Insight2 Impact

Advancing Financial Inclusion

Inclusive insurance enhanced through the use of client data

Typologies, use cases and adoption challenges

October 2018

In partnership with

fsdafrica

UKaid
from the British people

Authors

Renée Hunter, Krista Nordin and Mia Thom

About insight2impact

Insight2impact | i2i is a resource centre that aims to catalyse the provision and use of data by private and public-sector actors to improve financial inclusion through evidence-based, data-driven policies and client-centric product design.

i2i is funded by the Bill & Melinda Gates Foundation in partnership with the MasterCard Foundation. i2i is jointly hosted by Cenfri and Finmark Trust.

For more information:

Visit our website at **www.i2ifacility.org**.
Email Mari-Lise du Preez (i2i's Partnerships Manager) at **mari-lise@i2ifacility.org**.
Call us on **+27 21 913 9510**.

About FSD Africa

Financial Sector Deepening (FSD) Africa is a non-profit company that aims to increase prosperity, create jobs and reduce poverty by bringing about transformation in financial markets in sub-Saharan Africa (SSA) and in the economies they serve. It provides know-how and capital to champions of change whose ideas, influence and actions will make finance more useful to African businesses and households. It is funded by UK aid from the UK Government. FSD Africa also provides technical and operational support to a family of financial market facilitators or financial sector deepening trusts (FSDs) across SSA called the FSD Network. FSD Africa has partnered with Cenfri to develop insurance markets in SSA.

For more information:

FSD Africa, Nairobi, Kenya
info@fsdafrica.org
@FSDAfrica
www.fsdafrica.org

Contents

Acknowledgements

The authors of this note would like to thank all organisations and experts consulted in the research process for their valuable input, review and comments.

Methodology

This note is based on desktop research and telephonic interviews with 15 inclusive insurance providers and six experts in the field.

Definitions. In this note, *client data* is defined as data that provides insight into an individual client or the characteristics of a segment of clients, who they are, what they need and how they behave. *Inclusive insurance* is defined here as insurance provided in emerging markets that aims to go beyond wealthy consumer groups. A *client-data use case* is understood as a distinct observation of an example of *client data* application to a specific insurance activity.

Desktop research. Desktop research was done to provide an overview of the different types of client data available. This produced the list of client-data types to consider in this note, shaped the terminology used and informed the selection of providers for the interviews.

Interviews. Interviewees were selected to cover a range of value chain actors, geographies, product types and specialist perspectives. Insurers, insurtechs and data solution providers were interviewed. In terms of geography, the aim was a strong representation of developing countries. In terms of organisations, the aim was a balanced representation of innovative insurtechs, more traditional insurers, micro-insurers, and data solution providers focusing on insurance. A limited number of innovative insurtechs from developed markets were included where their solutions were applicable to the developing world. Finally, interviewees were selected to provide a balance between life, non-life, health and agricultural insurance. This sample was selected to allow the research process to identify a broad range of client-data use cases in insurance that are applicable to the developing world. See Appendix 1 for an overview of the institutions interviewed.

Interviews focused on the client-data types used by organisations, the use cases and rationale for use and the challenges faced in implementation.

Limitations of the study. This note draws on the views and experiences of a sample of providers. Efforts were made to select a wide range of providers to identify key trends in innovation with client data, but they may not be representative of the market as a whole.

Acronyms used

AI	artificial intelligence
API	application programming interface
ARPU	average revenue per user
CIC	Cooperative Insurance Company Kenya Limited
CDR	call detail records
CRM	customer relationship management (database)
FSP	financial service provider
GIS	geographic information system
IoT	internet of things
IPRS	Integrated Population Registration System
KYC	know-your-customer
MFI	microfinance institution
MM	mobile money
MNO	mobile network operator
MSMEs	micro-, small and medium-sized enterprises
PIN	personal identification number

1 Introduction

Insurance has always been a data-driven industry, with quantitative assessments forming the backbone of the service. The explosion of big data in recent years, combined with new technologies, enables insurers to collect and analyse these vast new datasets. These advancements open up possibilities for insurance providers to save time and costs, reach new markets and explore new business models (Breading, 2017).

In developing markets, insurers face a number of challenges that limit their ability to reach consumers with products that effectively meet the consumers' needs. In this note, we use the term "inclusive insurance" to denote insurance provided in emerging markets that aims to go beyond wealthy consumer groups. The challenges for these providers have been organised into five broad categories by Smit et al. (2017): (i) A lack of information on consumers; (ii) Consumers being beyond the current reach of distribution channels; (iii) Consumers having new and different needs from those of traditional insurance clients; (iv) Consumers being inexperienced with formal financial services; and (v) Constrained business models (see Figure 1). These challenges all hinge on the dual themes that insurance fails to provide low-income consumers with enough value, and that the business case in the low-income market is challenging.

The effective use of client data allows financial service providers (FSPs) to (i) understand individual customer behaviour and needs, and (ii) do so viably at scale (Maeder, 2017). CGAP (2014) estimated that the effective use of client data by inclusive insurers led to efficiency gains for insurers, which reduced operating costs by between 15% and 30%. The potential impact for business performance is real and relevant.

Figure 1: Inclusive insurance delivery challenges

Source: Authors' own, adapted from Smit et al. (2017)

However, many insurers are not yet widely using client data in their business and, as a result, are not realising the potential business and client value gains. Research to date has focused mostly on the potential applications of data in insurance across insurance activities – see, among others, Arnoldussen & Hauner (2016), Breading (2017), CGAP (2014), Chitta et al. (2017) and Maeder (2017). This note adds depth to the discussion by focusing on data use cases for specific insurance activities, and the opportunities and challenges therein. Box 1 on the next page provides a brief description of relevant client-data types.

Box 1: Description of different client-data types

Biometric identifiers	Biometric data refers to data on biological characteristics that are unique to an individual. There are several types of biometric identification (including facial structure, fingerprint, hand geometry, retina, iris, signature, pitch and cadence).[1]
Call detail records (CDR)	These are records that are generated for every user event on a mobile network. They document the details of the communication transaction, whether it be a phone call, text message, etc. They capture a broad spectrum of information, including unique caller ID, origin of the call, calling number, receiving number, start time, end time, duration, call location and call type.
Complaints and call centre data	This comprises the recording of interactions with call centre agents and other client feedback mechanisms that the FSP may have in place. This data includes, at least, content of the complaint and the timestamp. Depending on internal data structure, this information might be included in, or linked to, an FSP's customer relationship management (CRM) database.
Credit bureau data	Credit bureau data is purchased from an external provider (the bureau) that collects and synthesises information on individuals' financial histories. This may include (re)payment records of various loans, as well as regularity of bill payments – depending on the bureau. The synthesised history is summarised in a single credit score, implying an assessment of the individual's willingness and ability to repay, which is what FSPs receive from the bureau.
Credit history data	This comprises a record of repayment behaviour, either from the FSP doing the analysis or shared by external credit providers. Variables include principal amount, original payment plan, days outstanding and amount outstanding.
Customer interviews	This is a type of qualitative market research that an FSP may conduct. Such interviews might take place one-on-one or in focus groups, and they can be conducted through a variety of channels (including face-to-face, telephonically or digitally). Topics and variables covered can differ to suit an FSP's specific purpose. A common type of interview is the customer satisfaction interview, conducted by customer service departments.
Geographic Information System (GIS) data	A GIS variable consists of two components: spatial data and attribute data. The spatial data is a geographical location, whereas the attribute data describes the "what", "when" or "how" connected to that geographical location.[2]
Government/ open data	This could be any data that is made publicly available by government and could include data on demographics, health trends, market performance, weather data, etc.
Information provided at sign-up	The information collected here will vary per FSP but includes, at least, know-your-customer (KYC) information, which is used to verify the identity of clients. This will be supplemented with any information the FSP deems necessary for the adequate servicing of the client, such as information on assets owned, occupation and family structure. Depending on internal structure, some or all of this information might be captured in an FSP's CRM database.
Image (meta-) data	The data contained in an image file. This might include meta-data such as timestamp and geographic location, as well as the image pixels itself.
Mobile money (MM) data	This comprises recorded financial transactions through mobile money. See "transactional data" below.

1 See for more information Cooper et al. (2018)

2 See for more information http://access.i2ifacility.org/Alternative_data_sources/

Psychometrics[3]	Psychometric data captures an individual's attitudes, skills, beliefs, intelligence, ethics, honesty, personality and reactions based on a set of circumstances. The data is captured at the time of the insurance application through questions on a survey and is used to measure characteristics such as confidence, outlook, conscientiousness, autonomy, opportunism, numerical reasoning skills and honesty. Examples of data collected are how long it took the applicant to answer the question, response time variance, whether they changed their answer and consistency, among similar questions.
Satellite/aerial imagery data	A satellite image is an image of, for example, the whole earth or part of the earth, taken from a satellite. Imaging satellites are operated by governments and businesses around the world. Images are then licensed to other governments and businesses, such as Google Maps and Apple Maps. There are three main types: visible light imagery, infrared imagery and water vapour imagery. Examples of physical or environmental variables that can be extracted from satellite imagery are: evapotranspiration, rainfall, soil moisture, vegetation and landscape characteristics. Aerial imagery is similar but is captured by cameras other than satellites (by drones, for example). Aerial imagery therefore generally covers a smaller area than satellite imagery but can be of higher quality.
Sensor/Internet of Things (IoT) data	The IoT is a network of physical devices embedded within electronics, software, sensors, actuators and network connectivity that enable these objects to collect and exchange data. The data collected is the output of the device, in response to some type of input from the physical environment. There is a variety of types of sensors, such as accelerometers, thermometers, pressure meters, photoelectric meters and smart grid sensors. Sensors are often used specifically in the health space, and they record data such as movement patterns, calories burned, heartbeat, temperature, speed, and so forth.
Smartphone data	This kind of data is obtained from mobile applications. It includes data that a user inputs into applications, data generated through the user interaction with the application and data that the application has permission to extract from the smartphone. Examples of the latter include personal, identifiable information, geographic location, battery life, contact lists, financial notifications via SMS and, depending on the applications connected, different types of social media data.
Social media data	This comprises all forms of data collected on social media. Attributes include posts, views, likes, shares, follows, unfollows, impressions, clicks, content of messages, sentiment of messages, size of network and connections between individuals.
Survey data	In this report, survey data refers specifically to nationally representative survey data. This is collected by third parties and is made available either through open access or as a paid resource and provides individual anonymised information on individuals in the market on a variety of topics. Examples include various financial inclusion surveys that provide insights into how individuals use financial services and what their needs are, as well as household expenditure surveys, labour force surveys and surveys on healthcare usage.
Transactional data	This data comprises recorded financial transactions generated by the FSP and in some cases shared externally. Variables recorded include unique identifier, timestamp, amount, sending account and receiving account.

Source: Authors' own and Nordin (2016)

3 Psychometric data has been applied successfully in the low-income market in credit provision. The authors had expected to observe similar examples in the inclusive insurance space at the outset of the project, but none of the interviews conducted revealed any such examples

2 Client data and its role in inclusive insurance

Although a significant amount of new data is emerging, this does not mean that it is as yet possible to apply it in a way that makes business sense, particularly in challenging inclusive insurance markets. To start off with, there is a lack of clarity around the availability of different data types, their potential use cases, and the costs and benefits involved in them. These questions were explored during interviews, and this section discusses different client-data types and the ways in which inclusive insurance providers are using these data types for key business challenges.

2.1 What is client data?

Data is an overarching term that can be used to describe any fact or statistic collected for reference or analysis (Nordin, 2016). In this note, client data is understood as data that provides insight into an individual client or the characteristics of a segment of clients: who they are, what they need and how they behave. This is a deliberately broad definition and may include a client's identity, location, movements, assets, needs, wants, history or behavioural profile. There is a variety of emerging and traditional client-data sources that offer potential to address the challenges faced by inclusive insurers.

2.2 Client-data types

Vast amounts of data are generated on an ongoing basis by individuals' actions through a variety of channels. Each of these data types consists of different variables, collected in a different way. Some are automated, some manual, some internal to the FSP, some purchased externally, some anonymised and some highly individual. Box 1, on the previous page, provides a brief description of the client-data types relevant in inclusive insurance, as observed in this research.

"Vast amounts of data are generated on an ongoing basis by individuals' actions through a variety of channels. Each of these data types consists of different variables, collected in a different way."

Chatbots not considered a data type, but significant for data collection and action. Whereas chatbots are not an example of a data type, they are relevant to inclusive insurance due to their ability to automate the handling and actioning of various data types. Chatbots can be used to automate the collection of data and to apply artificial intelligence (AI) to the immediate analysis of the provided data. This method is most useful for data that is generally collected in a conversational manner, such as information provided upon sign-up and complaints and call centre data.

Unstructured data included across data types. Unstructured data is not a client-data type, but rather a format that can apply to different client types as listed above (for example, social media data or call centre records). Through emerging technology, this data format is increasingly accessible for analysis.

2.2 Client-data use cases across the insurance value chain

The current and potential use cases of client-data types in inclusive insurance are broad and the business case for each is evolving. We've used the insurance value chain as a framework to better understand how different types of client data are used to address business challenges (see Figure 2).

Figure 2: Insurance value chain

Source: Smit et al., 2017, informed by ILO Impact Insurance Facility (Merry et al., 2014)

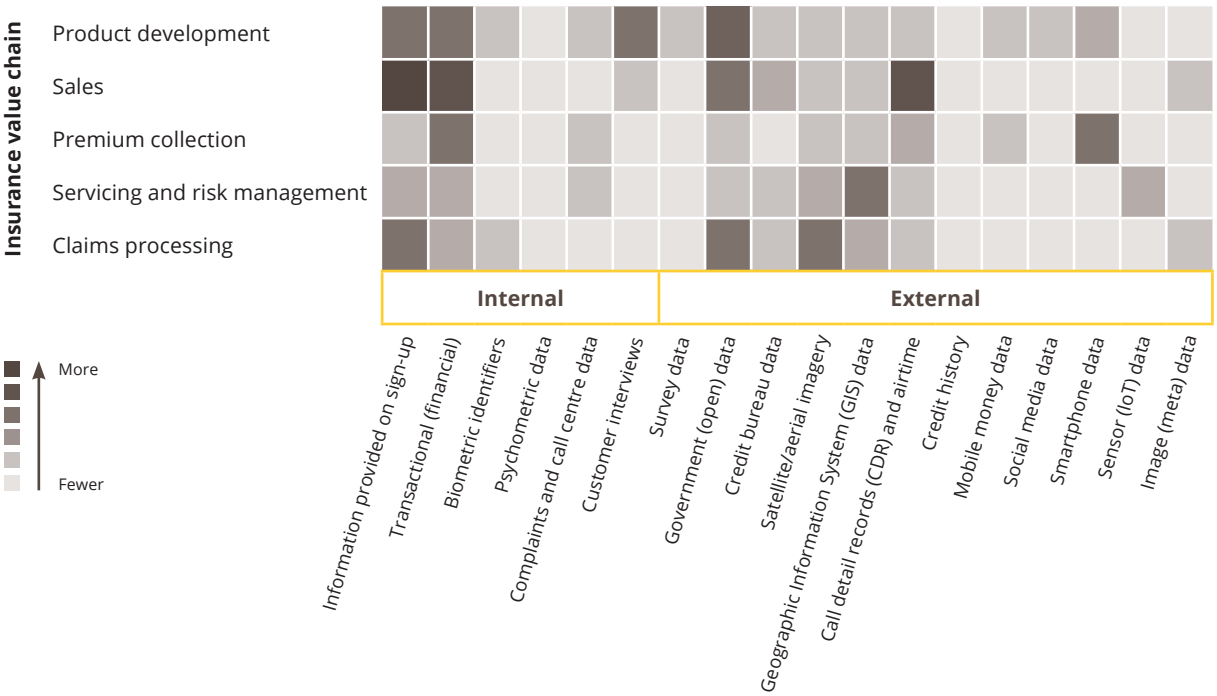
Box 2: Recapping the insurance value chain

The insurance value chain, as set out in Figure 2 above, groups all activities undertaken by insurance providers in serving their customers. The first step in the value chain is **product development**, which involves the process of researching, designing and pricing the product. The second step is **sales**, which is the process of reaching consumers and extending the product to them. The third step is **premium collection**, which involves the systems and mechanisms that are in place to facilitate the payment of insurance premiums, including any nudges or other mechanisms to ensure timely payment. The fourth step, **servicing and risk management**, is done at the back-end of the insurance provider. Servicing refers to the processing of an insurance policy and all communications around it, e.g. monitoring the payments, sending out notifications, verifying information and handling complaints. Risk management is primarily a back-end activity, but it is increasingly brought to the consumer-facing side of the business, in providing the consumer with information to allow them to engage in risk-decreasing activities. Finally, the fifth step is **claims processing**, which relates to all activities that relate to the processing of an insurance claim, including the lodging of the claim, verification and pay-out. This value chain is a closed loop, as all steps ideally inform the product development phase.

The 15 insurance providers interviewed listed a total of 93 client-data use cases across the insurance value chain. Figure 3, below, maps the prevalence of these client-data use cases against the insurance value chain activities. The various steps in the insurance value chain are plotted against the Y-axis, with the different client-data types plotted on the X-axis. The cells represent the number of examples observed of a client-data type addressing a challenge within a step in the insurance value chain. The darker the colour, the more use cases have been observed for that specific data type for the specific activity.

In this discussion, we understand a client-data use case as a distinct observation of an example of client-data application to a specific insurance activity. In some instances, use cases are integrated with other use cases to form a broader insurance application. For example, Discovery's Vitality[4] solution makes use of information provided upon sign-up, transactional data and sensor (IoT) data, which feed into the processes of product development, sales, servicing and risk management and claims processing. For the purposes of identifying trends across client-data types as well as insurance activities, each application of a data type to a step in the value chain (each cell in Figure 3) is referred to as a use case.

Figure 3: Client-data use cases per insurance value chain



Source: Authors' own

4 https://www.discovery.co.za/vitality/how-vitality-works

Product development. Insurers reported difficulties in identifying target customers and understanding these customers' needs during product development. The largest number of distinct client-data use cases (as illustrated in Figure 3) were observed in the product development stage of the insurance value chain, most of them making use of internal data that insurance providers are likely to collect themselves. Examples of how client data is already being used to address some of these challenges are the following:

- Agricultural insurers (such as **Pula**[5] or **Acre**[6]) use GIS and satellite data to identify the crops farmed by customers and understand the associated risks they are exposed to.
- In another example, the **Vanguard Group**[7] in Ghana recently conducted a broad social media analysis to identify insurance needs among their target market that are not met by existing product features.
- **Digital Fineprint**[8] uses social media data to identify owners of MSMEs and the types of risks they might be exposed to. This allows the insurer to group potential clients with similar risk profiles together and approach them with appropriate products.
- In an example of a more qualitative approach, **Hollard**[9] makes extensive use of self-conducted customer interviews and focus groups, as well as nationally representative survey data to identify needs and inform their product development.
- Mobile insurance providers, through their partnership with mobile network operators (MNOs) can use average revenue per user (ARPU) as a proxy for affordability in designing products for developing markets. **BIMA** is an example here.

Sales. Within sales, the challenges identified by insurers were: limited consumer awareness, challenges in optimising their sales channels (including agent networks), identifying appropriate physical marketing locations, onboarding clients cost-effectively, ensuring adequate information be collected on the customer, and customer education. A range of different client-data types that support a more efficient and effective sales process were reported. These included solutions that (i) automated the collection of data required for client onboarding, thus shortening the sales process; (ii) enabled easier and more reliable verification of client data; (iii) targeted sales efforts towards those customers most likely to purchase the product; and (iv) tailored automated communication strategies to encourage product take-up.

Whereas the greatest number of *distinct use cases* were identified within product development, we observed the greatest number of *insurance providers* applying the identified use cases within their sales process.

5 https://www.pula-advisors.com/
6 https://acreafrica.com/
7 http://www.vanguardassurance.com/
8 https://digitalfineprint.com/
9 http://www.hollard.co.za

Some examples include:

- **The CIC Insurance Group**[10] has linked up with the Kenyan Integrated Population Registration System (IPRS) to facilitate rapid cross-referencing of individuals' sign-up data. This has greatly reduced the time involved in completing a sign-up process and provides the insurer with greater confidence in the identification of clients, which means that fewer clients are rejected due to identification issues.[11]
- **BIMA**[12] makes use of MNO data to identify clients with a specific ARPU. Based on ARPU, affordability levels of clients can be identified, and clients can be directly targeted with the appropriate product for their income levels, increasing the efficacy of the sales process.[13]
- **Inclusivity Solutions**[14] use data on prospective customer engagement to drive their marketing and sales content in tracking uptake data. Following experiments with different marketing text messages, they were able to identify which wording leads up to a 65% larger take-up of mobile insurance. This data-driven approach has proven to be a more cost-effective and accurate way of optimising marketing than, for example, customer surveys and interviews (Tappendorf, 2018).
- **ToGarantido**[15] utilises chatbots to automate the collection and processing of data required for the sales and onboarding process. AI is applied to the sign-up data provided by clients, allowing for immediate risk assessment, identity verification and insurance activation.[16]

Premium collection. The main challenge for insurers in collecting premiums is high lapse rates, often due to ineffective premium collection mechanisms, payment system limitations and behavioural challenges to move from the intent to pay to payment. Client data can be used to time the collection of premiums (or the timing of payment reminders to clients) to maximise the likelihood that policyholders will be sufficiently liquid to pay when collection takes place. The use of client data remains relatively limited in this stage of the value chain: Among our interviewees, only three insurance providers reported applying client data here. Some high-potential examples do emerge:

- Based on the GIS locations of insured farmers, agricultural insurers (such as **Pula**) are able to target farmers for premium payment per area. The timing of premium collection can be adjusted for harvest periods of specific crops that are prevalent in that area.
- MNO data on airtime top-up behaviour can be used to estimate at which point in the month clients will have a certain balance in their accounts. This allows mobile insurance providers (such as **BIMA** or **MicroEnsure**[17]) to adjust their premium collection dates to when clients are likely to have enough money in their accounts or mobile wallets.

10 https://cic.co.ke/
11 See case study in Annex.
12 http://BIMAmobile.com/our-services/
13 See case study in Annex.
14 http://inclusivitysolutions.com/
15 https://togarantido.com.br
16 See case study in Annex.
17 http://www.microensure.com

Servicing and risk management. Insurers reported challenges in educating customers on the features of their insurance policy and their risk behaviour. A number of examples emerged of how client data can be used to better facilitate this stage of the insurance value chain. The most promising examples relate to risk management, involving continuous risk-assessment by the insurance provider in the back-end, to allow for optimal forecasting. Even more promising and disruptive are the examples that feed information back to the consumers, encouraging and enabling them to make informed decisions to change their behaviour. In reducing clients' risky behaviour, these client-data use cases provide value for both the consumer and insurance provider, as the insurance provider increasingly aids in preventing adverse events happening in the first place. Interestingly, out of all five steps in the insurance value chain, this one showed the *greatest diversity* in client-data use cases – with a sizeable number of insurance providers each applying distinctly different use cases.

- **Lumkani**[18] makes use of a network of IoT heat detectors to measure fire risks in informal settlements over time. This data provides insurers with real-time risk assessment of the insured areas, and the linked fire alarms allow for faster action in the case that a fire does occur.
- **Aerobotics**[19] makes use of satellite and drone imagery to provide its clients with precision farming dashboards. These dashboards provide insights from the satellite and drone data on various farming metrics, such as moisture, pesticides and vegetation growth. This allows for continuous risk monitoring of the covered farms and timely preventive action to be taken.[20]
- **AgriGO**[21] is an example of how farmer activity data can be successfully used to inform farming advice notifications. Based on key inputs such as crops and location of the farm, basic nudges about the timing of activities (such as planning and weeding) are sent to farmers via SMS. These notifications are responsive to updates that farmers provide on their activities via USSD (Hunter, 2018a).[22]
- **Discovery Vitality**[23] health insurance makes use of a wide variety of alternative client-data types to encourage less risky behaviour from insured clients. Clients receive Vitality points for activities, such as buying healthy foods (captured from clients' transaction data), gym visits (captured through partnerships with gym networks), being physically active (captured through IoT devices, such as Fitbits) and various health check-ups (captured through clients' health records). This gamification approach applied in providing points and rewards reduces health risks for insured clients over time, and it has resulted in health costs decreasing by 25% for those involved in the programme (Omarjee, 2017).

18 https://lumkani.com/

19 https://www.aerobotics.io/

20 At the time of writing, Aerobotics offered the dashboards as a precision farming service and was in the process of launching their services for the insurance industry.

21 https://agrigo.rw

22 At the time of writing, AgriGO was employing its nudges towards agricultural credit provision, and was actively seeking to expand its services to include insurance. It is a promising example of nudging towards less risky behaviour specifically in the low-income market.

23 https://www.discovery.co.za/vitality/how-vitality-works

Claims processing. Efficiency and transparency in claims assessment, including fraud detection, swift payment and identity verification are key insurance challenges. The use cases identified are limited, but some providers reported the use of internal client data to automate their claims process. Examples of more innovative applications of client data are:

- Through its precision farming dashboard, **Aerobotics** is able to easily access before- and after-images of the insured farm. This allows the insurance provider to cheaply assess whether a risk event has indeed occurred without needing to send an individual to the farm and affected field.
- Through their Insurance Data System (IDS), **TransUnion**[24] provides an overview of claims data across insurance providers. This enables insurance providers to conduct a check for claims that have been submitted, cross-referencing with any other recent claims. This significantly reduces time and cost involved in identifying claims that are likely to be fraudulent, which is especially relevant for smaller-value claims.
- **CoverApp**[25] uses the metadata of digital images – in this case the timestamp and geolocation code – to detect possible fraud. Images provided at sign-up and during the claim are cross-referenced for this timestamp and location data. This approach results in a high level of certainty for the insurer and greater convenience for the client.[26]
- **Inclusive Financial Technologies**[27] provides a variety of identification tools to FSPs, including insurance providers. Its flagship identity verification API has reduced the time required to confidently confirm an individual's identity to 15 seconds. Its image-based selfie recognition tool, which compares an individual's selfie[28] with the picture on their ID, likewise brings down the time and cost involved in identity verification for claims processing (and sales) (Hunter, 2018b).

Across the value chain, early emphasis is on the use of internal client data. Regardless of where in the insurance value chain, most of the current use cases start with data that is internal to insurance providers – mostly sign-up and transactional data (14 out of 93 use cases each). This is reflected in priorities as relayed by insurance providers: Most indicate focusing on maximising value derived from data that is already present within the organisation, before focusing on improving customer experience and cost-savings further down the line. Moreover, for FSPs at the early stages of their data journey, it might be easier to access internal data, rather than engage in data sharing or purchase agreements with external providers (De Villiers & Chetty, 2018).

24 https://www.transunion.co.za
25 http://coverapp.co.ke/
26 See case study in Annex.
27 https://inclusiveft.com
28 A photograph that one has taken of oneself, typically with a smartphone.

3 Barriers to using client data in inclusive insurance

Although all providers interviewed find the use of client data an imperative to improve insurance value propositions and business models, they also reflected on the significant difficulties they faced in adopting existing use cases or in moving towards new opportunities.

The interviews identified five main categories of difficulties, and they were confirmed by literature:

1. Strategic priorities and organisational changes
2. Uncertain business case
3. Poor data quality and availability
4. Partnership constraints
5. Regulation

Strategic priorities focused on incremental change. Interviews showed that client data is a key strategic priority for some providers and less significant for others. More established institutions in particular reflected on the complexity of changing culture and legacy systems to harness new types of client data. Providers interviewed typically started by cleaning up internal data and using external data for distinct use cases, rather than with a more advanced and comprehensive data and governance strategy. This allows for a staged integration of client data into decisions and reduces the upfront need for comprehensive governance, resourcing and systems change. Examples cited by providers included updating systems to obtain a 360-degree view of the client across legacy systems (especially after multiple mergers), harnessing customer complaints data for product development and the use of public datasets on motor vehicle registration to reduce fraud.

Uncertain client-data business case hampering adoption. Due to the nature of inclusive insurance and low margins driven by low premiums (coupled with the requirements for data-driven decision-making, high upfront investment costs and trained experts), the business case is unclear for firms to leverage client data to overcome some of their business challenges. Additionally, several providers and experts described how experiences with failed or costly data initiatives may form, and have formed, a deterrent for experimenting with future client-data use.

Poor data quality and low availability of client data within firms. The difficulty most frequently mentioned by interviewed providers that impedes the adoption and use of client data is the availability of high-quality client data available to firms. Poor data quality and low availability of client data in insurance providers stem from manually captured customer information (which tends to be incomplete or inaccurate) and/or data being stored in silos on legacy systems (which is difficult for providers to merge to form a holistic picture of individual consumers). Significant investment is required to turn such data into useful information for decisions.

> “The difficulty most frequently mentioned by interviewed providers that impedes the adoption and use of client data is the availability of high-quality client data available to firms.”

Partnerships showing potential of data sharing, which is difficult in practice. A significant number of interviewees acknowledged the need to partner to access additional data outside of their own internal data. Those interviewed providers that *have* partnered reflected on the difficulties involved in setting up and maintaining such partnerships and the limitations to access client data. Even when partnerships are structured in such a way that they are beneficial for all parties involved, there can be reluctance from potential partners to enter into client-data-sharing agreements. Reasons given by those providers that have entered into partnership agreements include: fear of giving up competitive advantage and, significantly, uncertainty around repercussions from regulators. Once a partnership is created, further difficulties emerge. The insurance provider can lose direct access to its clients,[29] and it can become difficult to merge databases between the parties to create a single view of the customer. However, for most, the benefits or partnership still outweighed the challenges.

Use inhibited by onerous or uncertain regulation around data-sharing. For some providers, onerous regulation has inhibited data-sharing arrangements, either between partners domestically or within providers across borders. However, providers highlighted that *uncertain* regulation can be a more severe constraint to innovation with client data, as it's not clear whether it will ultimately be possible to realise investments in data sharing or client data.

29 Typical partnerships observed are between mobile network operators and insurance providers. Once they partner, the insurance provider loses direct access to its clients, as all channels to its clients are intermediated by the MNO.

4 Conclusion

In conclusion, a range of client-data types are emerging with use cases across the insurance value chain. In emerging economies, the main applications of client data that have been observed are in product development and sales, where it is used to solve difficulties in affordably identifying and reaching clients. However, emerging use cases can be observed in the tailoring of premium collection, risk management and claims processes, and more insurance providers will likely move in that direction once certain data challenges have been dealt with.

Insurance providers able to experience big wins in getting their internal data in order first. Many providers have realised the strategic importance of capitalising on existing client data collected in their business. Approximately one-third of all client-data use cases observed through our interviews are related to such endeavours. For example, a significant opportunity in the insurance space is for providers to centralise their data systems and integrate all silos so they can achieve a single customer view. While this might not sound as exciting as applying advanced analytics to a variety of client-data types for detailed client profiles, it is an essential first step in making the data assets that do currently exist usable for insights. The fact that more providers are focusing on getting their internal data in order is promising, and this might mean that we will observe a large number of new applications in the near future, once the foundational work has been completed (De Villiers & Chetty, 2018).

Organisational factors as important as technical considerations. Many insurance providers indicated that organisational factors such as buy-in, reporting structures and skills can be just as prohibitive to successfully applying client-data uses as technical considerations are. A strong data culture often requires a cultural shift, as well as across-the-board buy-in and commitment to the investments required.

Many innovative use cases, but the business case remaining unclear. The application of many client-data types is still in an experimental stage. The most common use cases were in client data gathered internally or through partnerships with MNOs related to airtime spending patterns, or in some cases bank partnerships. A number of other client-data types have been tested with some success, but many still struggle to deliver a compelling business case or require organisational changes that are difficult to accommodate and require a high degree of risk taking. Many such models are also still limited to better serving existing consumer groups, rather than to penetrate further into the unserved market. Significant opportunities remain both in cost-saving and time-saving – invaluable in the large-scale, low-margin space of inclusive insurance – as well as in reaching new customer markets that were previously "invisible" to traditional data. More work needs to be done to understand the cost and benefits involved in these applications and to deepen capacity on how to extract value from adopting new client-data models.

Need for clear data regulation that protects consumers while enabling businesses to invest to deliver more value. Data protection and data privacy regulation is undergoing significant changes globally. The related regulatory uncertainty and, at times, onerous compliance requirements affect the incentives to invest in client data and data-sharing arrangements to deliver value to consumers. Transparent and defined regulation that allows for responsible innovation with client data will be significant to unlock the full potential of data to better serve and include consumers.

Appendix 1:

Interviewees

Providers:

- **Microinsurance providers**

  - Acre Africa
  - BIMA
  - MicroEnsure
  - Pula Advisors
  - Pioneer Philippines

- **Traditional insurers**

  - CIC Insurance Group Ltd – Kenya
  - Discovery
  - Hollard Insurance

- **Insurtech**

  - CoverApp
  - Digital Fineprint
  - Lumkani
  - ToGarantido

- **Data solution providers**

  - Aerobotics
  - Cignifi
  - TransUnion

Experts:

- Accion Venture Labs
- Bowman's
- Centre for the Economic Analysis of Risk (CEAR), Georgia State University
- ILO Impact Insurance Facility
- Microinsurance Network (MiN)
- National Insurance Commission (NIC) Ghana
- University of Pretoria Natural Hazard Centre

Appendix 2:

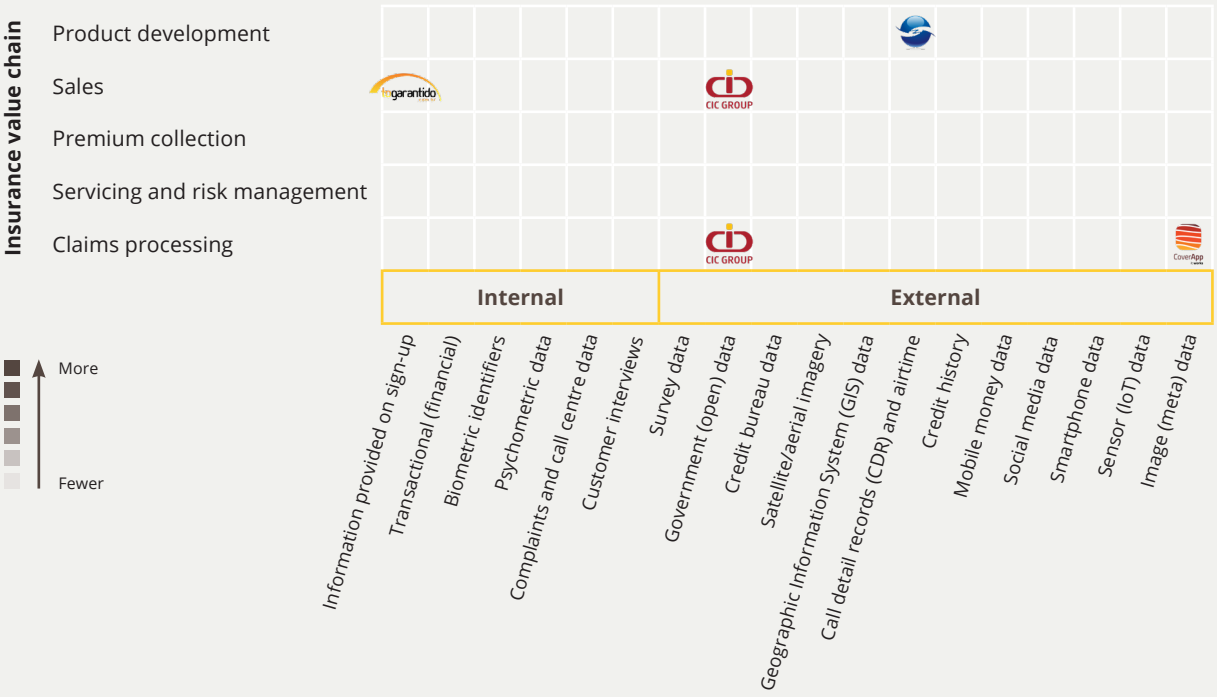
Client-data case studies

This appendix discusses four illustrative case studies from our research. These case studies illustrate how certain insurance companies are leveraging client data to overcome the business challenges that they are facing. They were selected to provide coverage of the breadth of business challenges across the value chain and are therefore not proportionally representative of the most applied use cases identified through this research.

All applications described in the case studies are either currently being applied in the low-income space or show promising application.

The four companies selected are BIMA, CIC, CoverApp and ToGarantido. Figure 4, below, illustrates which types of client data are used in each example, and at which stage of the insurance value chain.

Figure 4: Client-data use cases per insurance value chain activity



Source: Authors' own

BIMA[30]

About the company

Introduction	BIMA is a mobile microinsurance and mHealth[31] provider, launched in 2010, with headquarters in Stockholm and London.
Number of customers and target market	26 million customers reached, 75% of whom access insurance for the first time. 93% live on less than USD10 per day, 54% live on less than USD2.50 per day, 90% are self-employed.
Where they operate (geography)	Ghana, Senegal, Tanzania, Bangladesh, Cambodia, Fiji, Haiti, Indonesia, Pakistan, Papua New Guinea, Philippines, Sri Lanka, Honduras, Paraguay and Nicaragua.
Product and distribution channels	• *Products:* Health, life and personal accident insurance and mHealth (including tele-doctor consultations, prescription and delivery, preventative health tips and health coaching programmes, discounts at hospitals and pharmacies, etc.) • *Distribution:* MNOs and other strategic partnerships, call centres and field agents
Overall client-data use	BIMA uses a variety of client-data sources and market research approaches at each step of the insurance value chain. Examples of data used by the company are: • Qualitative customer research data, claims data and test-and-learn data for product development • MNO data for both sales and premium pricing • Transactional data for premium collection • Tele-doctor engagement data for servicing and risk management

30 https://www.bimamobile.com

31 mHealth (or mobile health) refers to different ways of delivering medical services through the use of mobile technologies.

About the use case

Specific business challenge

This case study specifically focuses on BIMA's use of MNO data for product development. The specific challenge in product development that they are solving, through leveraging this data source, is how to best price their products to remain affordable for their target customers.

Client data used to solve the challenge

BIMA leverages client data from their partner MNOs. The information they receive from their partner MNOs includes: customer identifier, phone number, average revenue per user (ARPU) and in some instances top-up behaviour.[32] Upon entering a new market, BIMA partners with a local MNO and receives the call list with relevant metrics per individual. These metrics are analysed to inform initial affordability levels in the market. Given the relatively simple nature of the insurance product, it can be launched almost immediately (pending regulatory approvals). In certain cases, iterations of the product may be tested through a test-and-learn approach, or the insights may be supplemented by qualitative market research, focus groups or claims data.

Advantages and disadvantages of this approach

The advantages of this data type are that it is already collected within the partnership, it is available for large groups of (potential) customers at once, one can derive granular insights from the individual-level data, and it includes contact details.

The disadvantages of this data type are the fact that ARPU is only a proxy for income, and the assumptions underpinning this proxy may not hold across markets; and ARPU, as a metric, masks certain behaviour. It cannot, for example, distinguish between large ad hoc top-ups or regular low-value top-ups, thereby only showing part of the story. Moreover, many individuals have multiple SIM cards, which makes ARPU data misleading on its own. Finally, the amount and nature of the data shared depend heavily on the nature of the individual partnerships.

Why this approach?

In the past, BIMA has attempted partnerships with other organisations aside from MNOs (such as MFIs or solar providers). However, MNOs have been proven to be the best strategic partner, due to their large customer base, relatively higher level of technology (allowing for easier integration) and the level of trust that they enjoy from their customers.

Mobile financial services have a high rate of customer activity, with approximately 50% of customers considered active. This means that BIMA has a higher chance of signing them on in any given month, making this data-driven sales channel highly effective.

Future plans

BIMA intends to create platforms that allow for more day-to-day digital direct engagements with clients in order to build fuller client profiles and understand client risk. The BIMA Doctor App will be expanded to make it relevant for clients daily and to facilitate collecting data on pre-risk events and health trends. Its plan is to use this data across the insurance value chain, informing product design, sales, pricing and servicing.

32 The amount and nature of the data they receive are highly dependent on the partnership with the MNO.

CIC Insurance Group Ltd – Kenya[33]

About the company

Introduction	The CIC Insurance Group Limited is a regional insurance provider, founded in 1968. CIC has three subsidiaries (providing life, general insurance and asset management).
Number of customers and target market	1.3 million clients. Target market ranges from the upper middle class to underserved customers.
Where they operate (geography)	Kenya, Uganda, South Sudan and Malawi
Product and distribution channels	• *Products:* Over 50 different products in agriculture, health, motor and life insurance, and asset management • *Distribution:* Channels range from (micro-)agents and brokers to SACCOs and partnerships with other companies, such as MNOs and petrol stations.
Overall client-data use	CIC mostly uses client data for product development. Examples of data used are: • Demographics • Products held per clients • (Limited) complaints data • Government database of population information

About the use case

Specific business challenge

This case study focuses on CIC's use of the Kenyan national ID database (IPRS) for sales and claims processing. The specific challenge in sales and claims processing that it is solving is the lack of verifiable personal information, which adds considerable time to the sign-on process and compromises fraud detection.

33 https://www.cic.co.ke

Client data used to solve the challenge
CIC makes use of the Kenyan Integrated Population Registration System (IPRS). Launched in 2015, this database connects various government departments and pools all personal data that is centrally held about an individual. Data currently stored in the IPRS includes:

- Births
- Deaths
- Registration of refugees and aliens
- Driver's licences
- National Social Security Fund
- National Hospital Insurance Fund
- Voter registration
- SIM cards
- Bank accounts

CIC linked to the IPRS in 2016 through API integration with the IPRS database. The IPRS can be accessed via browser, SMS and email. The cost is negligible with a pay-per-search fee structure. For each sale or claim to be processed, CIC uses an individual's ID number to access their verified personal information on the IPRS.

The information captured in the IPRS is often more extensive than what is included in the CIC CRM. This information can be used to update CIC's internal client data, allowing them to better segment their customer base in the future.

Advantages and disadvantages of this approach
The benefits of the IPRS are that it provides a single version of the data, which is verified and endorsed by government. It is cheap and easy to use, with no complex integration of systems required.

The disadvantages of the IPRS are that certain key information is not (yet) captured by the database (such as tax PIN numbers), and certain functionalities are not (yet) added (such as automatic death notifications). These functionalities are currently being added.

Why this approach?
Previous approaches have included the manual correction of data during customer engagement, collaborating with subsidiaries to cross-reference records, and partnerships with other FSPs or insurance providers. But concerns around data-sharing between organisations have made these approaches generally unsuccessful; and, being a government-led initiative, the IPRS doesn't have these challenges. Moreover, the IPRS is cheaper than other solutions would have been, given that the government has set it up for public good.

Impact of solution
The time required for many of CIC's processes has been reduced significantly. Moreover, the easy access to complete customer records has given CIC greater security and has allowed it to onboard clients that may previously have been difficult to identify.

Future plans
CIC's future plans for its data activities are to extract more value from what it currently has. The additional information added to its CRM through the IPRS database allows for richer customer segmentation and analysis, contributing to improved product design.

ToGarantido[34]

About the company

Introduction	ToGarantido is an online microinsurance broker, founded in 2014.
Number of customers and target market	1,000+ active customers. Focus on the new Brazilian middle class, specifically those who lost their jobs in recent years: 95% of customers are accessing insurance for the first time.
Where they operate (geography)	Brazil
Product and distribution channels	• *Products:* Three different health plans, designed in partnership with Chubb, an affinity insurance provider • *Distribution:* Through the ToGarantido chatbot platform
Overall client-data use	ToGarantido's platform applies to insurance sales, and it uses sign-up data through a chatbot.

34 https://www.togarantido.com.br

About the use case

Specific business challenge
This case study describes ToGarantido's use of sign-up data and a chatbot for sales. The specific challenge in sales that this approach tackles is expensive client onboarding.

Client data used to solve the challenge
ToGarantido's chatbot, "Iara", collects information required from customers to facilitate the onboarding process. The insurance products on offer through the ToGarantido website are designed to be relatively simple to understand and have a basic underwriting structure. The data collected includes an individual's national ID number (to identify them), ZIP code (to broadly assess risk, based on location), email address and cellphone number.

Clients that search ToGarantido's website for information about insurance products are prompted to engage with its chatbot. Using a decision-tree approach, the chatbot guides the customer through pointed questions and collects the relevant data. Based on the basic information provided, the chatbot suggests a specific product and gives the individual a quote based on a simple risk assessment. The purchase is then completed through the chatbot.

Advantages and disadvantages of this approach
The benefit of this data is that it is generated automatically through the sales process in a format that is ready for use by the chatbot and insurance provider in the backend. There is no need for data-sharing agreements or restructuring and cleaning existing datasets. Moreover, the savings on the sales process are considerable, as agents do not have to be trained, deployed and managed, while prospective clients still engage in a human-like interaction.

The disadvantage of the current bot is that it does not apply natural language processing. This means that there are certain questions that the chatbot cannot deal with, which reduces the extent to which automation is possible.

Why this approach?
Previously, ToGarantido ran its chatbot through Facebook Messenger, which meant clients had to be redirected from the ToGarantido website to a different app. This formed a break in the process, which increased the risk of losing clients mid-way and decreased trust. Moreover, Facebook Messenger placed certain technical limitations on the ToGarantido chatbot, whereas the current independent platform allows the company more freedom.

Impact of the solution
With its activities and results to date, ToGarantido has proven that selling microinsurance online is a viable model, allowing for other insurance providers to go this route and access an entirely new target market.

Future plans
ToGarantido plans to expand its data and analytics for more accurate risk assessment, to decrease the premium amounts. For example:

- Bots could use the information provided by ToGarantido's clients to further complete the profile through a search of unstructured information online.
- Experiments with Facebook could facilitate the entire sales process through a Facebook page.
- Online bookings of doctor's appointments could provide more information about clients and their health activities.

CoverApp[35]

About the company

Introduction	CoverApp is an insurtech platform developed by AB Consultants. It was launched in 2017.
Number of customers and target market	150+ customers, excluded by the current high premiums on offer in the market for asset cover
Where they operate (geography)	Kenya
Product and distribution channels	• *Product:* Moveable asset cover • *Distribution:* Through the CoverApp
Overall client-data use	The CoverApp application facilitates the entire customer journey for insurance. It is an example of innovative data use for sales and claims, through the use of image (meta)data.

About the use case

Specific business challenge

CoverApp offers three movable-assets insurance plans through its app.[36] This case study focuses on CoverApp's use of image metadata for sales and claims processing. The specific challenge that it solves in these two value chain steps is how to make the processes less onerous and expensive, to remain an affordable service and improve client satisfaction.

Client data used to solve the challenge

CoverApp utilises image metadata to complete the underwriting process and to assess claims. Customers provide an image of the asset to be

35 https://www.coverapp.co.ke

36 CoverApp is owned by Bima Kenya Agency, and acts as an agent.

covered through the app, of which the image itself, timestamp and geographic location are analysed.

To sign up to the various products offered through CoverApp, individuals are requested to complete a digital form and upload an image of the insurable item. All items are covered at three possible predetermined fixed amounts, and for premiums as low as USD1 per month. The simple design of the products removes the need for individual assessment and pricing, thus reducing costs. The claims process works similar to the sign-up process: Individuals are requested to fill out a form on the app and provide either an image of the damaged item or of the police form, stating it was stolen. In case of suspicious claims, an inspector from the insurance provider will be requested to investigate. The image itself provides proof of the insured asset's existence and state. The timestamp indicates whether the image was taken close to the time of signing up or submitting the claim. The geographic location indicates whether an image of home contents was indeed taken at the home address of the individual. There are different image requirements for different insurable items (for example, motor insurance requires images of various specific car parts, whereas electronics would require images of serial numbers).

Advantages and disadvantages of this approach

The advantage of this data type is the ease with which customers can submit multiple data points through one file.

The disadvantage is that the current approach is still largely manual, as some subjective interpretation is required to determine the validity of a sign-up or claim.

Why this approach?

The main mission of CoverApp's founders is to make insurance accessible and convenient for a young middle-class population. To move away from an insurance process that has traditionally relied heavily on spreadsheets and documents, the company needed to find a data type that provided as much information as possible through as little engagement with the customer as possible. Images, with their rich body of metadata, provide this solution.

Impact of the solution

Given that CoverApp was launched in December 2017 (recently), it is difficult at the time of writing to point out the impact that the solution has had. However, the more automated process has proven to lead to time-savings and cost-savings, which makes insurance more affordable, attractive and valuable to a target market whose members have until now not insured their home contents.

Future plans

The main enhancements in the near future hinge on reducing human involvement. CoverApp's future plans revolve around applying increased AI to the analysis of the collected data. This would further reduce time and cost involved in the current sign-up and claims processes. Incorporating AI, a risk-tiering tool will be developed to assist with the claims process. Based on certain characteristics of the data and the claim submitted, CoverApp will be able to categorise claims into three different "levels" of riskiness. Subsequently, human claims assessors will only get involved with the riskier claims, and the lower levels of risk will be handled automatically.

Bibliography

Arnoldussen, L., & Hauner, W. (2016). Innovative risk solutions using big data. Baden-Baden: Munich Re. Retrieved from https://www.munichre.com/site/corporate/get/params_E-1576919391_Dattachment/1283175/MunichRe-Presentation-16-10-24-Baden-Baden.pdf

Breading, M. (2017). Analytics in insurance: Start fast, accelerate value. Boston: Strategy Meets Action (SMA). Retrieved from https://www.avanade.com/~/media/asset/research/sma-analytics-in-insurance.pdf

CGAP. (2014). Projecting impact of non-traditional data and advanced analytics on delivery costs. Retrieved from https://www.slideshare.net/mobile/CGAP/projecting-impact-of-nontraditional-data-and-advanced-analytics-on-delivery-costs

Chitta, K., et al. (2017). Top 10 trends in insurance in 2017 - What you need to know. Capgemini. Retrieved from https://www.capgemini.com/wp-content/uploads/2017/07/insurance_trends_2017_web_0.pdf

Cooper, B., et al. (2018). Biometrics and financial inclusion: A roadmap for implementing biometric identity systems in sub-Saharan Africa. Cape Town: Cenfri. Retrieved from https://cenfri.org/publications/biometrics-and-financial-inclusion-a-roadmap-for-implementing-biometric-identity-systems-in-sub-saharan-africa/

De Villiers, L., & Chetty, R. (2018). Survey: Data in FSP decision-making - Findings from six African countries. Cape Town: insight2impact. Retrieved from http://i2ifacility.org/insights/articles/do-fsps-in-africa-use-data-to-drive-their-decisions?entity=news

Hunter, R. (2018a). AgriGO: A farmer's tool to grow greater financial harvest. Cape Town: insight2impact. Retrieved from http://www.i2ifacility.org/system/documents/files/000/000/069/original/AgriGO_-_A_farmer's_financial_tool_to_grow_greater_finanical_harvest_i2i_July_2018.pdf?1532604835

Hunter, R. (2018b). Inclusive: Identification that removes barriers to financial inclusion. Cape Town: insight2impact. Retrieved from http://i2ifacility.org/system/documents/files/000/000/068/original/Inclusive__Identification_that_removes_barriers_to_financial_inclusion.pdf?1532604586

Maeder, P. (2017). The smart insurer: More than just big data. Report Issue 004. London: BearingPoint Institute. Retrieved from https://www.bearingpoint.com/files/BEI004-17-The-smart-insurer.pdf&download=0&itemId=389133

Merry, A., Prashad, P., & Hoffarth, T. (2014). Microinsurance distribution channels: Insights for insurers. Paper No. 33. Geneva: ILO Impact Insurance Facility.

Nordin, K. (2016). Data and analytics for business decision-making: Breakdown of data value chain. Scoping Note 2. Cape Town: insight2impact. Retrieved from http://access.i2ifacility.org/Publications/ScopingNote2_DataValueChain.pdf

Omarjee, L. (2017). Discovery gets global recognition for its Vitality model. Fin24. Retrieved from https://www.fin24.com/Companies/Health/discovery-gets-global-recognition-for-its-vitality-model-20170629

Smit, H., et al. (2017). Insurtech for development. Research Study. Cape Town: Cenfri. Retrieved from https://cenfri.org/wp-content/uploads/2017/11/InsurTech-Research-Study_March-2017.pdf

Tappendorf, T. (2018). Why focus groups aren't enough: A surprising lesson in marketing insurance in Africa. Retrieved July 23, 2018, from https://nextbillion.net/marketing-insurance-in-africa/

info@i2ifacility.org
@i2ifacility
www.i2ifacility.org
+27 21 913 9510

Establlsihed by

FSD Africa, Nairobi, Kenya
info@fsdafrica.org
@FSDAfrica
www.fsdafrica.org

Department for International Development
enquiry@dfid.gov.uk
@DFID_UK
www.gov.uk

About i2i

Insight2impact | i2i is a resource centre that aims to catalyse the provision and use of data by private and public-sector actors to improve financial inclusion through evidence-based, data-driven policies and client-centric product design. i2i is funded by the Bill & Melinda Gates Foundation in partnership with the MasterCard Foundation. i2i is jointly hosted by Cenfri and Finmark Trust.

About FSD Africa

FSD Africa is a non-profit company which aims to increase prosperity, create jobs and reduce poverty by bringing about a transformation in financial markets in Sub-Saharan Africa (SSA) and in the economies they serve. It provides know-how and capital to champions of change whose ideas, influence and actions will make finance more useful to African businesses and households. It is funded by the UK aid from the UK Government. FSD Africa also provides technical and operational support to a family of 10 financial market development agencies or "FSDs" across SSA called the FSD Network. FSD Africa has partnered with Cenfri to develop insurance markets in SSA.