FINANCIAL DATA AGGREGATION

Paving the Way for Smoother Mortgage Transactions

FORMFREE®

EXECUTIVE SUMMARY

This white paper provides lending executives with an overview of the current uses and limitations of direct-source financial data in the mortgage space and highlights opportunities for advancement in the field.

Mortgage lenders aggregate data from consumer financial accounts to verify loan applicants' ability to repay. Already, financial data aggregation is paving the way for smoother mortgage transactions, but challenges presented by current account authorization and data retrieval methods are still creating bumps in the road.

The most widely used method of authorization, credential-based access, has many potential points of failure, including user input errors and financial institution security measures that inadvertently block legitimate requests for account access. The three core methods of data retrieval — Open Financial Exchange (OFX), screen scraping and direct-source API access — each have weaknesses that can disrupt the collection and integrity of datasets.

Despite these challenges, data aggregation already is enabling faster, more fraud-resistant loan transactions, and advancements in data aggregation technology and standardization will create even more exciting opportunities for mortgage lenders.

WHAT IS FINANCIAL DATA AGGREGATION?

Data aggregation is the process of gathering data and presenting it in a summarized format for data analysis. Since the accuracy of insights from data analysis depends heavily on the amount and quality of data used, aggregation efforts typically involve the retrieval of large quantities of data directly from one or more authoritative sources.

Data aggregation is useful in many business contexts. Marketers use data aggregation to better understand their audiences and shape marketing campaigns. Product strategists use data aggregation to optimize pricing. And in mortgage lending, aggregation of the financial data required to make underwriting decisions is an essential ingredient in the digitally driven credit decisioning process.

Lenders can use aggregated information from consumer bank, retirement and brokerage accounts to verify applicants' assets, employment and income and gain a holistic view of their overall cash flow and creditworthiness — in essence, the consumer's financial DNA. When compared to manually collecting and reviewing information from printed or digital bank statements, the combination of direct-source data aggregation and straight-through processing provides greater purchase certainty and yields a more accurate assessment of applicants' ability to pay while minimizing delays, errors and fraud risk.

Consumers are increasingly willing to grant lenders permission to grab their transaction data directly from financial institutions; however, imperfections in the data aggregation process have led most lenders to continue supporting document-driven methods of financial data collection alongside more streamlined approaches. Until these potholes are paved over, mortgage lenders will not be able to realize the efficiencies of requiring all applicants to share their financial data.

To unpack the pain points in financial data aggregation, one must first understand its two component processes: authorization and retrieval.

AUTHORIZATION

Authorization is the process by which loan applicants grant data aggregators permission to retrieve their account, balance and transaction information from a financial institution.

Credential-based access (a username and password) is the most widely used method of authorization. With this method, the consumer selects their financial institution from a list and supplies the username and password associated with their account. The data aggregator then securely presents the credentials to the financial institution to gain account access.

The potential failure points of this method are many fold. Today, nearly eight in ten Americans bank online — but that leaves 20% who do not have online banking access. Those who do may not provide the correct username/password combination, or they may provide the right username/password combination but specify the wrong financial institution.

Another potential issue occurs when the financial institution identifies a mismatch between the data aggregator's IP address and the IP address typically associated with the consumer. When this happens, the consumer may be asked to submit to a multi-factor authorization protocol such as answering security questions or supplying a one-time code. But consumers aren't always able to correctly answer security questions, and one-time passwords sent via phone or email may not reach the consumer if their contact info has changed since opening the account.

A newer authorization model called **token-based access** sidesteps many of the challenges presented by credential-based access. Modeled on open standards like OAuth (used by Amazon, Google, Facebook, Microsoft and Twitter), token-based access lets consumers grant aggregators access to their financial data without sharing their passwords with the aggregators. With token-based access, control of the consumer experience is transferred to the financial institution for first-party consumer authentication. After successful login, the consumer consents to third-party data sharing and control is passed back to the data aggregator.

This method avoids IP mismatch issues, since credentials are entered directly by the consumer from a known device. It also gives consumers greater self-determination over the sharing of information, as consumers must explicitly consent to sharing information with data aggregators and can revoke that consent at any time.

DATA RETRIEVAL

Data aggregators rely on several methods to retrieve bank data. The oldest of these is the **Open Financial Exchange (OFX),** a data-stream format for exchanging financial information that's been around since the earliest days of personal financial management software in the 1980s. OFX is no longer widely used, in part because it lacks the data richness of newer methods.

The most prevalent method of data retrieval is what's called **screen scraping** or HTML harvesting. Data aggregators use computer scripts to navigate through an online banking site much as a human would, accessing relevant data and sending it back to the aggregator for processing. Screen scraping offers superior richness of data, as a sufficiently sophisticated script can access any information that would be available to the consumer. But it's also the most fragile method, as any changes to the financial institution's user interface can "break" the screen scraper. Screen scraping is also resource-intensive from a server capacity standpoint, driving up server load without any meaningful way to differentiate data aggregator activity from actual logged-in users.

The newest method of data retrieval relies on **direct-source API** access. These APIs, or application programming interfaces, allow data aggregators to download consumer data via API payload. Because they are built and controlled by individual financial institutions, each API connection has different data fields, delivers different fidelity of data and offers a different depth of transaction history. The unpredictable quality and completeness of data retrieved via direct-source API remains the greatest challenge to this approach.

THE ROAD AHEAD

Already, data aggregation is enabling faster, more fraud-resistant loan transactions. Access to financial data happens at digital speed and eliminates the traditional paper chase, and lenders can quickly and easily refresh data without burdening consumers if a second verification is needed between loan qualification and closing. Yet even more exciting are the opportunities that lie ahead in the data aggregation space.

- ### MORE DATA FOR A FULLER FINANCIAL PICTURE

    Historically, financial data repositories have been designed with a focus on recent data — i.e., transactions completed within the last 30 days. Data aggregators are now working with financial institutions to establish real-time access to historical data going back 12 months or more. With years' worth of transaction data, creditors will be able to make much more deterministic assessments of borrowers' creditworthiness.

- ### GREATER CONSUMER SELF-DETERMINATION

    Some data aggregators have been thwarted by financial institutions' efforts to obstruct access to consumer financial data. Although often done in the name of protecting consumer privacy, data obstruction actually undermines a bank's relationship with its customers. That's because banks don't own a consumer's financial data — they are merely custodians of it. Just as financial institutions have a responsibility to protect consumer data, they also have a responsibility to share data with third parties upon a consumer's explicit request. By educating consumers on their ownership of their own financial data, we can democratize access to consumer-permission data.

- ### GLOBAL STANDARDIZATION

    Groups like the Financial Data Exchange (FDX) are working to establish global, industry-wide standards for the sharing of consumer financial data. If achieved, this data standardization would establish baseline minimums for API data payloads and greatly improve the quality and consistency of aggregated data. Such standardization would also pave the way for more widespread use of technologies like federated identification, which allows a single electronic identity to be used across multiple financial institutions.

There is no doubt that data aggregation will continue to play an increasingly prominent role in the way not just mortgage lenders, but all creditors, assess consumers' ability repay. Once the "bumps in the road" have been paved over, data aggregation will empower lenders and their investors to understand consumers' financial DNA at electronic speed and with unprecedented levels of confidence. Ultimately, data aggregation methods that employ direct-source data extraction and straight-through processing will yield an improved lending experience and cost savings for lenders and consumers alike.

ABOUT FORMFREE

FormFree® is a market-leading fintech company whose revolutionary products AccountChek® and Passport® make for a more inclusive credit decisioning landscape by enabling lenders to understand people's true ability to pay (ATP®). To date, thousands of U.S. lenders and brokers have ordered millions of FormFree's patented verification reports representing trillions of dollars in loan verifications. FormFree delights borrowers and lenders with a paperless experience, dramatically reduces origination timelines and offers automated analysis and standardized delivery to lenders and investors using a secure ReIssueKey®. For more information, visit **https://www.formfree.com** or **follow FormFree on LinkedIn**.