

CAHF Data Agenda for Africa

The mission of achieving access to affordable housing finance across the continent is severely hindered by the lack of publicly available, credible information on housing finance in many African countries. Since its establishment, CAHF has focused heavily on data—its collection, interpretation and sharing—in order to draw market intelligence which can be used to promote investment and improve policy making in the affordable housing sector. This document sets out CAHF’s broad, high level data agenda for Africa.

A core requirement of the Data Agenda for the fulfilment of other aspects of CAHF’s mission, is the need to ensure that the data CAHF collects, analyses and shares is credible and accessible. To achieve this, CAHF requires a coherent, strategic approach to data which recognises how all our projects both depend upon and help to promote our Data Agenda. The Data Agenda for the continent clarifies how CAHF’s data work is located within its broader mission, and sets out clear objectives over a medium-term period (5 years). It is an internal document, providing guiding principles to frame how we work and how we prioritise opportunities.

This document is a living strategy, reflecting CAHF’s most recent thinking on its Data Agenda. The document will be revisited at least annually, and updated to reflect developments in the data landscape as well as CAHF’s internal needs, the needs of its partners, and progress on the implementation of identified priorities.

1. Goal of the Data Agenda and links to CAHF’s Theory of Change

As a first step, it is useful to explicitly revisit CAHF’s broader mission and strategy. CAHF’s Data Agenda finds its roots in the organisation’s broader vision.

CAHF’s vision, mission and goal

The **vision** of CAHF is an enabled affordable housing finance system in countries throughout Africa, where governments, business, and practitioners work together to provide a wide range of housing options accessible to all.

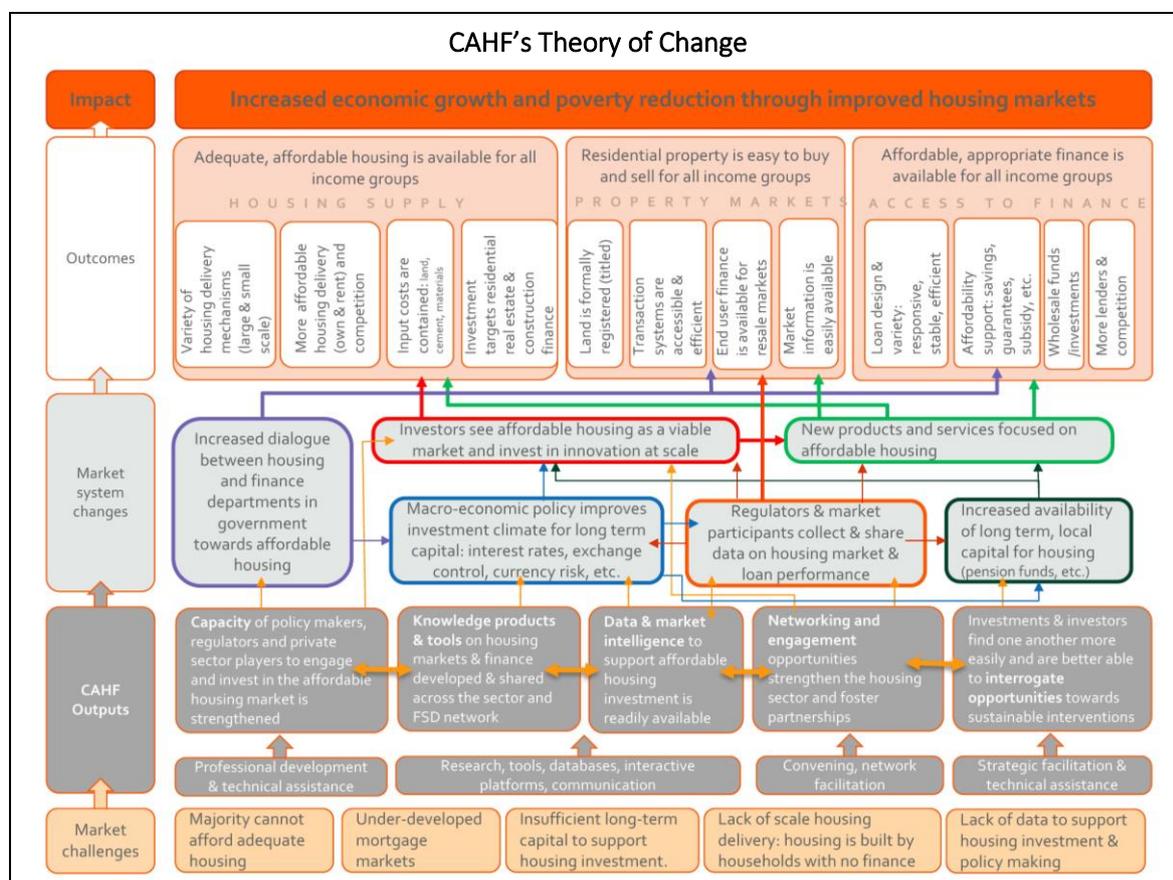
CAHF’s **mission** is to make Africa’s housing finance markets work, with special attention to access to housing finance for the poor. We pursue this mission through the dissemination of research and market intelligence, supporting cross-sector collaborations and a market-based approach.

The overall **goal** of our work is to see an increase of investment in affordable housing and housing finance throughout Africa: more players and better products, with a specific focus on the poor.

Data is at the heart of CAHF’s Theory of Change (TOC). Lack of data is one of the five key market challenges identified by CAHF, which it seeks to address explicitly with targeted outputs. As per the TOC below, these core outputs include:

- Knowledge product and tools on housing markets and finance developed and shared across the sector and FSD network
- Data and market intelligence to support affordable housing investment is readily available

The TOC then recognises that these two data-related outputs would directly contribute to market system changes including macroeconomic policy improvements and the willingness of regulators and market participants to collect and share data on housing and housing finance markets. These shifts would then lead to new products and services focused on affordable housing, and more investors viewing affordable housing as a viable market.



In line with this, CAHF’s **data vision** is the development of a housing and housing finance data universe in Africa that provides regulators, policymakers, developers and housing financiers with accurate, relevant and current data.

CAHF’s **data mission** is to catalyse the provision and dissemination of this data and to support the development of a community of in-country housing and housing finance practitioners who are aware of the importance of accurate data and who have the capabilities to engage with data to generate relevant insights.

In short, CAHF would like to create a data-driven approach in the affordable housing sector.

2. Data Agenda objectives

In line with its TOC, CAHF has already developed a number of knowledge products and tools, and databases. CAHF has also invested heavily in uncovering and disseminating data on housing and housing finance markets that is not currently available, or is not sufficiently accurate or granular to inform decision-making.

Over the next five years (to March 2024), CAHF is committed to pursuing the following **Data Agenda objectives**:

1. Collect, assess and curate available data as relevant to CAHF's overall mission and increased investment in affordable housing across Africa.
2. Identify data gaps that inhibit market activity and close key data gaps either by
 - Encouraging key data owners to make existing data available (existing or secondary data that exists but is not disseminated)
 - Encouraging key partners to improve the quality or coverage of data they collect
 - Commissioning partners to gather additional data (primary data)
3. Integrate, optimise and expand its set of knowledge products and data tools to produce a set of key data outputs that are widely referenced and that directly support CAHF's broader objectives
4. Build the capacity of partner institutions to analyse data and engage with data-driven outputs
5. Promote the importance of accessible, credible data as a catalyst in the affordable housing market

Each of these three areas is discussed briefly:

2.1 Collect, assess and curate available data

Due to CAHF's extensive experience collecting data on housing finance across the continent, it is well-placed to assess the quality and coverage of that data. It is also able to compare alternative data sources and data collection strategies across these various markets, and to use this knowledge to build the data capacity of its partners.

CAHF also stores and curates the data it collects for reference internally by its own analysts as well as by external partners. The way it manages its own data is therefore critical.

2.2 Identify and close key data gaps

Many entities (many of which are key CAHF partners or stakeholders) already collect data that is useful for supporting housing investment decisions. These include the following public entities:

- Central Banks who might collect data from Financial Services Providers and who typically track foreign investment and financial flows, as well as the progress and performance of mortgage and other credit markets in their respective economies;
- National statistical agencies who conduct household surveys and censuses and collect data on demographics; housing circumstances and conditions; price levels (including housing and construction prices); and other factors;
- Land registries who record property ownership and transactions;

- Revenue authorities who monitor personal income tax, VAT payments and imports;
- Capital market authorities who maintain, monitor and publish on the progress and performance of national capital markets;
- Publicly funded loan liquidity facilities;
- Public housing companies, whether at national, state or city level;
- Municipalities who grant planning and other forms of statutory permission for construction and investment in housing;
- Pension fund authorities who might monitor the investment activity of pension funds within their own local areas as well as globally; and
- Other public agencies as might be particular to a local context.

In addition, private sector organisations also generate data in the course of their activities. These include:

- Investors, capital market participants, and others (whether local, regional or international) who provide investment capital into various links of the housing delivery and housing financing value chains;
- Credit providers and credit bureaus¹ who track credit granted and performance;
- Other financial services providers (FSPs) including insurers, pension fund administrators, savings and credit cooperatives, and other deposit-taking institutions;
- Developers, housing cooperatives, and builders who create housing stock;
- Landlords and property management agents who offer and manage rental housing stock;
- Estate agents and other market makers who match supply and demand; and
- Other private organisations as might be particular to a local context.

As CAHF works with these players, we identify relevant data sets and data owners, and leverage the trusted relationships we have developed (while building new relationships as well) to encourage participation in the Data Agenda for Africa with the provision of more data, more often. A key aspect of CAHF's advocacy in this regard is demonstrating the importance of the data they generate (often as a by-product of their core activity) to the development of housing and housing finance markets.

In many cases, however, the necessary data is not available or the quality or coverage of the data is limited. As CAHF identifies these gaps, we commission targeted research to generate this data, typically directly from market participants. In most cases, CAHF is likely to replicate the process of collecting data, each time building on key learnings and experiences and driving efficiencies and improved quality of the output.

Critical data is incorporated into specific knowledge products, building their usefulness and strengthening CAHF's position as a trusted partner and source.

2.3 Integrate, optimise and expand CAHF's existing knowledge products and data tools

CAHF's existing mature data products include **CityMark** and the **Yearbook**. The products have established audiences and have been through a number of update iterations. There is significant scope to build on these products both in terms of content and process.

CAHF has also laid the groundwork for a **Data Portal** that would house relevant data sets and facilitate access to data by its partners and more broadly by market participants. The portal will be

¹ In some countries credit bureaus are publicly-owned entities.

the point of entry through which all CAHF's data products are accessed and explored. Aside from enabling access to data, the portal will be used to develop and launch dashboards and showcase research content generated by CAHF and other partners. In the future, it is intended that the portal will crowdsource market analytics, allowing users to analyse, engage and explore data across datasets, sharing their questions and their findings towards a better overall understanding of affordable housing market dynamics in Africa.

2.4 Build the capacity of partner institutions to analyse data and engage with data-driven outputs

While CAHF's data portal will enable partners and sector participants to access data, CAHF also uses its data as a basis for its engagement and networking with stakeholders. Through its on-going interaction and engagement CAHF will build the capacity of policymakers and other sector participants to find and analyse data, and to use data more deliberately in decision-making processes. In this regard, CAHF is also able to co-host engagements and training sessions with the AUHF, in order to expand reach and impact. These engagements provide CAHF with opportunities to encourage still more data to be shared and referenced in support of evidence-based policy making.

CAHF also seeks to achieve this by creating local, regional and continent-wide 'communities of interest' comprising local researchers, consultants and sector participants who can engage thoughtfully and in a data-driven way on housing and housing finance opportunities. CAHF will identify opportunities to host webinars, in-country or regional workshops to build data partnerships, and grow specific data gathering or analytical skills within its partner base.

2.5 Promote the importance of accessible, credible data as a catalyst in the affordable housing market

Together, CAHF's data products and approach reinforce and elevate the importance of data in the development of affordable housing markets. CAHF's products provide wide access to a core set of metrics and data sources across the continent that are regularly updated and continuously improved. Its approach seeks to build rigour, replicability and scalability into its research and data gathering approach.

3. Guiding principles

A number of guiding **principles** frame the way CAHF engages with data, as well as with its various consultants and partners:

- **Drive replicability, efficiency and scale:** CAHF's data collection efforts may initially require significant human effort and direct contact between consultants or analysts and data owners. It is critical that all methodologies and data sources are replicable (e.g. data can be traced back to its source). In addition, CAHF should seek to shift as much as possible to automated data gathering methodologies.
 - **Replicability:** All sources must be documented with clear citations in publications to ensure that the same data can be collected in subsequent iterations of research. Research tools (including data gathering templates or questionnaires) should be curated to ensure that CAHF is able to replicate (or improve on) data gathering activities.
 - **Efficiency and scale:** In some cases, data collection processes that are manual (i.e. that rely on high levels of human intervention) could be automated. This can be achieved through partnerships with data aggregators who might be collecting but not publishing data. It may also be achieved through more innovative data collection approaches.

- **Innovation:** The world of data is being fundamentally reshaped by new technologies. It is critical that CAHF constantly identifies and explores or pilots alternative, emerging technology-driven data collection methodologies such as web-scraping and image processing. In addition, CAHF can play a critical underwriting role, actively supporting partners and consultants who wish to try new approaches to collecting data.
- **Emphasis on the local:** CAHF will support local sourcing of data and the growth of local attention to and engagement with data in support of affordable housing in the following ways:
 - Prioritisation of local and primary sources: Strive to obtain data as close to the source as possible.
 - Prioritising local capacity: CAHF actively invests in local capacity to generate and engage with high quality data and research outputs.
- **Adherence to best practises:** CAHF will define and champion best practise in its Data Agenda in the following ways:
 - Data management: CAHF will develop and promote data management standards for this Data Agenda, that include sound documentation, clear attribution, adherence to principles of database tidiness, and so on.² CAHF will circulate and actively advocate for the adoption of these standards by its partners.
 - Honesty and transparency: Be honest about limitations of the data and highlight any potential pitfalls to users.
 - Risk management: Ensure that no liabilities accrue to CAHF as a result of any data it might disseminate
 - Accessibility: In keeping with the best practises of Open Data, CAHF will strive to liberate data – encourage data owners to make data available for free.
- **Collaboration:** CAHF’s objective is to see that available data is made accessible to users who would find it valuable – not necessarily to provide it directly itself. CAHF will work with partners across the affordable housing sector to promote a practise of collaboration in achieving the Data Agenda. This will require:
 - Respect for privacy: Compliance with data privacy regulations and protection of the privacy of all research subjects.
 - Feedback and ongoing interaction with partners: Engagement with data owners/entities that generate data to ensure they generate data that is useful for the housing sector, *and* that showcases the importance of housing in the economy/household sector etc.

4. Basic approach

In line with CAHF’s data vision and mission, partnerships and capacity building are core to its approach.

4.1 Partnerships

CAHF has relationships with a number of partner organisations who share its vision for the housing and housing finance sectors. These partnerships include a number of the organisations who supply

² For other data management principles see *Draft Feasibility Study for the Development of a Data Tool*, 22 June 2018. Prepared for CAHF by Open Data Durban and LUMEC.

data, as listed in section 2.2. above. On the **data supply** side, the objective of these partnerships is to encourage and facilitate the provision of relevant, accurate and current data by these entities. In line with its mission, CAHF's objective is to see that this data is made available in an accessible format which promotes its data vision, to users who would find it valuable – not necessarily to provide it directly.

Supply-side data partners			
Category of partner	Examples	Their interest or relationship with data	CAHF's role or approach with these partners
Regulators	Prudential regulators (central bank), market conduct or consumer protection regulators, housing or planning authorities and municipalities	In some cases these regulators already collect useful data that they (or CAHF) could publish. In other cases, they could collect useful data that should be published.	Share insights on practises in other countries and to highlight the important role of the data they could / do collect.
Statistical agencies and other entities that conduct household surveys	National Bureaus of Statistics, World Bank	These entities collect data as part of their on-going activity. Data includes national censuses, health and demographic surveys, living standards surveys and so on	Encourage entities to add more / more relevant questions on housing and to disseminate data in an accessible way or allow CAHF to do so
Administrators	Deeds registries, Revenue collection agencies, Credit bureaus	These entities are not directly involved in the housing or housing finance value chains but generate useful data in the course of conducting administrative tasks.	These entities are often not aware of the potential value of their data as an input into decision-making in the housing and housing finance sectors.
Aggregators	Privately owned online platforms, including iBuild, rental market listing agencies, online estate agencies	Concerns about publishing data can based on privacy, or an intention (often vague) to monetise their data.	Explore opportunities to publish data in partnership (for example, CAHF could prepare and disseminate cobranded outputs)
Academic institutions	Various universities and business schools on the continent and beyond. Because academics are trusted to explore data impartially and without seeking economic advantage, they may be able to access data sets that would otherwise be difficult to obtain, typically from private sector providers who are reluctant to make data available but who might trust academics	Would like to access and analyse data and publish research on housing and housing finance markets.	Work with specific departments and academics on critical topics where leading edge technical expertise can be leveraged (including data science, imagery etc).
Specific privately-owned companies or industry bodies	These might include cement manufacturers, retailers, developers and financial institutions who might be	Often reluctant to share data. But might be willing to consider publishing or participating in industry-wide research	Leverage CAHF's role on the AUHF Identify specific indicators (e.g. cement prices) that

that play a critical role along the value chain	willing to share data on their specific activities		companies may wish to explore
--	--	--	-------------------------------

It can take many years to unlock a new data source. CAHF therefore needs to prioritise in which entities it invests in building relationships and trust.

CAHF also partners with entities that, like CAHF, would like to see the more data available. These entities seek data (**data demand**) and include entities that share CAHF’s vision for data and similarly prioritise data within their theories of change. These include donor-funded entities such as Reall or Habitat for Humanity, multilateral institutions such as the World Bank, investors, developers and other participants in the affordable housing sector, as well as policy makers and regulators. CAHF typically supports these entities, often managing the process of collecting data and key indicators across the continent. These partnerships can enable CAHF to participate in important conversations and leverage off relationships with globally recognised organisations to build recognition with other entities. In such cases, partnerships can strengthen CAHF’s agenda to liberate data that is collected but not yet disseminated on an on-going basis (see supply side partnerships above).

CAHF will establish a community of practitioners who recognise the importance of data in building the affordable housing sector, adhere to good data practise and can support CAHF’s data vision and mission in support of this Data Agenda for Africa.

4.2 Capacity building

CAHF will build the capacity of the partners it works with in a number of ways, including:

- The provision of training to in-country consultants who partner with CAHF on various data-gathering projects;
- Establishing and promoting best practise data standards that cross national boundaries;
- Establishing key data protocols for the collection of affordable housing-relevant data;
- Sharing of best practise; and
- The development of a CAHF ‘way of working’ that outlines a set of data norms and standards that is clearly communicated and included in all project plans and proposals.

In addition, CAHF will build the capacity of its internal team to deliver on core outputs that it wishes to maintain internally. CAHF will develop a set of guiding principles to determine which capabilities it should build internally and which it can rely on external partners to provide.

5. Data needs

A critical deliverable of the Data Agenda, as per objective 1, is to develop a set of core indicators as part of a coherent framework that recognises the differing needs of key categories of users of the data. Because of its extensive experience in the sector and its history of engagement with a wide range of players across the continent along the housing value chain, including private sector participants, public sector entities, policymakers as well as multilateral and donor-funded entities, CAHF is uniquely positioned to set the housing data agenda for Africa, identifying priority data sets and key indicators that need to be made available. At the same time CAHF obviously must have a good sense of the needs of its key audiences or partners on the demand side. These needs may not be well articulated – because users might not know what they don’t know – and CAHF may need to lead this, framing needs *for* its partners and not only responding to partners. In addition, CAHF must

have a good sense of the data that is or could be made available (the supply side) to close the key gaps that currently impact on market development. These are likely to differ by country.

Because there are many indicators that CAHF already publishes and uses in its work, CAHF will structure a comprehensive list of indicators in line with key frameworks that CAHF already uses. This approach will enable CAHF to identify key indicators that might be of value for its key audiences, and communicate clearly why indicators matter.

In general, key audiences include:

- Policymakers and regulators
- Private sector participants (including FSPs, investors and developers)
- Donors and funders

It is already clear that while policymakers and regulators (and donors) might be happy to work with more macro level data, FSP's, investors and developers might need to access more granular levels of detail. In some cases this data may not be useful for the sector as a whole, but would relate only to a particular opportunity. CAHF would need to assess whether in providing this data, it would be serving its broader agenda (for instance opening up a set of data, or assessing the feasibility of accessing a data set).

6. Conclusion

This initial draft of the data agenda prepared in March 2019 will be refined during the course of the year, with an update to be prepared in 2020. That update will reflect the implementation of specific interventions included in the plan for 2019. At the heart of this plan is a focus on consolidation, documentation and data management, providing a solid foundation for CAHF's future data work. This includes the implementation and communication of clear data protocols to ensure that data is collected, documented and maintained in line with sound data management principles by CAHF's internal team and its consultants. Existing databases are being cleaned in line with 'tidy data' principles³ while processes are being developed to streamline the collection and verification of data for the Yearbook and FSD indicators. The 2019 implementation plan also includes the clear articulation of CAHF's priority data sets and indicators in line with a principled framework. These in turn will enable CAHF to clearly communicate its housing and housing finance data agenda for the continent to an external audience, and position CAHF clearly as the thought leader in this regard.

³ See Tidy Data, by Hadley Wickham, The Journal of Statistical Software, vol. 59, 2014 available at <http://vita.had.co.nz/papers/tidy-data.html>