandusescu, A. (2020). *What Does it Mean? Shifting Power Through Data Governance*. Mozilla. https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/

GPAI. (2021). *Enabling data sharing for social benefit through data trusts* [An Interim Report for the 2021 GPAI Paris Summit]. https://gpai.ai/projects/data-governance/primer-on-data-and-economic-justice.pdf

IAEA Project. (2017). *Integrated and Sustainable Management of Shared Aquifer Systems and Basins of the Sahel Region*.

IMDA Singapore. (2019). *Trusted Data Sharing Framework*.

Leon Usigbe. (2019). Drying Lake Chad Basin gives rise to crisis. *Africa Renewal*. https://www.un.org/africarenewal/magazine/december-2019-march-2020/drying-lake-chad-basin-gives-rise-crisis

Open Data Institute. (2022). *Measuring the impact of data institutions | report*. https://theodi.org/insights/reports/measuring-the-impact-of-data-institutions-report/

Open Data Institute. (2021). What are data institutions and why are they important? *The ODI*. https://theodi.org/news-and-events/blog/what-are-data-institutions-and-why-are-they-important/

Soni, S. (2020a, June 12). Explorer | The Data Economy Lab. *The Data Economy Lab*. https://thedataeconomylab.com/stewardship-explorer/

Soni, S. (2020b, July 31). Data Sharing for Public Good: Theoretical Bases and Policy Tools. *The Data Economy Lab*. https://thedataeconomylab.com/2020/07/31/data-sharing-for-public-good-theoretical-bases-and-policy-tools/

ANNEX

Interview guide

Presentation/objective

- Presentation of the researcher

- Context : Climate change is leading to conflicts, causing disasters, driving displacement and making life harder for refugees in settlements. Several organizations (local, regional and international) are collecting, stewarding, exchanging and using data to better understand key drivers of migration and the impact of Climate Change, and improve migrant's conditions.

- Interview goal: We would like to better understand the role of individual organizations in the data ecosystem (collecting, stewarding, or using data) and their data governance practices, the relationship and data exchanges between different players and the degree to which the most affected communities are included in this process.

- Project goal: This initial assessment would lay the ground for the ultimate goal of the project: to co-design a framework for trustworthy data governance and sharing respectively for the organization and the overall climate migration data ecosystem, empowering local communities to play an active role in the data value chain.

- Process:

◦ *Approximately 1 hour of discussion*

◦ *Approach: The interview consists of 5 sections/themes. I would appreciate it if you could share your perspective as freely as possible.*

◦ *Conditions: If you agree to give me your permission, I would like to record our conversation so that I can listen to it again and make an analysis when the time comes. I assure you that the recording will be kept in a safe place. Furthermore, only I and the research team will be allowed to listen to the recording. All recordings will be deleted at the end of the research.*

◦ Signature of consent form.

Theme 1: Awareness about Climate migration

◦ what motivated your interest in this issue?

◦ what are you doing?

◦ what aspects of climate migration interest you?

◦ What other organizations do you know that are working on this issue in the Lake Chad Basin?

◦ What local communities do you work with to get information around climate migration? Are they aware of the issues around climate migration?

Theme 2 : Data governance

1) Data collection process

- Can you describe how you go about collecting data on climate migration?

- What kind of data do you collect? (e.g. personal data on migrants/displaced communities, environmental data, etc.) What are the different attributes (e.g. text, imagery, video, sound, etc.)?

- What techniques and tools do you use to collect this data? (e.g. questionnaires, portal,

- Do you collect data directly from climate migrants? If so, what are the difficulties you face in the process and how do you overcome them? Do you talk about the importance of data with climate migrant communities? (for this last question, we can refer to the data literacy section)

Transition question: Once the data is collected, what do you do with the data?

2) Data storage, access, exchange/sharing

a) Storage

- Where and how do you store the data you collect?
- What technical systems (platforms, software, etc.) are you using to store your data?
- In which format are you storing your data?
- Before storing data are you proceeding with data quality assessment?
  - If yes, can you describe the process and the criteria you are using?

b) Access

- Do you have any dataset?
- What is the intellectual property regime that governs access to this database?
- Under what conditions is the dataset publicly accessible?
- Is the dataset accessible to everyone?
- Or it is only accessible to specific stakeholders or communities?

- What are the conditions to have access to this dataset?
- Which people or organizations are responsible for managing or ensuring access to a dataset?
- Which people or organizations are responsible for managing or ensuring access to a dataset?
- Are there any barriers (technical, legal, cultural) to making the dataset publicly available?
- What do you propose to overcome these barriers?

c) Exchange/sharing

- Which people or organizations contribute to the dataset?
- Are there any barriers (technical, legal, cultural) to making this data publicly available?
- With which organizations are you sharing your data?
- From which organizations are you receiving your data?
- Are you monetizing data? If so, who keeps the profits? Are any benefits shared with data subjects?

3) Data sensitivity/privacy and legal compliance

- Which kind of personal data are you collecting (Identity, health data, biomedical data)?
- What legal and technical means does your organization have in place to ensure the protection of this data?
- Do you also collect data from indigenous populations? If so, what measures are you taking to ensure that rights held by indigenous peoples and of other group rights over data (data sovereignty) are respected?
- When it comes time to deal with legal compliance,
    - do you have any data protection policy?
    - do you have some data-sharing agreements with other organizations?
    - do you have some agreements with other organizations, on data ownership and other intellectual property rights?

Theme 3: Opportunities and Planned uses of data

- What are you planning to do with all the data you collect?
- What services or added value does your dataset offer (or you are planning to offer) to people, local communities ?
- Who uses your data to create things (products, services, analyses, insights, stories or visualisations)
- Which people or organizations benefit from this dataset because it enables them to make decisions.
- What are the impacts of changing how data is accessed, used and shared?
- What aspects need to be improved so that we can make good use of the data on climate-induced migration?
    - What methods could be used to improve data flows between actors in the climate-induced migration data ecosystem?
    - Which potential users and communities could benefit from new data infrastructure in a sector?
    - Which new stakeholders should be involved?
- What aspects need to be improved so that we can make good use of the data on climate-induced migration?

Theme 4: Data literacy (specific to individuals or local communities)

- When people talk about data, what do you think of? What does it mean to you?
- Do you know why and for what purposes organizations collect data from you?
- Do you always feel comfortable and agree that your data is being collected?
    - If not, why?
    - If so, what benefits do you see in sharing your data?
- Once your data is collected, do you know where it goes? And do you always have access to it?
    - If so, how do you access it?
    - If no? What prevents you from accessing it? Would you like to have access? And what recommendations would you make to these organizations that collect this data, so that you can easily access it?
- What uses do you think could be made of the data collected from you?
- What uses do you want to make with the data collected from you?

Theme 5 - Participant portrait

Objective: to evaluate the posture adopted by the participant.

- Interest in this study
  - How did you hear about this study?
  - What made you want to participate in the research?
- Could you tell me a little more about yourself...?
  - field of study, occupation, role in the organization they are representing
  - age, place of living
- Tell me about your experience with climate migration

Thank you for your time and participation!

Consent form

Presentation of the research project and the researchers responsible

The goal of the research is to study the exchange of data on climate migration between your organization and the local communities from which data are collected. It aims to co-design a framework for trustworthy data exchanges within the climate migration data ecosystem, improving how data is being collected, stewarded, shared and used to better serve the needs of communities and empowering them to play an active role in the data value chain.

This project is funded by GPAI/CEIMIA and it is carried out by Lama Saouma, AI Initiative lead at CEIMIA; Thomas Hervé Mboa Nkoudou, Researcher in residence at CEIMIA; and Teki Akuetteh, GPAI Expert.

Before agreeing to participate in this research project, please take the time to read and understand the following information. This document explains the purpose of this research project, its procedures, benefits, risks and drawbacks. We invite you to ask any questions you may have to the person presenting this document or to the researcher in charge, whose contact information is listed below.

Participation Procedure

Participation will consist of taking part in a one-on-one interview with the responsible researcher, lasting between 30 minutes and 01 hour, which will cover the following: Awareness about Climate migration - Data collection - Data governance - Data literacy.

The interview will be recorded (sound only).

Advantages, disadvantages and compensation

Participating in this research offers you an opportunity to reflect your opinion and discuss in confidence how data is being collected, stewarded, shared and used.

During the interview, you are free not to answer or to answer only partially to certain questions that would cause you discomfort.

We will try to minimize the amount of travel required by letting you determine the most convenient location for the interview. If necessary, the individual interview can be conducted by telephone or via the Internet.

Confidentiality and data management

The following measures will be taken to ensure the confidentiality of the information provided by participants:

- the names of participants will not appear in any report;
- The various research documents will be coded and only the researchers will have access to the list of names and codes;
- Only researchers and their assistants will have access to data;
- Members of the research team are bound by a written commitment to confidentiality;
- The research will be published in reports, scientific journals and books, and no participant will be identified or recognized in the publications;
- All research materials, including lists, consent forms, electronic copies and recordings, will be destroyed 18 months after the research project is completed. Only irreversibly anonymized data will be retained for future use in other research.

For more information

If you have any questions about the research or the implications of your participation, please contact the researchers in charge: Lama Saouma, AI Initiative lead at CEIMIA (lama.saouma@ceimia.org) and Thomas Hervé Mboa Nkoudou, Researcher in residence at CEIMIA (thomas.nkoudou@ceimia.org).

Acknowledgements

Your collaboration is invaluable for the realization of the research. The research team appreciates the time and attention you have given us for your participation. We thank you for your participation.

Signature
I, the undersigned _____ freely consent to participate in the research project entitled: **Data institutions for climate migration in Lake Chad Basin**. I have read the application form and understand the purpose, nature, benefits, risks and drawbacks of the research project. I am satisfied with the explanations, clarifications, and answers provided by the researcher, if any, regarding my participation in this project.

Participant's signature _____ Date _____

I explained the purpose, nature, benefits, risks, and drawbacks of the research project to the participant. I answered the questions posed to the best of my knowledge and verified the participant's understanding.

Signature of the researcher _____ Date _____

Complaints and criticism
Any complaint or criticism about this research project can be addressed to our Executive Director, Sophie Fallaha (sophie.fallaha@ceimia.org)
Centre d'Expertise International de Montréal en Intelligence Artificielle,
7260 Rue Saint-Urbain,
Montréal, QC H2R 2Y6,
Canada