Submission: Screen scraping – policy and regulatory implications

Submission by:

Wednesday 25th October, 2023

Basiq (ABN: 95 616 592 011)
7 Raglan Street
Manly, NSW 2095

Executive Summary

Basiq was founded with a clear mission: to empower consumers with more choices and control over their financial lives. I felt a growing frustration seeing the limited ways people could interact with their finances, confined mainly to traditional banking channels. This drove me to create a platform that would enable developers to build solutions directly catered to consumers' needs.

The key to making this vision a reality was to unlock the data, ensuring that consumers could share it with any application or service they found valuable. Data is a game changer - it levels the playing field allowing new service providers to have a similar understanding and insights about consumers as their existing banks do. This understanding can then be used to offer personalised, innovative and new financial services to consumers.

At the time Basiq was taking its first steps, there were no official APIs, and the concepts of Open Banking or CDR were not introduced in Australia. Supporting and advancing this technology demanded a considerable investment. The inherent nature of scraping services is brittle, requiring ongoing investment to ensure they remain operational. Over the years, the task became more challenging with banks deploying anti-bot tools, making data retrieval from some institutions nearly impossible.

Alongside this journey, I have also witnessed a positive shift in consumers' understanding and expectations regarding their personal privacy. They have become more cautious, demanding greater control and protection over their shared data.

The introduction of Consumer Data Rights (**CDR**) has provided a more reliable and secure way to access this crucial data, addressing many existing challenges. Our organisation fully backs and supports the CDR initiative. Our support takes various forms, including assisting organisations in meeting compliance requirements, creating developer-friendly tools for easy integration, and providing regular feedback to enhance the program.

We firmly believe that CDR is the future and a viable alternative to scraping. However, there are challenges that could potentially disrupt organisations accustomed to scraping services. We've detailed these challenges in this document, urging the Government to take these into consideration. The disruptions posed are significant, and for some businesses, they could severely impact their future success and their ability to continue providing value to their customers.

We remain committed to the CDR program, and I am more than willing to elaborate on any of the points discussed, and if necessary, bring in our valued customers to provide direct feedback. This engagement would undoubtedly help in shaping the way forward for the CDR program.

Kind Regards,

Damir

Responses to consultation questions

How is screen scraping currently used?

1. a) What is the scope and purpose of the data that is captured? Is the data that is captured only banking data, or does it include data from other sectors?

Screen scraping operates by using various tools to interact with user interfaces, such as websites, to gather specific data. For example, it can extract financial information from a bank's website. This technology has empowered many organisations and service providers to access vital financial data, which can then be utilised for various purposes like evaluating a person's creditworthiness or enabling financial management services.

The term 'screen scraping' is broadly used and covers many different methods of data access, including:

- **Utilising unofficial APIs that power the banks' own websites**. In modern architecture, banks often develop APIs to support their mobile applications or internet banking portals. Where available, some aggregators leverage these APIs to obtain the necessary data.

- **Employing server-side web browsers.** When unofficial APIs are not available, developers may use a web browser, feeding it commands such as logging in, and navigating to various pages that hold account and transaction data to programmatically extract the required information. This process essentially simulates the experience a consumer would have using their desktop computer to access their financial data.

The rise of screen scraping was largely driven by the necessity to access data in a more organised and automated manner, especially in instances where a more modern and secure method of data sharing through official APIs is not widely available or adopted by key financial institutions. Nonetheless, screen scraping continues to be a common practice in scenarios where APIs are not available or where screen scraping provides a more straightforward or cost-effective solution.

Currently, screen scraping facilitates access to financial data from various organisations such as:

- Banks

- Superannuation
- Non-Bank Lenders
- MyGov (for retrieving Centrelink financial statements)

Within industry, screen scraping supports many different use cases some of which include:

- Personal Financial Management - gathering financial data to help consumers understand their spending.
- Wealth Management - conducting financial assessments of individuals.
- Account Verification - verifying ownership of bank accounts, and capturing account details for setting up payments.
- Affordability Assessment - evaluating the financial situation of individuals or businesses for lending decisions.
- Round-ups - accessing transaction data to manage roundup payments.

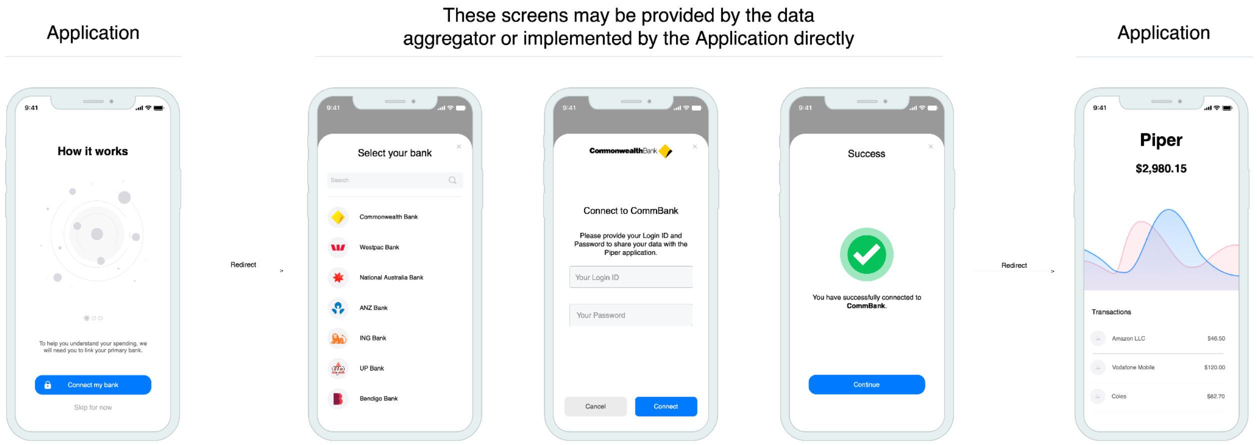
The usage of screen scraping services extends to various organisations including:

- Lenders
- Wealth and Investment Organisations
- Cryptocurrency Platforms
- Buy Now Pay Later (BNPL) Services
- Proptech Companies
- Debt Collectors and Hardship Management Agencies
- Personal Finance Management Platforms
- Salary Advance Services

It's important to note that while the above mainly illustrates the use of screen scraping in the financial sector, other industries also depend on screen scraping services in the absence of official APIs. This includes product comparison services, AI services that utilise scraping for model development, and various data aggregation services across numerous industries. Depending on the use case, screen scraping may require consumers' login credentials to access account-level information, or it might simply tap into publicly available data that doesn't require a login or password.

1. b) What steps do consumers, screen scraping service providers and businesses using screen scraping take in the screen scraping process? What information is provided to consumers through the process?

To explain the screen scraping process, we've outlined a typical flow below, highlighting the steps consumers generally follow when sharing their financial data through screen scraping services.


Application

These screens may be provided by the data aggregator or implemented by the Application directly

Application

How it works

To help you understand your spending, we will need you to link your primary bank.

Connect to CommBank

Please provide your Login ID and Password to share your data with the Piper application.

You have successfully connected to **CommBank**.

Often, a data aggregator provides API services and user interface libraries that an organisation can integrate within their application. However, organisations seeking more control over the user experience may opt to design and control the entire UI flow themselves. This approach offers the added advantage of creating a seamless data connectivity process for the consumer. By doing so, they ensure that establishing a connection to an institution feels native and smooth within the application, avoiding any jarring transitions during the process.

Due to this flexibility, the interface and interaction can vary a lot across different use cases and businesses, as it's up to them to decide on the best way to set up the UI.

However, at a basic level, an organisation would need to provide at minimum the following

- **An institution picker** - a list of institutions that the aggregator supports, letting consumers choose their institution.
- **An authentication form** - this is used to capture the login and password required to interact with the target institutions and acquire the consumer's financial data.
- **A link to the terms and conditions** - that explains how their data will be handled and used.

1. c) When is the consumer’s data accessed as a one-off, and when is longer-term or ongoing access obtained? Where ongoing access is in place, how are consumers made aware of this and can they cancel access at a later point?

In screen scraping services, the duration of data access doesn't have a predefined expiration period. Instead, it's the responsibility of the organisation and the aggregator to ensure the data is used appropriately.

The type of data access, whether it's one-off or ongoing, is determined by the specific use case at hand. For example, lending scenarios usually require one-off data access, while services like personal financial management require ongoing access to provide accurate and continuously updated insights.

Since screen scraping isn't regulated, the onus is on the aggregator and the application service provider to decide how to manage data access and when to delete the data. Most aggregators offer API services that allow organisations to Create, Read, Update, and Delete (CRUD) consumer records. Aggregators usually mandate that service providers offer features allowing consumers to cancel data sharing. This requirement prompts application service providers to include features or screens within their applications where consumers can manage or revoke data access.

In essence, while the control and management of data access in screen scraping services largely rests with the organisations and aggregators, organisations such as Basiq enforce measures that ensure consumers have control, especially in stopping ongoing data access whenever they choose.

1. d) Do you use screen scraping for purposes other than data collection (for example to undertake actions on behalf of a customer)?

Basiq strictly uses screen scraping for data collection purposes only; we do not engage in write access or take any actions on behalf of consumers. In instances where screen scraping services are utilised, we uphold a responsible approach.

For example, if a bank updates their terms of service, Basiq suspends the connection until the consumer independently logs into their banking app and accepts the new terms. This practice underscores our commitment to ensuring a clear boundary between data collection and any other actions, adhering to a responsible use of screen scraping technology.

What are the risks of screen scraping?

2. Are there any other risks to consumers from sharing their login details through screen scraping?

We believe that the risks outlined in the discussion paper are comprehensive and no additional risks have been identified. We would also like to add that aggregators can also implement measures to mitigate risks arising from the sharing of login details (major aggregators in Australia take security seriously, aiming to uphold data security standards comparable to those maintained by banks). However, it's crucial to note that the onus is on the aggregator to prioritise security and implement the appropriate controls. The commitment to security and the effectiveness of the implemented controls significantly impact the level of risk associated with the use of login details through screen scraping.

For example, Basiq has attained ISO 27001 certification and is actively working towards SOC2 compliance. Basiq has also made the decision to apply the CDR framework to our scraping services. This ensures that regardless of the data acquisition method (CDR or scraping), Basiq handles data in a consistently secure manner.

3. Do you have any data, case studies, or further information about the risks of consumers sharing their login details through screen scraping?

Sharing of data with 3rd parties

One of the shortcomings in data-sharing practices using scraping is the lack of transparency when organisations need to share consumers data with third parties, especially compared to the CDR framework. In data scraping methods, consumers often lack full awareness, or are entirely unaware, of who their data is being shared with.

Under CDR, service providers are obligated to clearly present consent mechanisms and disclose all third parties involved in data sharing. In contrast, with scraping, this level of transparency is generally not required and important information is often hidden within the terms of service.

The core issue with this practice lies in the lack of awareness for the consumer regarding how their data will be managed by the third parties. Moreover, there's uncertainty about whether these third parties have the necessary security measures in place to handle the consumers' data appropriately. This situation underscores the need for a more transparent process where consumers are fully informed and have control over how and with whom their data is shared, along with assurances on the security measures safeguarding their data.

This could lead to a range of issues concerning consumers' data, including but not limited to:

- **Misuse of data:** In some instances, data may be used for purposes beyond what the consumer initially intended. This can be particularly problematic as it might target vulnerable consumers with specific products to which they may be susceptible.
- **Uncontrolled data sharing and serious data breaches:** There's nothing to prevent third parties from further sharing consumer data with additional organisations. This means consumers' data can be duplicated and stored in various places, which is not only a breach of consumer trust and privacy but also elevates the risk of data breaches.

While the primary focus here is on screen scraping services, it's important to also shed light on other methods through which financial data is accessed, outside the control of the CDR framework. Below are some examples of other methods and entities that engage in the collection of consumer financial data, which merit consideration:

- **Digital Wallets:** Companies like Google and Apple, dominating the tap and pay market, actively collect consumer transaction data via their digital wallets. Unlike organisations operating under the CDR framework, these digital wallet providers aren't subjected to the same stringent rules, creating a disparity in data handling standards. Although the Australian Competition and Consumer Commission (**ACCC**) is reportedly investigating this issue, it's an example of data collection activities outside the CDR realm. More details can be found in this article: https://www.channelnews.com.au/apple-googles-tap-pay-under-threat/.

- **Market Insights Data:** Some major banks have established departments dedicated to providing spend insights using anonymised consumer data. While this data provides valuable market insights to businesses, the practice operates outside the CDR framework. This not only raises concerns regarding consumer awareness and consent but also points to the broader issue of unregulated financial data handling.
    - https://www.mi-3.com.au/10-05-2021/cba-and-quantium-announce-consumer-data-insights-platform

  - https://www.westpac.com.au/corporate-banking/corporate-institutional/datax/

These examples underline that while screen scraping is a primary method for accessing financial data, it's not the sole method. There are other channels through which consumer financial data is being accessed and utilised outside the rules of the CDR framework. This situation calls for a more holistic examination and potentially an extension of CDR principles to other data collection practices, ensuring a uniform, consent-driven, and transparent handling of consumer financial data across the board.

4. Could you provide any examples of actions your organisation takes to prevent or block screen scraping (if you hold the consumer's data, such as a bank), or when your company's use of screen scraping has been blocked (if you provide screen scraping services or you partner with a screen scraper to provide your services), and why?

One of the primary challenges with screen scraping services is their susceptibility to disruptions, making them less reliable compared to official APIs available in the market. Such disruptions can significantly affect businesses, especially those that integrate screen scraping as a critical part of their consumer onboarding process.

Here are some common ways screen scraping services can face disruptions:

- **Changes by Banks:** When banks alter their user interface or internal APIs, which screen scraping services depend on, it can cause the scraping services to fail. In such cases, the aggregator needs to review and update the code to adapt to the changes and ensure successful data retrieval. Such disruptions can take anywhere from days to weeks to resolve. During this period, organisations relying on these scraping services may be unable to facilitate data sharing for their consumers. This can have direct operational impacts; for example, if a consumer is applying for a loan and the scraping service for that bank is temporarily non-functional, the lender won't be able to process the loan application. This is likely to drive the consumer to seek alternative options.

- **Anti-bot Measures:** Banks may deploy anti-bot tools like Akamai to deter screen scraping. These tools use technologies such as fingerprinting to discern between a bot (automated code) and an actual user. Once these anti-bot measures are in place, they are extremely challenging to bypass. Examples of this include Macquarie Bank and Up Bank, both of which have actively taken steps to block scraping services.

These scenarios illustrate the fragile nature of screen scraping services and how they can be easily disrupted by changes or measures implemented by the data-holding institutions, thereby affecting the seamless operation of businesses relying on such services for data access.

5. Could you provide any examples of how your organisation or entities you partner with manage the risks associated with screen scraping?

We have taken a multi-faceted approach to manage the risks associated with screen scraping, ensuring robust data governance and security controls. Below are some of the measures we have adopted, along with how we work closely with our partners to mitigate risks:

- **Voluntary Implementation of CDR Framework:** We have proactively adopted the CDR framework across our scraping services. This ensures that whether the data is acquired via CDR or scraping services, we adhere to the same stringent data governance and security standards as mandated by the CDR framework.

- **Formal Risk Management Framework:** We follow a structured risk management framework to identify, mitigate, and address risks effectively. This framework guides our actions and strategies in managing the inherent risks associated with screen scraping.

- **ISO 27001 Certification:** Holding the ISO 27001 certification ensures that we have established policies and procedures to protect data within our organisation, showcasing our commitment to maintaining high standards of data security.

- **Progressing Towards SOC2 Compliance:** We are actively working towards achieving SOC2 compliance to further bolster our security environment, demonstrating our continuous effort to adhere to industry-leading security standards.

- **Utilisation of Managed Cloud Services:** By leveraging managed services from our cloud provider, we operate under a shared responsibility model. We intentionally avoid accessing the underlying operating system or network interfaces, entrusting our cloud provider to handle issues like zero-day vulnerabilities, which in turn enhances our security posture.

- **Regular Penetration Testing:** We conduct routine penetration testing of our systems, carried out by a reputable third-party specialising in this area. This practice helps in identifying and addressing any potential vulnerabilities in our system.

- **Industry-Standard Encryption Techniques:** We ensure that all data, whether at rest or in transit, is secured using industry-standard encryption techniques and services, thereby safeguarding the data from unauthorised access.

- **Security Compliance Process for Customers:** We mandate all our customers to undergo our security compliance process to ensure they have adequate controls in place, promoting a culture of security awareness and preparedness.

- **Collaborative Customer Onboarding:** We engage with our customers both during pre and post onboarding phases to ensure they have implemented appropriate controls. This collaborative approach helps in fostering a secure and compliant environment for data handling and processing.

Through these measures, and in collaboration with our partners, we strive to manage the risks associated with screen scraping effectively, ensuring a secure and compliant operational framework.

Reforms and reviews related to the screen scraping market

6. Are there other proposed reforms or legal frameworks that relate to the use of screen scraping? Are there any other international developments that should be considered?

We are not aware of any other reforms or proposed legal frameworks beyond what has been previously discussed.

The Consumer Data Right

8. a) Can you provide examples of data that is being accessed through screen scraping that cannot currently be accessed using the CDR or vice versa?

The differences in data access between screen scraping and the Consumer Data Right (CDR) framework can be nuanced but are essential in understanding the scope and limitations of CDR. Below are some examples of data accessibility differences between screen scraping and CDR:

- **Account Selection:** A key distinction is the level of control consumers have over which accounts to share. With screen scraping, all accounts and transactions become accessible upon connection establishment, unlike CDR where consumers can choose the accounts they wish to share. While this might seem advantageous from a consumer control perspective, Basiq is aware that this may make CDR less suitable for certain businesses or use cases (e.g. CDR might not be suitable for lenders in meeting their responsible lending obligations or to otherwise conduct a comprehensive financial assessment). The ability to obscure certain accounts could potentially hinder accurate affordability assessments, increasing the risk of irresponsible lending.

  We've proposed that the CDR framework could be enhanced to allow ADRs to specify the types of accounts needed for their services, instead of leaving this choice solely to consumers. This adjustment could bridge the gap between the data accessibility of screen scraping and CDR, ensuring a more thorough and accurate analysis for crucial financial decisions.

- **Specific Data Attributes:** There are some instances where screen scraping provides additional data, including Date of Birth and Running Balance for each transaction. Whilst these attributes are important, their availability varies across different institutions and may also depend on the specific account in question.

- **Enhanced Data Attributes through CDR:** Despite the aforementioned disparities, CDR tends to offer a richer and more structured dataset compared to screen scraping. For instance:
    - Screen scraping frequently fails to provide reliable access to comprehensive personal consumer data, whereas the Consumer Data Rights (CDR) framework excels in this area. This discrepancy arises because banks under CDR are obligated to return a complete set of consumer attributes, while scraping services are limited to capturing only the data visible on the front end - which tends to vary significantly per institution.
    - CDR provides access to Merchant Category Codes (MCC), aiding in better expense classification, which is beneficial for lenders aiming to understand consumer expenditure patterns. Basiq has found that via CDR accuracy is increased by approximately 10 per cent.
    - Additionally, CDR makes critical metadata like interest rates and bank fees for specific accounts and products available. This information is invaluable for comparison services or competing providers looking to offer better terms to consumers.

Basiq considers that removing CDR data accessibility deficiencies (compared with screen scraping) will help maximise the potential of the CDR as a reliable and feasible data sharing framework. Some small but pivotal changes to address these discrepancies will only strengthen CDR’s more robust, structured, and data rich environment.

8. b) Are there particular restrictions related to data use and disclosure under the CDR that influence choices to continue using screen scraping, or vice versa?

1. Ease of User Experience in Screen Scraping:

A notable difference between screen scraping and CDR is the user experience when accessing data holder screens. In screen scraping, the process is simplified to a single screen where consumers provide login credentials. Once submitted, the aggregator can connect to the institution and retrieve the consumer's financial data. This is a smoother process compared to CDR, where consumers are redirected to their bank's portal to authenticate, select accounts, agree to terms, and then get redirected back.

The CDR process can potentially cause consumer drop-offs due to the multiple steps involved. Moreover, once redirected to the data holder's portal, the organisation loses control over the consumer's journey, which could lead to app abandonment, especially if the data holder's user interface is not user-friendly or responsive.

Recommendation: A CDR customer experience (CX) agency or team should continuously evaluate the user interfaces of Data Holders under the Consumer Data Right (CDR) to ensure they comply with the stipulated rules. This evaluation can also yield valuable feedback on usability, which should be shared publicly to motivate Data Holders to enhance the user experience. Alternatively ADRs could be required to track and report on user drop-off rates with various Data Holders and collaborate with those showing high drop-off rates to improve the user engagement.

Basiq has elaborated on consent flow issues in its response to the *CDR Consent Review: CDR rules and data standards design paper*.

2. Ambiguity in CDR Accreditation for Complex Business Structures:
Businesses with complex data sharing arrangements find it challenging to navigate the CDR framework, especially in determining the suitable accreditation without violating CDR rules. Unlike screen scraping, where such constraints don't exist, CDR's accreditation model can be a hurdle for organisations with intricate operational models.

For instance, consider an organisation with franchisees that operate under the parent company's software service and AFSL licence. It remains ambiguous how such organisations should access CDR data, as well as how to delineate the application of security and privacy requirements. Specifically, it is unclear which entity—the parent organisation or individual franchisees—bears the responsibility for implementing these requirements and how they should be executed. This complexity is further heightened when the software service is designed to share data with certain parties under a trusted advisor model, when consumers want direct access to their own data, or when data needs to be shared with third parties. Navigating such multifaceted scenarios is exceedingly challenging due to their complexity.

Recommendation: It would be helpful if relevant government bodies could identify these complex models and release clear guidance on how CDR can support their business operations, outlining accreditation options, their advantages, disadvantages, and obligations.

3. CDR Limitations on Certain Use Cases:
Some use cases, like setting up payment instructions with third-party payment providers, face challenges under CDR due to data sharing restrictions. For instance, if a consumer revokes consent after a payment setup, the payment provider may need to erase the account details, preventing future transactions. Screen scraping doesn't have such constraints, allowing a broader range of use cases.

Recommendation: Clear guidance from the government on addressing and supporting such use cases within the CDR framework would be beneficial.

4. Growing List of CDR Exceptions:
The current list of exceptions within the CDR framework is expanding, and in some instances, critical accounts necessary for particular use cases are not accessible via CDR, but are available through screen scraping.

A notable example is the case of Corporate Trust accounts used by organisations such as Real Estate agents. Due to these exceptions, accounting firms and property management organisations find it difficult to leverage CDR for reconciling rental payments. For instance, National Australia Bank (NAB) does not provide access to these accounts (available via NAB Connect) unless the individual also has an internet banking account. Commonwealth Bank of Australia (CBA) supports some of these account types, but not others, depending on how the organisation's entity was set up in their system.

These examples illustrate that while some accounts are inaccessible via CDR, they are available through screen scraping, indicating a gap in the CDR framework's coverage.

Recommendation: While the rationale behind some accounts being unavailable is understood, there was no industry involvement to ascertain the impact of these exception rules. It's recommended that organisations relying on this data have a pathway to request access to it. Engaging with industry stakeholders to understand the implications and exploring solutions to bridge this gap would be a beneficial step forward.

Business Banking and CDR Challenges
Sharing business banking data is easier through screen scraping than through the Consumer Data Right (CDR) framework. With screen scraping, a business simply provides

its bank login details to a third party, which then collects the needed data from the bank's website.

However, with CDR, the process is much more complex. A business needs to go through a long procedure to allow data sharing. Sometimes, this includes filling out forms, getting signatures from authorised people in the business, and then waiting for approval which could take up to a week. This process is not only time-consuming but also varies from one bank to another, making it very hard for businesses to share data consistently across different banks.

Recommendations:
We recommend the Government ask banks to make this process simpler and to provide easy-to-follow instructions on their websites for businesses wanting to share data. Additionally, it could be helpful if banks automatically allowed data sharing for business users who already have access to their data on the bank's website.

Traffic Threshold NFRs are too low

The current rate limits for CDR endpoints are too low for use cases (e.g. Personal Finance Management or wealth/investment roundups) that require more frequent refreshes of banking data for more consumers. Based on the current 50 request per second only 300-400K connections can be supported for a software product by a Data Holder. This number is not sufficient for any of the large Data Holders who hold data for millions of consumers.

Recommendations:
We have provided detailed requirements on how these thresholds can be raised to match the needs of Data Consumers from the Fintech industry which you can find here.

The outlined points highlight the trade-offs between the streamlined user experience of screen scraping and the more regulated, yet sometimes cumbersome, process under the CDR. Addressing these concerns through clear guidelines, enhanced user experience, and a more inclusive approach to various business models and use cases can help bridge the gap between these two data access methodologies.

8. c) Are there requirements in other regulatory frameworks that affect the viability of CDR as an alternative to screen scraping?

We are not aware of any additional information or considerations beyond those already delineated in the consultation paper.

8. d) Can you provide suggestions on how the CDR framework could be adjusted so that it is a more viable alternative to screen scraping?

Improving Awareness and Education on CDR with Data Holders
In numerous instances, we've noticed that when individuals or organisations reach out to their banks to inquire about open banking under the CDR framework, they are incorrectly informed that the bank does not share financial data with third parties. This misinformation, often disseminated by bank call centres or front-office staff, indicates a significant lack of awareness regarding the CDR framework within these institutions.

Recommendation:
We suggest initiating a comprehensive educational campaign targeted at data holders, especially their customer-facing staff, to bolster awareness and accurate understanding of the CDR framework. This will help in ensuring that inquiries are handled correctly, and the reputation and effectiveness of the CDR program are not undermined due to misinformation.

Addressing Delayed Responses to Tickets Raised with Data Holders
We have come across situations where tickets raised with data holders regarding CDR issues remain unresolved for extended periods, showing a lack of engagement from the data holders' side. It's crucial to remember that behind many of these tickets, there's a consumer awaiting service and a third party eager to provide service, both contingent on the timely resolution of these tickets.

Recommendation:
We recommend establishing a robust Service Level Agreement (SLA) that includes defined severity models to ensure that issues raised are attended to with the requisite urgency. This will help in expediting responses and actions from data holders, thereby enhancing the overall efficiency and consumer satisfaction associated with the CDR framework. By categorising and prioritising tickets based on their severity and impact, data holders can ensure that critical issues are resolved promptly, which in turn, would make CDR a more attractive and viable alternative to screen scraping.

Expand CDR as a Universal Standard Across All Industries

We propose that the Consumer Data Rights (CDR) framework be extended to serve as a national standard in Australia. This would enable all organisations to adopt it as a benchmark for demonstrating secure and consumer-centric data handling practices. Furthermore, this standard should be applicable to all data scraping services, ensuring uniformity in data protection and consumer transparency across industries.

The CDR framework does not explicitly dictate the underlying technology being employed, whether it's via an official API or programmatically accessed via screen scraping services. This flexibility suggests that CDR principles could indeed be applied to screen scraping operations, thereby ensuring a structured and regulated approach towards the collection and usage of consumers' financial data.

In anticipation of such a regulatory evolution, our organisation, Basiq, has proactively aligned its operations with CDR principles, particularly concerning consent and data governance. We ensure that all screen scraping services necessitate a consent akin to CDR guidelines, and we employ the same rigorous data governance controls as required for CDR data handling.

We have previously highlighted to the Government that while the CDR framework is exemplary in establishing controls on data handling, its reach is limited to organisations voluntarily participating in CDR data collection. We propose a consideration towards making CDR a standard framework within Australia, allowing organisations outside of CDR designated sectors to actively seek accreditation.

Achieving such accreditation would enable organisations to showcase their adherence to robust data management controls, filling a critical gap. While certifications like ISO and SOC address security and procedural aspects, they fall short on specific guidelines concerning consumer consent and data handling. Transitioning CDR into a standard framework would bridge this gap, providing clear guidelines on consumer data handling alongside ensuring security and procedural integrity.

This proposal not only elevates the data protection standards but also aligns with the growing consumer expectation for transparency and control over their data, regardless of the technological methods employed for data collection and processing.

The Statutory Review recommended that screen scraping should be banned in the near future in sectors where the CDR is a viable alternative.

9. a) How should the Government determine if the CDR is a viable alternative?

To gauge the feasibility of the Consumer Data Right (CDR) as a substitute for screen scraping, the government could take the following pragmatic steps:

1. Adoption Rate: Analyse the uptake of CDR among businesses and consumers in comparison to screen scraping. Data aggregators, which offer both scraping and open banking services, can provide insights into this.

2. Successful Completion Rate: Measure and evaluate the consumer through-put rate in relation to the CDR (i.e. the percentage of consumers who successfully share their data through CDR). This will help in understanding the consumer's ease and willingness to use CDR.

3. Ease of Implementation: Compare the technical ease and resource burdens in connection with implementing and maintaining CDR accessibility against screen scraping. This evaluation could include analysis of the resources, time and technical expertise required to implement and maintain each system. Understanding friction points in relation to CDR implementation and maintenance will help understand the feasibility for organisations to transition from screen scraping to CDR.

4. Data Completeness: Assess whether CDR provides all the necessary data required by various sectors / businesses (in comparison to screen scraping) to be able to provide their underlying service / product. This can be measured by examining the types and amount of data that can be accessed through both methods (e.g. gap analysis).

5. Engagement with Stakeholders: Conduct discussions with key industry stakeholders to gather their preferences and the challenges they face with CDR as compared to screen scraping. This engagement can provide first-hand insights into the practical challenges and benefits of CDR from the perspective of those who would use it.

9. b) What are your views on a ban on screen scraping where the CDR is a viable alternative?

General Statement:

We believe that the Consumer Data Right (CDR) framework is a viable alternative to screen scraping services. The CDR provides modern protection for consumers regarding how their data is handled and offers a more reliable method for data acquisition compared to screen scraping.

We appreciate the Government's efforts to further improve the CDR through various consultations, such as consent improvements, operational enhancements, and addressing issues concerning data sharing for business entities.

However, we highly recommend that the suggestions outlined within this document are taken into serious consideration, as they are crucial for ensuring businesses support the CDR and are satisfied with the outcome. We believe that with the initiatives that are already underway, addressing the items stated here will be important for ensuring the future success of CDR.

Furthermore, we also recommend that after evaluating the submissions for this paper, the Government engages with organisations that have a long history of using scraping services (and have a large volume of customers on their platform using it), as well as with the data aggregators. These organisations can help explain the challenges and share their satisfaction levels, while the aggregators are key to ensuring that these organisations can transition over to the CDR. This engagement will promote a better understanding and collaboration, ensuring a smooth transition and establishing a solid foundation for the long-term success of the CDR.

Challenges for organisations transitioning from scraping:

We have identified several challenges and considerations that organisations might face while transitioning from scraping to the Consumer Data Right (CDR) framework. It's encouraging to note that many of these challenges are already under review for improvement within the program.

- **Improve ADR User Experience (UX):** The present consent mechanism has certain UX challenges, as highlighted in the recent consent review. Our recommendations in response to the design paper 321 could bring about significant improvements. We have detailed some additional considerations in this paper for further exploration.

- **Improve Data Holder UX:** Although the consent review didn't cover Data Holder UX improvements, these are crucial and need attention. Recommendations to tackle this challenge have been provided in this paper.

- **Consumer's ability to select which accounts to share:** This change could impact organisations in lending sectors, as they are accustomed to having full visibility of consumer accounts for responsible lending assessments. Despite the change, we believe organisations can adapt with some operational fine-tuning. Ideally, we recommend eliminating the concept of account selection by consumers, as it's not intuitive without an understanding of the underlying technology. This practice could potentially derail use cases—for instance, a consumer sharing only a term deposit account while applying for a loan, overlooking credit card and keycard accounts.

- **12-Month Consent Limitation:** The consent period has a hard stop after 12 months. For businesses with long-term customer relationships, this necessitates re-engaging with consumers for re-consent, causing potential drop-offs and friction. For example, an automated charity roundup would halt after 12 months, requiring the business to seek re-consent. We find this restriction on companies using the Representative access model unfair and advocate for its removal, given that the consents already define the scope and purpose, with data minimisation measures in place.

- **Quarterly Notification Reminder:** Although addressed in the consent review, we emphasise a need for more flexibility with the quarterly notification reminder. If a consumer is actively using the service, the notification requirement should be waived, reserving notifications only for passive users.

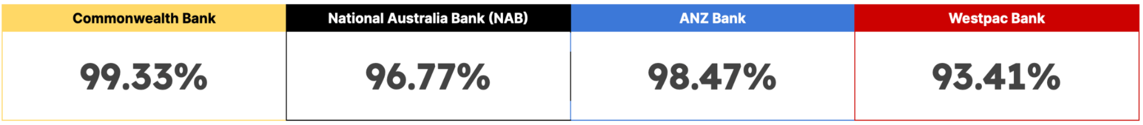
- **Accreditation Models:** We recommend the need for operational enhancement and more information on accreditation models to support complex businesses and use cases. We echo this sentiment and seek further support from regulatory bodies to understand how diverse business models and use cases can be accommodated under the CDR framework.

- **Account Availability:** Some banks do not provide access to certain account types under CDR, unlike scraping — for example, Corporate Trust accounts. This limitation could pose challenges for organisations requiring access to a broad spectrum of account types.

Benefits for organisations transitioning from scraping:

Here are several compelling reasons why organisations currently using scraping are inclined towards transitioning to the Consumer Data Right (CDR) as opposed to continuing with scraping:

- **Reliability:** Via CDR, the chance of an API successfully returning data is higher, making it more reliable than scraping. On average, Basiq experiences a 5% to 10% increase in successful data retrieval attempts via CDR and the successful attempt rate has never fallen under 91%. With Screen Scraping the successful attempt rate in a given day could go as low as 40%. Moreover, while scraping services may not be available for some institutions (such as Macquarie, Up Bank and HSBC), CDR is accessible, presenting a considerable advantage. This enhanced reliability is a significant draw for businesses towards CDR.



Commonwealth Bank	National Australia Bank (NAB)	ANZ Bank	Westpac Bank
99.33%	96.77%	98.47%	93.41%

This report above provides a sample of the number of successful CDR data retrieval "jobs" completed by the Basiq platform with the top four major banks between 2023-09-01 to 2023-09-21. Note the Consumer Data Right (CDR) framework mandates that institutions self-publish statistics related to error rates and performance, which can be accessed via the link below. Please note that these results may differ from those on the Basiq platform due to variations in measurement methods. See link: https://www.cdr.gov.au/performance

- **Speed and Performance:** Speed is crucial in reducing onboarding time for customers, which in turn can enhance the user experience significantly. The prompt data retrieval through CDR facilitates a smoother onboarding process and a better overall application experience for the consumer.

- **Richer Data Attributes:** CDR offers a richer set of data attributes and entities compared to scraping. This wealth of data can offer deeper insights to businesses, aiding in personalised service offerings and uncovering new use case scenarios. For instance, accessing detailed account information like interest rates can empower a lender to tailor more competitive loan offers.

- **Direct Engagement with Data Holders:** One of the significant challenges associated with data scraping is the potential for disruptions in data access. When such disruptions happen, there is often a lack of established communication channels or infrastructure to support the continuity of operations. In such scenarios, the burden falls on the aggregator to swiftly resolve connectivity issues. CDR provides a clear solution to this problem by incorporating data holders within the ecosystem, allowing Accredited Data Recipients (ADRs) to register and log tickets when issues arise. This proactive approach ensures higher uptime for businesses, enabling them to better serve their customers and operate efficiently.

Addressing Data Quality Issues in Consumer Data Rights (CDR)

It's important to clarify misconceptions surrounding data quality issues in CDR and Open Banking. The quality of banking data, especially transaction data, inherently presents challenges. This is because the data originates from payment infrastructure that was not initially designed for CDR or Open Banking purposes. Many of the concerns raised stem from those who haven't had direct experience working with banking data.

We believe that it falls upon Accredited Data Recipients (ADRs) or data aggregators to enhance the quality of this data. For instance, at Basiq, we allocate substantial resources to effectively categorise transaction data, enabling us to offer deeper insights into the nature of each expenditure.

Another aspect of data quality issues arises when data holders inadvertently send incorrect, invalid, or incomplete data. It's important to note that these issues are generally not malicious but rather a consequence of the complex systems banks operate. In some cases, duplicated account data may exist across multiple systems, with varying update frequencies. When banks implemented APIs, they sometimes made arbitrary decisions about which tables or systems to source data from. However, such instances have been limited, and data holders promptly address and resolve any identified issues.

Lastly, organisations that have developed their own analytical and data enrichment algorithms, often based on scraping data, may encounter variances when processing the same CDR data. This discrepancy arises because scraping provides data visible to consumers, whereas CDR data may originate from different tables and systems that are more raw in nature. While these differences exist, they are not inherently problematic but merely reflect distinctions in data sources. Organisations working with CDR data have the need to familiarise themselves with these nuances to adapt their systems accordingly.

9. c) What timeframe would be required for an industry transition away from screen scraping and why?

Given that the CDR is already in place and operational, the transition away from screen scraping can be expedited compared to starting from scratch. However, addressing the earlier mentioned challenges is crucial for a smooth transition. Here’s an updated outline considering the existing CDR framework:

1. Immediate Actions (0-6 Months):

- Engage with stakeholders that rely on scraping services and with data aggregators that provide scraping capabilities to gather feedback on the current CDR framework and identify specific challenges and areas of improvement.

- Establish a task force to address the identified challenges, comprising government officials, industry stakeholders, and technical experts.

- Begin addressing technical and regulatory challenges with the existing CDR framework based on feedback and assessments.

2. Short-term Actions (6-12 Months):

- Implement the necessary changes to the CDR framework to address identified challenges.

- Conduct targeted campaigns to educate stakeholders on the improvements and encourage the shift away from screen scraping.

- Work with data aggregators to develop and disseminate clear guidelines for businesses on how to transition from screen scraping to CDR.

3. Medium-term Actions (12-24 Months):

- Monitor the adoption of CDR and collect feedback for further improvement.

- Continue engagement with industry stakeholders to ensure smooth transition and address arising challenges promptly.

- Evaluate the progress and impact of the transition on different sectors and adjust strategies as necessary to ensure the objectives are met.

4. Long-term Actions (24+ Months):

- Maintain a continuous improvement process for the CDR framework based on industry feedback and technological advancements.

- Continue monitoring compliance and engagement with stakeholders to ensure the long-term success of the CDR as a viable alternative to screen scraping.

With the existing CDR framework, a complete transition away from screen scraping could be envisaged within a 12 to 24-month timeframe, provided that the identified challenges are promptly addressed and there's active engagement with industry stakeholders to facilitate the transition.