alt.AI

FAILURE TO ACCESS

Testing transparency in data protection and artificial intelligence in South Africa

FA.I.LURE TO ACCESS
Testing transparency in data protection and artificial intelligence in South Africa
ALT Advisory, 2022

This report forms part of a research series produced for **ALT AI**, an ALT Advisory Special Project focused on safeguarding fundamental rights in an AI-driven future. Access all publications at ai.altadvisory.africa.

The research team includes Tara Davis, Murray Hunter, Wendy Trott and Zahra Abba Omar. This report was designed by Wilna Combrinck.

Suggestion citation
ALT Advisory, *FA.I.LURE TO ACCESS*, September 2022, accessible at ai.altadvisory.africa.

Acknowledgements
This work was carried out in the context of the Africa Digital Rights Fund ("ADRF") with support from the Collaboration on International ICT Policy for East and Southern Africa ("CIPESA").

Disclaimer
Please note that while every attempt has been made to ensure that the information in this report is up-to-date and accurate, there may be errors and omissions. The research in this report is provided for general guidance on matters of interest and does not constitute legal advice. ALT Advisory is not responsible for any errors or omissions, or for the results obtained from the use of this information.

Connect with ALT Advisory
altadvisory.africa
@altadvisory
connect@altadvisory.africa

TABLE OF CONTENTS

ACRONYMS

ADRF	The Africa Digital Rights Fund
AI	Artificial Intelligence
CIPESA	The Collaboration on International ICT Policy for East and Southern Africa
DHA	The Department of Home Affairs
PAIA	Promotion of Access to Information Act 2 of 2000 (South Africa)
POPIA	Protection of Personal Information Act 4 of 2013 (South Africa)
UN	United Nations

GLOSSARY

Artificial Intelligence	A broad term for a computer or software system's ability to be programmed to 'think' like a person, for example, to analyse information, look for patterns, or make decisions.
Algorithm	A set of rules or instructions that a computer or software system is programmed to follow in order to process information or perform a task.
Automated processing	Any tech-enabled processing of personal information without ongoing human involvement.
Data Subject	The person to whom personal information relates.
Personal information / personal data	Information relating to a data subject that identifies the data subject. This includes but is not limited to contact information, information relating to race, gender, sex, pregnancy, national, ethnic, or social origin, information relating to medical, financial criminal, or employment history, and biometric information.
Processing	Any operation or activity concerning personal information which includes but is not limited to the collection, recording, collation, storage, alteration, and use of personal information.
Responsible Party	A public or private body which determines the purpose of and means for processing personal information. This body is responsible for ensuring compliance with South Africa's data protection legislation, POPIA.

1. INTRODUCTION

Who is using artificial intelligence in South Africa – and what are the implications for data protection?

This report documents our attempts to test how a selection of prominent companies and government agencies in South Africa have implemented procedures, in terms of data protection law, for people to access information about how their personal information is being processed, and whether these access procedures can shed light on how artificial intelligence ("**AI**") is being used for consumer profiling and other data processing in South Africa.

The real and potential harm posed by AI technologies to human rights is the subject of a growing body of research, activism, and policy work across the world. Yet, as we note in other research produced for this series, the pace of technological developments far outstrips the pace of regulation.

Aside from gaps in regulation, in practical terms, there is limited public knowledge about how companies are deploying AI in South Africa, beyond information the companies themselves have made available in their marketing and publicity material.

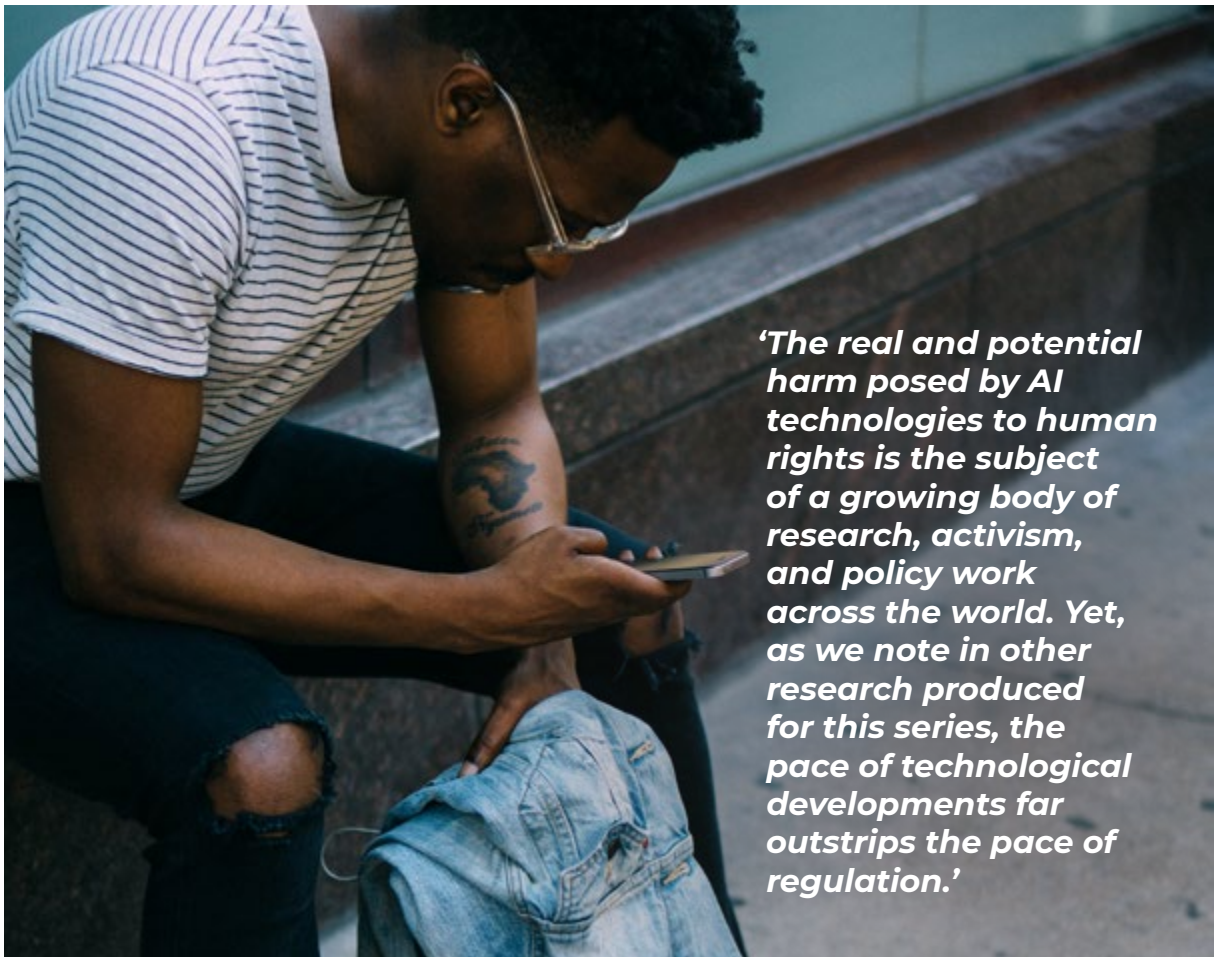
However, South Africa's data protection law, the Protection of Personal Information Act 4 of 2013, provides one possible avenue to seek transparency on the use of artificial intelligence, through its mechanisms for any person to access information about the processing of their personal information, in their capacity as a data subject. (In this regard, see a more detailed outline of the Legal Framework in section 2.)

We tested this mechanism by submitting requests to 14 prominent companies in the spheres of banking, insurance, retail, and e-commerce, and two government bodies.

Following the procedures outlined in each company's privacy policy or manual developed in terms of the Promotion of Access to Information Act 2 of 2000, we formulated a template request for information, to be submitted to each entity. This request sought to confirm what personal information was held about the requester; the identities of any third parties who had been given access to the personal information; whether that information had been subject to automated processing (in other words, processed using AI) – and to what end.

Though the law provides that we should be able to exercise and protect these rights, it proved to be inexplicably difficult to ascertain whether companies were using AI to process our own personal information, and if so, how. Our attempts to use the formal procedures to access this information suggested that in the majority of companies, the procedures are not properly established or implemented, with low levels of awareness and understanding among key staff of the relevant laws, the companies' own procedures, and issues related to data protection and artificial intelligence. Five of the 16 institutions (31%) failed to respond to any information request, and nearly half (45%) of those that did respond provided no substantive answer to our questions about their possible use of automation. Those answers we did receive were generally vague and undetailed.

While this is very preliminary research, it suggests a long and rocky path ahead to realise the rights of data subjects in relation to automated processing, and more generally towards more transparent and accountable use of artificial intelligence. The urgency for better regulation of artificial intelligence must be matched with better oversight and accountability of entities using AI through existing regulation, and better implementation and enforcement of data protection law as a whole.



Algorithmic (non-)transparency

The lack of transparency and accountability in uses of AI became a feature of an industry inquiry into racial bias in South African medical aid schemes' reimbursements to black health providers in 2021. The inquiry's interim findings found that a lack of algorithmic transparency and accountability limited the ability to scrutinise decisions taken by medical aid schemes – even for some of the companies themselves. Two of the three schemes under investigation, GEMS and Medscheme, could not provide information about the workings of the algorithm used in their decision-making, as it was purchased or licensed from an international provider that would not disclose these details.

While the inquiry's interim report did not find evidence that algorithms contributed to the discriminatory outcomes it identified, it remarked that:

"Without transparency, at least in relation to the factors driving algorithmic decisions, one can never properly assess if the algorithms … are not racially discriminatory and/or lead to racially discriminatory outcomes. In our view it is undesirable for South African companies or schemes to be making use of systems and their algorithms without knowing what informs such systems."

2. THE LEGAL FRAMEWORK

The Protection of Personal Information Act 4 of 2013 ("POPIA") is the legislative framework that regulates the processing of personal information in South Africa and was enacted to give effect to the right to privacy. POPIA provides each person with certain rights that empower us to understand how our information is being processed, and have a degree of control over how it is being processed.

POPIA provides data subjects with nine different rights but for the purpose of this report we are only concerned with two – the right to access and the right concerning automated processing.

The right to access information

The right to access is provided for in section 5(b) of POPIA which states:

> "[the right] to establish whether a responsible party holds personal information of that data subject and to request access to his, her or its personal information as provided for in terms of section 23."

Section 23 provides that the data subject should prove their identity and may request confirmation of whether or not an entity holds their information, for free. It further specifies that the data subject may request the record of the personal information that they hold, or a description of it – including information about the identity of all third parties who have had access to it. The data subject may be charged a fee for access to the record.

Certain sections of the Promotion of Access to Information Act 2 of 2000 ("**PAIA**") apply to the exercise of this right. A responsible party may refuse to disclose the requested information if one of the lawful grounds of refusal, provided for in PAIA, applies. The manner of access is also governed by PAIA which means that the form of making such requests must comply with PAIA.

In exercising this right, we submitted requests to several entities, asking the following:

> "Please provide me with records that show –
>
> 1. The personal information that you hold about me.
> 2. The identity of all third parties, or categories of third parties, that have or have had access to my personal information."

The right to access information plays an important role in enabling the exercise of additional rights – such as the right to request the correction or deletion of information, the right to object to the processing of information and the rights related to direct marketing and automated processing. It is impossible for a data subject to exercise these rights without first understanding whether the responsible party holds their personal information, and how they are processing it. This is evident in the discussion below concerning the right related to automated processing.

The right concerning automated processing

The right concerning automated processing is narrow in scope and requires sufficient information to enable a data subject to protect the right.

It is provided for in section 5(g) of POPIA and notes:

> "[the right] not to be subject, under certain circumstances, to a decision which is based solely on the basis of the automated processing of his, her or its personal information intended to provide a profile of such persons provided for in terms of section 71."

POPIA does not explicitly[1] define 'automated processing' but the term has come to be understood as the processing of personal information through the use of computers or computer software i.e processed using AI.

Section 71(1) elucidates the right by providing as follows:

> "Subject to subsection (2), a data subject may not be subject to a decision which results in legal consequences for him, her or it, or which affects him, her or it to a substantial degree, which is based solely on the basis of the automated processing of personal information intended to provide a profile of such person including his or her performance at work, or his, her or its credit worthiness, reliability, location, health, personal preferences or conduct."

Section 71 narrowly circumscribes the right. In terms of this section, it only applies to instances where automated processing was used to make a decision about a data subject – and such decision must have significant consequences. Further, the decision must have been made with no human intervention – it must be based "solely on the basis of the automated processing of personal information." Further still, such automated processing must have created a profile about the data subject and the profile must concern one of the seven listed categories.

In effect, the right only applies in very narrow circumstances.

---

1 POPIA does provide a definition for 'automated means' in section 3(4), being "any equipment capable of operating automatically in response to instructions given for the purpose of processing information." and when read with section 1, may provide a definition.

Despite this, it is important for data subjects to understand whether their personal information has been used for automated processing and if so – how. Without such information, it is impossible for a data subject to know whether a decision was made about them that was based solely on automated decision making. This makes it difficult for data subjects to protect their right in terms of section 5(g) read with section 71 of POPIA.

In making our requests, we wanted to establish whether entities would provide us with sufficient information to enable us to protect this right. We accordingly also requested access to records that show:

1. Whether my personal information has been used for automated processing.
2. Whether automated processing of my personal information provides or was intended to provide a profile about me.
3. If a profile was provided about me, whether it was used to make a decision about me, and what the decision concerned.

As detailed in section 4, we were unable to gather sufficient information to allow us to effectively protect the right concerning automated processing.

3. THE REQUESTS

Who we asked

We identified 14 prominent companies, in sectors that are widely associated with data processing, and where a member of the research team had an existing account or client relationship. We also selected two government agencies with mandates that may provide for extensive processing of personal information.

Category	Entity
Banking	Capitec, First National Bank, Nedbank, Standard Bank
Insurance	Discovery Health, Liberty, Old Mutual, Outsurance
e-Commerce	Netflorist, Takealot, Superbalist
Retail	Checkers, Pick n Pay, Woolworths
Public Bodies	Department of Human Affairs, National Department of Health

What we asked

Drawing on the legal framework outlined in section 2, we formulated a template request with five questions for each entity:

"Please provide me with records that show:

1. *The personal information that you hold about me.*
2. *The identity of all third parties, or categories of third parties, that have or have had access to my personal information.*
3. *Whether my personal information has been used for automated processing.*
4. *Whether automated processing of my personal information provides or was intended to provide a profile about me.*
5. *If a profile was provided about me, whether it was used to make a decision about me, and what the decision concerned."*

4. SUMMARY OF FINDINGS

16 requests submitted

8/16 responded

5/16 did not respond at all

3/16 provided a partial response

14/16 had publicly accessible privacy policies

13/16 explained how to exercise your rights

5/16 used a tailored request process

6/16 used the PAIA request process

2/16 used an automated request process

3/16 did not define the request process

6/16 requested ID

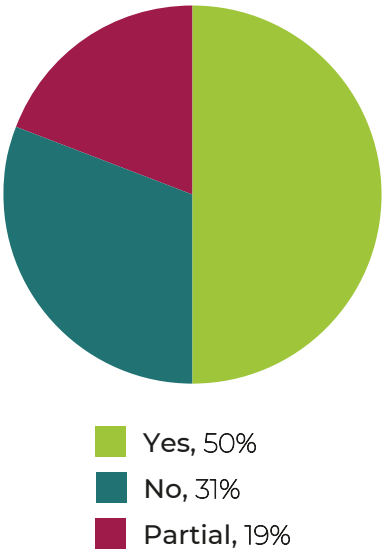
0/16 charged a fee

11/16 answered the question concerning the personal information they hold

8/16 engaged with the questions concerning automated processing

5. WHAT WE LEARNED: PROCESSES

Did the request get a response?



13 days
Fastest response (Liberty)

42 days
Slowest response (Standard Bank)

Before turning to an analysis of the answers we received, this section gives a brief overview of how the requests were made, and what we learned about the policies and practices in place (or, in some cases, not in place) for people seeking information about AI and data processing in these companies and agencies.

Variances in request procedures

There were some major differences in the prescribed request processes among the bodies we approached.

PAIA requests

Six companies' privacy policies explicitly provided that data subject requests should be made using a PAIA request form, in the same procedure that would be used for any other request for access to information.

Tailored request processes

Seven companies' privacy policies provided some form of tailored request process, such as an online, automated request portal, or a customer service hotline or email. It should be noted that most of these tailored processes were not suitable to our specific requests. In the case of Capitec Bank and Pick 'n Pay, the tailored process was non-functional (see section 4.2). In the case of Takealot and Superbalist, the automated request processes were designed only to provide a record of the requester's personal information and did not enable us to request further information about automated processing and decision-making. In these instances, we submitted follow-up requests using a PAIA form.

No process specified

Three entities, including the two public bodies we approached, did not specify how data subject requests should be made. In these instances, we resorted to submitting our request using a PAIA form. Notably, the two public bodies, the Department of Home Affairs and the National Department of Health, do not appear to have any privacy policy published on their websites, which presents a significant obstacle to the rights of data subjects.

Case studies: requests in process

To illustrate some of the variations and inconsistencies in many of the procedures we tested, and the significant obstacles for data subjects exercising their rights, we share further details of some of the request processes. Full details of the process considerations for each entity are included in Table 1, attached in the Annexure.

Capitec

> **Takeaways:** *A tailored request process that offers multiple ways for data subjects to contact the company but call centre staff did not seem to have the necessary information to handle or refer requests. However, Capitec has now launched an online request portal.*

Capitec Bank's website includes a 'Privacy Centre' with links to all its policies, FAQs and explainers on how its policies work, and contact details for the responsible parties at the bank. (Note: The Privacy Centre has been updated during the course of this research, and now includes links to automated request forms.)

At the time of our request, Capitec's Privacy Centre provided both a call centre number, and a phone number, for data subjects to make their requests. It notes:

> *"If you are a data subject wanting to exercise your data privacy rights with Capitec, you can contact our Business Support Centre on one of the numbers below or visit us in any of our nationwide branches. One of our consultants will gladly assist with capturing and actioning any of your requests, objections or information."*

Unfortunately, we were unable to make a data subject request using these numbers, or even reach a staff member who could advise us on how to make a request. We made 10 separate calls, and were transferred 8 times, to 5 different departments – including to a staff member in Capitec's internal human resources department, who could not say why we had been referred to them. In total we spent nearly 30 minutes speaking with or on hold with Capitec representatives.

However, we had more success after submitting a PAIA request to the email addresses provided on the Privacy Centre which we did on 29 April 2022. Capitec's Privacy Manager replied after 10 days to confirm receipt, and Capitec submitted a detailed response to the information request within 28 days of submission.

On 6 May 2022, the Privacy Centre was updated to include an online form for data subject requests, and Capitec's Privacy Manager noted our feedback about difficulties using the phone numbers provided, writing "Thank you for bringing the concern to my attention. I will raise it with the necessary teams to address it."

At the time of this report's publication, the phone numbers are still listed on the Privacy Centre in addition to a link to the online request form.

The substance of Capitec's response is detailed further in section 5.

Takealot

> **Takeaways:** *An automatic request process that was too narrowly designed to accommodate our request, took 28 days to be processed, and concluded without providing us any information.*

Takealot has developed an automated process for data subjects to exercise their rights. Registered users can log into their account at any time and may automatically access and manage some of the personal information that Takealot has about them. Data subjects accordingly do not need to lodge a formal request to access the following information: personal details including name, contact information and business details (if linked); newsletter subscriptions; order history (limited to the last three months); invoices; products included in a wish list and a history of the reviews they have submitted.

The website notes that this information "and more" may be requested by submitting a personal information request. To submit such a request, users simply click on a button that says, 'submit request' and one is automatically lodged. We received an email notifying us that the request had been received and we were provided with a request ID number.

The automated nature of the request meant that we were unable to submit our specific questions.

We received an email after 28 days noting that our "privacy-related request [had] been resolved." We were directed to the Takealot privacy portal to view the completed request. The privacy portal reiterated that the request had been resolved and that the personal information we requested had been sent via email which included a link where we could access and download the information. It noted that the information would expire in 60 days and the following was noted:

> *"Due to the size of some processed data, we have instead provided a description of this information:*
>
> - *Takealot records Customer Service interactions, including email, phone and social media communication.*
> - *To provide a better shopping experience, Takealot collects, records and analyses information about how you use our shopping platforms, including the types of products that you view or engage with, the features you use, the actions you take and the time, frequency and duration of your activities."*

However, the only email we received was one that sent us a link to the privacy portal, and we could not find the information there. Despite Takealot stating that the request had been resolved, we did not receive the email and were unable to access any response on the portal. There was no contact information provided to allow us to follow up. Although this may have been a technological error, we did not receive any information and for the purpose of this report we have accordingly considered Takealot as non-responsive in terms of substance.

Superbalist

> **Takeaways:** *An automatic request process that was too narrowly designed to accommodate our full request, but which was processed within a day. However, Superbalist failed to respond to a follow-up request in terms of PAIA.*

Superbalist has developed an automated process for data subjects to exercise their rights. To access information, registered users can log into their Superbalist account and automatically log a request for information by simply clicking on a button. The automated nature of the process makes it impossible to request access to specific information and we were accordingly unable to submit our questions.

The first time we submitted a request, they responded the next day, but we had difficulties accessing the response – their automated system incorrectly flagged the tracking number as "expired or invalid". However, we contacted the Superbalist team through their help centre and noted our difficulties with access. A member of their team responded promptly and we spent some time going back and forth until the issue was resolved.

We submitted a second request to test the automatic process and received a response on the same day. We received an email noting that the request was completed, and a link was provided where we could access and download the information. It noted that the information would only be available for 60 days and stated the following:

> *"Due to the size of some of your data, it won't be included in full within the download file you receive. We've provided a description of this information below:*
>
> - *Superbalist recording Customer Service interactions, including email, phone and social media communication.*
> - *To provide a better shopping experience, Superbalist collects, records and analyses information about how you use our shopping platforms, including the types of products that you view or engage with, the features you use, the actions you take and the time, frequency and duration of your activities. This data gathered while you use the shopping platforms is included as a summary only."*

The response stipulated that if we require the information in full, we would be required to make a request in writing which is subject to the prescribed fee.

The response did not fully address our queries, so we submitted a further request using PAIA on 12 May 2022. We still had not received a response by the time we published this report.

Checkers

> **Takeaways:** *A relatively easy process that was resolved within the stipulated time frames, without unnecessary formalities – even though the response did not address our questions.*

Checkers has prescribed a form that data subjects should complete when making a request for access to their information. The request was emailed to Checkers and they provided a substantive response, via email, after 25 days. The data subject was not asked to prove their identity and no fee was charged.

Although Checkers did not provide us with a full answer to our questions about the use of automated processing (see Section 5), their process was easy to follow, devoid of unnecessary formality and they responded within the 30-day period.

Pick n Pay

> **Takeaways:** *Customer service representatives were able to provide some answers about what personal information is held, and how it is used, but the formal procedures for information requests were non-functional.*

Pick n Pay, like Woolworths, provides for data subjects to exercise their rights through customer service channels. The Pick n Pay privacy policy states:

> *"If you have questions about this Privacy Statement or wish to exercise your rights in terms of access to, correction, or deletion of your information, please contact us via our Customer Care Line (080 011 2288) who will attempt to resolve your query."*

Though the privacy policy does also provide contact details for the Information Officer, we decided to follow the process outlined in the policy.

However, while our telephone request to Woolworths went relatively smoothly, we ran into several obstacles trying to follow Pick n Pay's process. In the first instance, the phone number provided in the privacy policy appears not to be active any longer; we were directed to call Pick n Pay's call centre. From there, our call was directed to the Smart Shopper department, Pick n Pay's loyalty programme. The representative we spoke to in the Smart Shopper department advised us that the only personal information they hold are the contact and identity details used to register an account, which are proactively available to customers through an online portal, combined with any transactions linked to the Smart Shopper card.

We were advised verbally that information about the frequency of purchases and purchase preferences are analysed to provide personalised shopping vouchers. However, the interaction was defensive, and we were unable to get the answer in writing. We asked to be transferred to Pick n Pay's Information Officer, but all lines were busy. When we called again, the customer representative we spoke with did not know who the Information Officer was initially; she was advised by a colleague and attempted to transfer our call, but the call dropped. We attempted to call back a third time and were placed on hold immediately. After waiting several minutes, we hung up and decided to submit a PAIA request.

Pick n Pay had not responded to the PAIA request by the time the research was concluded.[2]

Woolworths

> **Takeaways:** *Customer service channels dealt with our data subject request relatively well compared to similar procedures in other companies, though their response did not address all our queries.*

The Woolworths privacy policy provides for data subject requests to be made directly through customer service channels.

We found it comparatively easy to make our request verbally and staff appeared to understand the nature of our request and the company's own policies and procedures. A customer

2 We did receive a reply from Pick n Pay, nearly three months after the request was submitted and after the statutory 30-day period had lapsed. It detailed the categories of information held about the requester but did not address our queries related to automated processing.

service representative said the company held basic demographic information about the requester, and a transaction history, which would be provided by email. The representative also forwarded our query to the 'Wrewards' department (Woolworths' loyalty programme) and the IT department to determine if there had been any automated processing of personal information. We received a follow-up email six days later, which did not provide any further findings, and did not include a transaction history. Although there was no response to our query on automated processing, and the response did not appear to include all information about the requester, we considered the request procedure to be concluded, and decided not to submit further queries.

Discovery

> **Takeaways:** *An automated process that was not narrowly defined and allowed the requester to specify exactly what information we wanted.*

Discovery provides data subjects with the option to submit their requests through their automated portal or manually by completing and submitting a prescribed form. We used the automated request process which includes pre-populated content but has a textbox that allows you to provide further particulars or more details relating to the request. The automated nature of the process accordingly doesn't limit a data subject's opportunity to specify exactly what they would like access to, unlike the other automated request processes we encountered.

Discovery responded to the request within 29 days. The response is analysed further in section 6.

Public bodies

Our experiences in submitting data subject requests to the two public bodies we approached – the Department of Home Affairs and the National Department of Health – bear further mention.

It is concerning that these were the only two entities that appeared not to have *any* privacy policy accessible on their website. As such, despite having public mandates which may require the processing of very sensitive personal information, neither department appears to offer any public guidance on how data subjects can exercise their rights in terms of POPIA. In both instances, we decided to submit our query in the form of a PAIA request.

The Department of Home Affairs

At the time of our request, the Department of Home Affairs' PAIA manual available on its website was dated April 2013, which pre-dates the passing of POPIA. Nonetheless we submitted our queries to the Department using its provided PAIA form and received a response 24 days later. (For further discussion of Home Affairs' answers, see Section 6.)

Interestingly, at the time of this report's publication in July 2022, the Department's website continued to host the 2013 version of its PAIA manual, which is now accompanied with a note that "the PAIA Manual is currently being reviewed in conjunction with the POPI Act".[3] At the time of publication, nearly nine years had passed since POPIA was signed into law.

3 This note appears to have been added sometime after we submitted our request to the Department of Home Affairs, as it did not appear in available web caches of the website, including one as recent as 20 June 2022.

National Department of Health

At the time of our request, the National Department of Health did not have a PAIA manual available on its website, or a privacy policy. By failing to make these documents publicly accessible, the Department has obstructed people from taking even the first steps to exercise their rights in terms of data protection, and access to information.

Nonetheless, we submitted our queries in the form of a PAIA request, which was emailed to the Director General of the National Department of Health and the Department's Deputy Information Officer.[4]

Entities that did not respond to PAIA requests

First National Bank, Nedbank, Netflorist, National Department of Health, Pick n Pay.[5]

6. WHAT WE LEARNED: SUBSTANTIVE RESPONSES

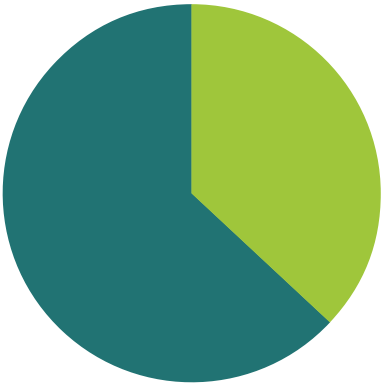
We turn now to an analysis of the substantive answers we received – noting that only half of the companies we approached provided any response at all to our queries on the use of automated processing.

It should be noted that there were huge variances in how the entities responded to our questions, especially those related to uses of AI. These answers were generally vague or hard to interpret or suggested a lack of engagement in the substance. As such, we are careful not to assume the responses we received to be exhaustive or free from inaccuracies, including where companies said that they made no use of automation. Rather, we describe them here to illustrate the uphill struggle for data subjects seeking to exercise meaningful oversight on the use of AI on their personal information.

4 Though the Department of Health website does not list which personnel are designated as the Information Officer and Deputy Information Officer in terms of PAIA, these details now form part of the court record in ongoing litigation unrelated to our research.

5 Section 27 of PAIA provides a 30-day period to make a decision on a request for information, after which the request is deemed to have been refused. However, we did eventually receive responses from First National Bank, Nedbank, and Pick n Pay, though these came substantially outside the 30-day period, after the research period was concluded.

Did they answer queries about automated processing of personal information?



Some response 37%
No response, 63%

6

companies that provided at least some answer to our queries on automated processing (of 16)

Capitec

> **Takeaways:** *The bank said it does not use automated processing. However, its use of an overly narrow definition of automated processing leaves this response open to interpretation, and other information the bank released suggested an element of automated processing may be in use.*

In response to our queries of whether the bank used automated processing, Capitec said,

> *"...we hereby confirm that your personal information was not used for automated processing to make a profile about you and therefore make a decision about that profile." (emphasis added)*

However, this was prefaced with a caveat, which read:

> *"An automated decision means any decision made in respect of a data subject, where the decision is based solely on processing of information by automated means. We take care to ensure that all decisions related to data subjects that may result in legal consequences about that data subject, is performed with human oversight and an opportunity for that decision to be challenged. Our policies prescribe that each department and division within Capitec are responsible for ensuring that business processes and decisions made include the appropriate human oversight, intervention or appropriate measures to protect a data subject's legitimate interest." (emphasis added)*

This caveat suggests that the response did not answer the question.

First, we asked Capitec if the account holder's personal information was used for automated *processing*, not only for automated *decision-making*.

Second, Capitec's response suggests that it has excluded any automated processing from its answer if this automated process included some aspect of 'human oversight [or] intervention.' However, our query very clearly applied to *any* use of automated processing, whether or not it included human oversight. Indeed, Capitec's own response seems to suggest that it does engage in automated processing of personal information, albeit with human oversight. Thus, its conclusion that our personal information has not been subject to automated processing is difficult to evaluate.

In addition, the pages of account information provided by Capitec (see below) also suggested a possible use of automated processing applied to the account. An annexure on 'Products and Services' suggested that the account was subject to 'Value-Added Services'. An explanatory note on this service strongly suggests the use of artificial intelligence:

> *To assist you with protecting your personal information we take ownership in analyzing your personal information and usage of your account for the following reasons to improve your financial life:*
>
> - *The type of transactions you perform to give you benefit in using it to your best interest*
> - *Provide you with personalised information relating to product and services that will benefit your needs based on how you use your account*
> - *Pre-assesses you in order to give you further access to products and services, limit changes, discounts and incentives*
> - *Statistical and other analysis to evaluate and improve existing and new personalized products and services to your benefit*

While it is possible that this service is provided only through manual analysis by Capitec employees, this seems unlikely given the bank's reported client base is 15-million people.[6]

Therefore, we consider it likely that Capitec does make use of automated processing, and that its answers to the contrary may reflect an overly narrow interpretation of our question.

Personal information held by Capitec

While this section deals primarily with how companies responded to our queries about automated processing, Capitec's answer about the personal information it held were detailed and worth noting.

Capitec's response included a 'Data Subject Access Request' page, which summarised a wide range of customer and account information linked to our request, and several annexures. The annexures included:

- A copy of the electronic agreement linked to the account;
- A photograph of the account holder taken in a Capitec branch when the account was opened;
- A table listing third parties with whom account information may have been shared, and the reason for sharing information with each party.

6 Capitec Bank '20 years of Capitec' 2021, accessible here.

The table of third parties given access to personal information is notable, because it was the most detailed response we received from any of the entities we approached. It lists 30 different parties, including public and regulatory bodies, and various software and support services used by Capitec. The list includes, for example, the Department of Home Affairs (for biometric verification), and Amazon Web Services (for data processing and storage).

Standard Bank

> **Takeaways:** *A confused response that did not shed light on whether the bank uses automated processing of data, but suggests limited engagement with the core issues.*

Standard Bank's responses to our queries relating to the use of automated processing appeared to be completely irrelevant to the question.

In response to the queries of whether the requester's personal information has been used for automated processing, Standard Bank answered:

> *After consulting with our Fraud team regarding this matter we confirm that there was no suspicious activity to your profile.*

It is not clear why Standard Bank's responses were confined to whether the account was subject to *fraud*, when our queries were explicitly about *automated processing,* which need not be fraudulent and may be undertaken by the bank itself.

In relation to our query about whether personal information was shared with any third parties, Standard Bank offered the same response:

> *After consulting with our Fraud team regarding this matter we confirm that there was no suspicious activity to your profile.*

Again, it is not clear why Standard Bank's response was focused only on whether or not third parties had accessed the requester's information through fraud and did not address whether the bank shared the requester's information with third parties in the ordinary course of business.

Based on these responses, we were unable to determine whether or not Standard Bank makes use of automated processing, but the answers do suggest a worrying lack of engagement with the issues. Further, it should be noted that the information that Standard Bank advised it held about the data subject was incorrect.

Superbalist

> **Takeaways:** *Superbalist did not respond to our queries about whether it used automation, but stores detailed customer behaviour data.*

After resolving the difficulties we encountered in making the first request (see section 5), Superbalist sent information in a spreadsheet under the following headings: Personal details, Addresses, Orders, Returns and exchanges, Shipments, Browsing and shopping behaviour and Marketing preferences.

The response included 186 lines of information which was presented as follows:

Count Lifetime,4,"Number of orders in since sign-up to current date";

Browse Fav Cat Last14days,0,"did user browse favourite category in last 14 days"; and

Browse Kids Hit Rate,0,"% of browsing time spent in Kids when logged in"

A significant amount of information was provided under 'browsing and shopping behaviour' which included information on the number of actions per orders which details the number of page views and clicks before completing a purchase, information on first and second favorite brands, favourite departments and categories such as "home + living" and "Dining". It also included information on hit rates, the devices that were used to access the site, favourite payment methods and the amount of times the favourite categories were browsed in the last 1 day, in the last 2 days, in the last 7 days, in the last 14 days and in the last 28 days.

We submitted a further request to Superbalist in terms of PAIA to clarify whether this or any other personal information was subject to automated processing. Superbalist did not respond. It is accordingly unclear whether Superbliast engages in automated processing and if so, how.

Outsurance

Takeaways: *Confirmed the use of automated processing, but did not provide any details of how, saying the information was 'Company Sensitive'.*

Outsurance's responses were fairly substantive, although their answers to our queries about automated processing were vague. Notably, Outsurance was one of the few institutions that provided the identity of the third parties with whom they had shared information. They also disclosed and apologised for an instance where they erroneously provided the data subject's information to a third-party supplier for a service the data subject did not want.

They responded to the automated processing questions as follows:

> *"With regards to your question on automated processing, there are many factors to our personalised underwriting model that are used to determine your premium. Your profile such as claims experience, driving experience, where the vehicle is parked, are all used in the premium determination and whether cover may be continued. Your premium reflects the outcome of this."*

In a follow-up they expressly confirmed that automated processing was used but declined to advise whether it is used in conjunction with human intervention, or any other factors that were considered. They noted that such information was "Company Sensitive." The nature of their response does seem to suggest that Outsurance uses automated processing to create a profile about a data subject but did not offer enough detail for us to establish this with certainty.

Checkers

Takeaways: *Uses automated processing for direct marketing.*

In its response to our queries, Checkers provided screenshots to prove the information that they had and advised on the internal investigation they had undertaken to establish which departments hold information about the data subject. Although no call recordings or

emails were provided, they confirmed that they did not have any, which was corroborated by the data subject.

Checkers briefly explained the data subject's communication preferences and what they had opted in to receive. Usefully, they explained the consequences of the preferences by noting "[t]his means that your cell phone number may have been used to provide direct marketing – personalised advertisements to you, based on your shopping habits."

They responded to most of our questions concerning automated processing but did not confirm or deny whether automated processing of personal information was, or was intended to, provide a profile about the data subject. They responded as follows:

> *"In addition to direct marketing, we also advertise on various other platforms like TV, radio, print, outdoor media and social media. Social media advertising occurs in terms of the user's agreed terms with each social media platform (e.g. Facebook and Google). We provide advertisement content to social media platforms and these platforms make the personalised advertisements visible to their users in terms of the terms and conditions agreed with each of their users.*
>
> *Your personal information is thus used for automated processing in terms of personalised advertising, but was not used to make any decision about you."*

It is accordingly evident that Checkers uses automated processing of personal information for advertising, but it is not clear whether a profile of a data subject is created to do so. There is a distinction between using a data subject's previous purchases to advertise the same or similar products to them versus drawing inferences from such purchases to create a profile about the data subject. Such profiles may be used to make assumptions about the kinds of products a data subject may purchase.

Checkers did, however, state that although automated processing was used for personalised advertising, it was not used to make a decision about the data subject.

Discovery

> **Takeaways:** *Discovery's answer was obfuscatory, but the company does appear to use automated processing for a range of purposes, including to profile data subjects.*

In response to the questions concerning automated processing, Discovery responded as follows:

> *"Discovery Health is not using your information to "profile" you and use it to inform decisions about you in any way which could be detrimental to you as a member of the Discovery Health Medical Scheme. Standard Operating Procedures and related systems have been used in interactions with you and for processing your data." [Emphasis added]*

They provided the purposes for which information is processed, in terms of section 8 of their Privacy Policy, which states:

> *"You understand and accept that We may process Your Personal Information for the following purposes:*

8.1. to verify the accuracy, correctness and completeness of any information provided to Us in the course of processing an application for membership or providing services related to Your membership;

8.2. for the administration of Your benefit plan;

8.3. for the provision of managed care services to You on Your benefit plan

8.4. for the provision of relevant information to a contracted third party who requires this information in order to provide a healthcare service to You on Your benefit plan;

8.5. to profile and analyse risk;

8.6. to share Your Personal Information with external healthcare providers for them to assess or evaluate clinical information, in the event that You are subject to such a clinical assessment."

Their response went on to note:

"Discovery Health does utilise automated processing to the extent that these are utilised in line with the overall privacy statement and in particular Section 8 above. Some of these systems are used on aggregated data where individual member data is anonymised or not referred to. Other systems are used to enhance servicing to optimise the routing and response to interactions."

[...]

"I would therefore like to offer you an assurance that your personal information and its protection is taken very seriously by Discovery Health, it is processed for the purpose for which it has been collected and that no data used by artificial intelligence systems has been used to inform a decision about you which would be to your detriment in any way."

Discovery's response was obfuscating. They noted that they are not using the data subject's information to profile them and inform decisions "*in any way which could be detrimental*". Their answer implies that they are using AI to create a profile about the data subject, but that Discovery does not consider this to be used to inform decisions that are detrimental to the data subject.

This assumption is corroborated by their statement that they use automated processing in line with their privacy statement. Specifically, they note "Discovery Health does utilise automated processing to the extent that these are utilised in line with the overall privacy statement and in particular Section 8 above." Section 8.5 of their Privacy Statement notes that one of the purposes for which they process information is "to profile and analyse risk". However, the extent to which automated processing is used to profile and analyse risk, if at all, is not provided.

Discovery's response is confusing, but it appears that automated processing is used to profile a data subject. However, it is unclear which decisions the profile is used to inform, if any.

This request related to the Discovery Health Medical Scheme only.

Liberty

> **Takeaways:** *Confirms it uses automated processing to assess risk in issuing an insurance policy, but the company's inability to explain the process highlights the issue of algorithmic transparency.*

Liberty's response to our queries about automated processing offer a stark comment on the challenges of algorithmic transparency – in that, initially, Liberty's own response suggested that it was unable to confirm how the requester's personal information was used for automated processing.

Liberty's initial response to our query, which was made through a formal information request in terms of PAIA, seemed to confirm that the Underwriting Department had used automated processing of the requester's personal information in the course of an application for life insurance. However, Liberty's response, which was issued through a client relations employee, took the unusual step of suggesting we ask the Underwriting Department directly for further details:

> *Your personal information such as smoker status, occupation details and income were used by our Underwriting Department to assess the risk. This is confidential information and as a result, I do not have the details of how they came to a decision. Please direct the query to our Underwriting Department so that they will provide additional information on how they made a decision. Our Underwriting Department can be contacted via [redacted email address]. They will then send the information directly to you.*

While this suggestion appears to have been motivated at least in part by a desire to preserve the requester's confidentiality, it is highly unusual – and not in line with the provisions of PAIA – for an official response in terms of PAIA to advise the requester to ask someone else in the same company or agency. (Indeed, one of the purposes of an access-to-information law is to ensure a centralised process for requesting information, which places a responsibility on the recipient to ensure that the requested information is provided unless the requester is not entitled to the information.)

Nonetheless, we emailed our queries to Liberty's underwriting department, which provided the following responses:

> *In regard to the underwriting process, the only automated part of the underwriting process was the request for and receipt of blood tests, both of which were normal.*
>
> *[...]*
>
> *The outcome of the risk scoring is to determine the terms of the insurance cover and consequent premiums quoted. The blood tests were used in the risk assessment. Since both tests were normal, they would not have affected the risk assessment negatively.*

Therefore, Liberty appears to confirm that it uses some measure of automated processing of data subjects in order to create a profile of a data subject. This is a more substantive and specific response than we received to most of our queries. However, as we observed elsewhere, Liberty's response appears to assume that the request is part of a dispute about a decision that may have been made. Therefore, while Liberty offers assurance that the results of the blood analysis were 'normal', and therefore did not contribute towards a *negative* profile, it does not shed further light on the use of automated profiling.

Old Mutual

> **Takeaways:** *Confirms the use of automated processing, but did not provide any details of how.*

Old Mutual responded to the questions on automated processing by noting:

> *"While we process your personal information via automated means, there was no automated processing with the intention of providing a profile about you."*

It appears that although Old Mutual uses automated processing, it considers its use of automated processing not to implicate the rights of data subjects as provided for in POPIA. However, the lack of detail in Old Mutual's response makes it impossible to assess how automated processing of the information is used.

Department of Home Affairs

> **Takeaways:** *Response suggested limited understanding of or technical capacity for automated processing.*

The Department of Home Affairs ("**DHA**") responded to our request but appears to have fundamentally misunderstood our questions. They did not confirm whether they hold information about the data subject, or provide access to such information, nor did they respond to any of the questions relating to automated processing. Their response focused on question 2 – the identity of all third parties that have had access to the data subject's personal information. In this regard they noted:

> *"there are no records of any application or request by any third party to have access to your personal particulars, or records as held by the Department of Home Affairs."*

They stated that the Identification Act 68 of 1997 prohibits the disclosure of the information to third parties and that "the disclosure of the information requested would, amount to unreasonable disclosure of personal information, and a violation of the right of privacy of the individual concerned."

We followed up with the DHA and asked for a response to our questions concerning automated processing. Their response alludes to the assumption that they do not conduct automated processing, but their focus remained on the sharing of information with third parties – they provided:

> *"The Department [of] Home Affairs does not have sophisticated network of automated facility within their main server. Therefore there is no way your information can be shared in any platform. Forward those questions to your other service providers that have sophisticated network of harvesting information in an automated manner, or sharing it in any way. DHA does not share information with third Parties."*

We did not follow up with the DHA again.

7. CONCLUSION

In sum, our experience shows that it remains frustratingly difficult for data subjects to exercise their right to access information. Without sufficient information, a data subject's ability to protect their rights concerning automated processing is severely undermined. Existing tools and mechanisms in law to enforce our rights at the most basic level are falling short of the moment. This may be due in part to the nascent enforcement of POPIA which only fully came into force in July 2021; the technical and legal complexity of AI and data protection issues may also be a factor. Yet, the struggles to enforce existing laws are even more pertinent, given ongoing conversations in the spheres of activism and policymaking on the need for robust new regulation of AI technologies.

We take some encouragement in noting that at least two entities (Capitec and the Department of Home Affairs) updated public information about their procedures after we made our requests. While these changes were minor, they suggest that continued public scrutiny could result in improvements to data protection processes.

Yet the dismal findings on the state of protection for data subjects' rights in relation to automated processing are even more alarming considering the positioning of our research team, and the entities we approached. We are conversant in legal and policy procedures for data protection and access to information and have significant social capital. We expect data subjects with less privilege and agency to face far greater struggles to enforce their rights. This is of great concern for efforts to promote transparent and accountable use of artificial intelligence.

Recommendations

Acknowledging the enormity of the task ahead, we propose the following recommendations:

Review and redesign of systems

- The huge inconsistencies in the quality and responsiveness of the procedures suggests that companies and government departments should urgently review, and where necessary overhaul, their procedures for data subject requests.
- The Information Regulator, and where necessary industry sector bodies, should provide guidance, minimum standards, and examples of best practice, to improve the quality of access procedures.

Action on algorithmic transparency

- The lack of clarity and, in some instances, apparent lack of candour, in the responses we received about the use of automated processing on our personal information is especially concerning for algorithmic transparency. In the short term, the Information Regulator should issue guidance to develop and clarify the definitions of automated processing, and the standards that should apply to ensure meaningful transparency on the use of AI in relation to data protection, and that data subjects' rights in relation to automation are properly protected.
- As a longer-term solution, policymakers should prioritise the development of substantive regulation of artificial intelligence, which should include considerations for an AI register, where all public and private sector parties disclose systems for automation, including information about the system's purpose, its underlying logic and working, and what measures are in place to mitigate risks of bias and abuse.

Further research and public scrutiny

- Noting that this report documents preliminary research, further research should be applied to use and test these procedures, including similar requests to other bodies not covered in this research.
- Members of the public should use these and other mechanisms, both to enforce their rights, and to draw further attention and resources to data protection processes.
- The urgency for better regulation of AI must be matched with better oversight and accountability of entities using AI through existing regulation, and better implementation and enforcement of data protection law as a whole.

ANNEXURES

Table 1 | Procedural concerns

Sector	Institution	Is their privacy policy publicly accessible?	Do they stipulate how to exercise your rights?	Format of request
Banking	Capitec	Yes	Yes	Tailored process
	FNB	Yes	Yes	PAIA request
	Nedbank	Yes	Yes	PAIA request
	Standard Bank	Yes	Yes	PAIA request
e-Commerce	Netflorist	Yes	No	Not defined
	Superbalist	Yes	Yes	Automated process
	Takealot	Yes	Yes	Automated process
Retailers	Checkers	Yes	Yes	Tailored process
	Pick n Pay	Yes	Yes	Tailored process
	Woolworths	Yes	Yes	Tailored process
Insurance	Discovery	Yes	Yes	Tailored process
	Liberty	Yes	Yes	PAIA request
	Old Mutual	Yes	Yes	PAIA request
	Outsurance	Yes	Yes	PAIA request
Public Bodies	Department of Health	No	No	Not defined
	Home Affairs	No	No	Not defined

Table 2 | Procedural concerns

	Institution	Was there proactive disclosure?	How long did they take to respond?	Was ID requested?	Did they charge a fee?
Banking	Capitec	No	28 days	Yes	No
	FNB	No	No response	Yes	No
	Nedbank	No	No response	Yes	No
	Standard Bank	No	42 days	Yes	No
e-Commerce	Netflorist	No	No response	No	No
	Superbalist	A limited amount	1 day / No response*	No	No
	Takealot	Yes	28 days	No	No
Retailers	Checkers	A limited amount	25 days	No	No
	Pick n Pay	A limited amount	N/A	No	No
	Woolworths	No	N/A	No	No
Insurance	Discovery	Yes	29 days	Yes	No
	Liberty	A limited amount	13 days**	Yes	No
	Old Mutual	Yes	33 days	No	No
	Outsurance	Yes	22 days	No	No
Public Bodies	Department of Health	No	No response	No	No
	Home Affairs	No	24 days	No	No

* *Provided a partial response through an automated disclosure, but did not respond to a further PAIA request*
** *Provided a partial response through information that is automatically available through a customer portal but did not respond to a further PAIA request*

Table 3 | Substantive concerns

	Institution	Did they respond?	Did they answer Q1?	Did they answer Q2?	Did they answer Q3?	Did they answer Q4?	Did they answer Q5?
Banking	Capitec	Yes	Yes	Yes	Partial	Partial	Partial
	FNB	No	-	-	-	-	-
	Nedbank	No	-	-	-	-	-
	Standard Bank	Yes	Yes	Yes	No	No	No
e-Commerce	Netflorist	No	-	-	-	-	-
	Superbalist	Partial	Yes	-	-	-	-
	Takealot	No	Yes (automatic)	-	-	-	-
Retailers	Checkers	Yes	Yes	No	Yes	No	Yes
	Pick n Pay	Partial	Yes (automatic)	-	-	-	-
	Woolworths	Partial	Yes	Yes	Partial	-	Partial
Insurance	Discovery	Yes	Yes	No	Yes	Partial	Partial
	Liberty	Yes	Yes	Yes	Yes	Yes	Yes
	Old Mutual	Yes	Yes	Yes	Yes	Partial	Partial
	Outsurance	Yes	Yes	Yes	Yes	Partial	Partial
Public Bodies	Department of Health	No	-	-	-	-	-
	Home Affairs	Yes	No	Yes	No	No	No

Q1: *What personal information do you hold about me?*
Q2: *What is the identity of all third parties, or categories of third parties, that have or have had access to my personal information?*
Q3: *Has my personal information been used for automated processing?*
Q4: *Has automated processing of my personal information provided or was it intended to provide a profile about me?*
Q5: *If a profile was provided about me, was it used to make a decision about me, and what did the decision concerned?*

Table 4 | Responses concerning automated processing[7]

	Institution	Is personal information used for automated processing?	How is the information used?
Banking	Capitec	No[8]	-
	Standard Bank	No[9]	-
Retailers	Checkers	Yes	For personalised advertising
	Woolworths	Yes	Transaction history is analysed to provide customised vouchers
Insurance	Discovery	Yes	In line with the purposes provided in their privacy policy[10]
	Liberty	Yes	Information from a blood-test was used as part of an underwriting process
	Old Mutual	Yes	Not specified
	Outsurance	Yes	To determine a data subject's insurance premium

Table 5 | Responses concerning automated processing

	Institution	Was automated processing used or intended to create a profile?	Was the profile used to make a decision about the data subject?	What did the decision concern?
Banking	Capitec	-	-	-
	Standard Bank	-	-	-
Retailers	Checkers	They did not say	Unclear	-
	Woolworths	They did not say	Unclear	-
Insurance	Discovery	Yes[11]	Unclear	-
	Liberty	Yes	Yes	Liberty stated that that the test results were factored into the terms and premiums of the policy. The results were deemed to be normal, and therefore "*would not have affected the risk assessment negatively*."
	Old Mutual	No	No	-
	Outsurance	No	No	-

7 Only entities that answered questions concerning automated process have been included in this table
8 See caveat noted in section 6
9 See caveat noted in section 6
10 See privacy policy purposes in section 6
11 This was not expressly stated, see section 6