



AN INFLUXDATA CASE STUDY

How Strive Uses InfluxDB to Feed Site Artifacts into a Machine Learning System

Gary Tomlinson

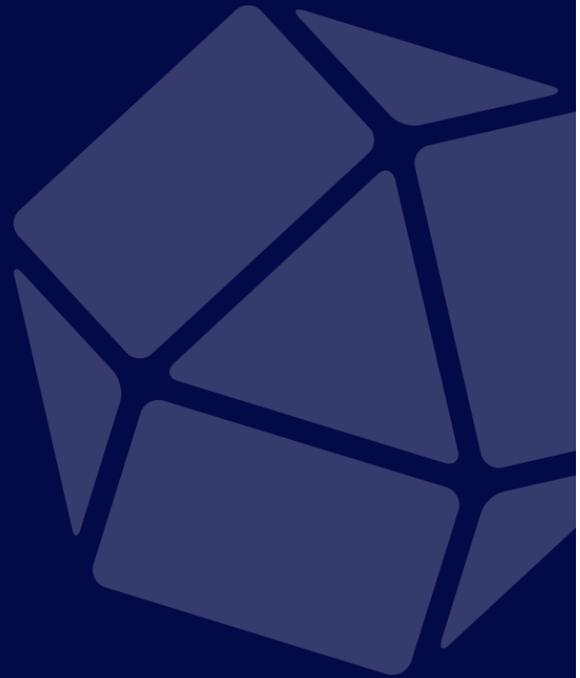
Director of R&D, Strive

Katherine Chavez

Director of Marketing, Strive



SEPTEMBER 2018



Company in brief

Founded in Seattle, Washington in 2014, Strivve (formerly Switch) is the first credit card updater for online accounts. It makes online purchasing faster, safer, and easier by enabling users to manage and update all their online accounts from a single place. Strivve leverages machine learning to automate common tasks across tens of thousands of the most popular websites.

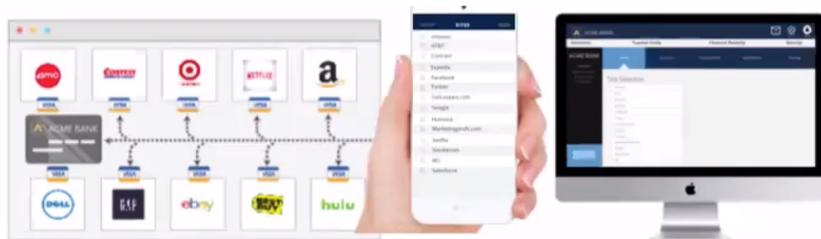
Strivve designs and builds Robotic Process Automation (RPA) solutions capable of: user-assisted placement of payment cards on site; automatic digital provisioning of new and replacement payment cards on sites; updating billing and address information on sites; and determining what payment cards are on a site.

Strivve's secure, automation technology helps issuers capture and increase card revenue by keeping cards Top of Wallet®. The company's secure, proprietary solutions add value to online purchasing processes for card issuers, merchants and consumers alike. Strivve's customers are financial institutions and issuers, and its potential partners are financial technology providers.

Case overview

Strivve wanted to build a platform that helps issuers get their cards into use immediately, to update online payment profiles, and activate new or replacement cards on cardholders' behalf. The task of managing cards to site relationships, however, is complex. To build their credit card updater, Strivve realized that they needed anonymized site navigation for machine learning to enable using remote process automation (RPA).

Strivve chose InfluxDB Cloud, the hosted version of InfluxDB, as the time series database for their platform. Using best-in-class security protocols, Strivve's platform called TopWallet™, also available in the form of an API called CardSavr™, discovers where cards are used for payments, navigates to login pages, and adds new or updated payment cards on behalf of each user. This unique functionality is enabled by streaming anonymized crowdsourced site artifacts into InfluxDB for learning, analysis and support. Strivve's technology allows issuers to capture lost revenue and gain visibility into usage of newly issued cards, as well as enhancing convenience and security for cardholders.



Strive using InfluxDB in TopWallet™ platform and CardSavr API to increase card revenue for issuers and save time for cardholders

“We all get our credit cards replaced a lot. Our online ecosystem is getting more complex with the number of accounts we have. We buy our groceries, clothes, almost everything that we have and our utilities online, often with credit cards. We wanted to simplify that, and make it less complicated to get those cards back on file.”

Katherine Chavez, Director of Marketing

The business problem

Strive wanted to create one-of-a-kind cardholder experiences that inspire loyalty to recapture and increase card usage and revenue for card issuers by automating the process of online credit card updating. Each year, over 700 million new payment cards are issued, forcing cardholders to update their cards online, and \$195 billion in transaction fees and interest payments are generated by card issuers annually. In 2017 alone, over one billion credit cards were reissued.



As cardholders, we all have to replace our card on every single site if we get a new card: to activate it then visit multiple sites and update it on each. Further, widespread fraud is requiring frequent credit cards replacement and resulting in a more complex online ecosystem consisting of multiple accounts per individual. Yet updating a card everywhere you pay is a painstaking, time-consuming process. Cardholders are experiencing these frustrations by the hundreds or thousands every day:

- Card issuers' brand becomes associated with this frustration
- Customer satisfaction declines
- Card revenue is negatively affected

When reissuing events occur and replacement cards are active but not used, significant revenue is lost by card issuers and merchants. With the growing number of cards in every wallet, online merchants are storing on file.

The greatest fear of any card issuer is a fraud event, rendering the card useless and requiring replacement. Almost everyone has more than one card from more than one issuer and tends to use the other card if the first one is compromised. It is critical for card issuers to get their cards out there faster, easier, and keep that loyalty. Merchants have revenue in the order of hundreds of millions of dollars per year on these cards. A card issuer's goal is not to know what sites their cardholders are using but to get their card on those sites.

Card issuers should help cardholders update their cards online to:

1. Increase revenue by gaining additional transaction fees and interest payments
2. Keep their card Top of Wallet: cardholders choose that card first when purchasing online
3. Enable cardholders to quickly and easily update websites with their new card or switch from competitor cards to theirs

Building the Strivve platform also involved entering the consumer space, which required brand awareness and consumer trust, especially in the context of recent security leaks suffered by even the largest corporation websites. Besides the need to store credit card information, Strivve needed to store login credentials on all sites used by each cardholder. This led it to begin partnering with card issuers and building relationships with major merchants to establish trust in its technology.

The technical problem

Early on in their technical journey, Strivve tried to identify its technical needs. At first, they attempted to gather as much data as possible. They had a lot of unstructured data and were using classic mining techniques but soon realized that this was not the right approach.

They started to look at what data would be useful out of everything that they were streaming and narrowed down the massive volume to only data of interest:

1. Anonymized site navigation for machine learning: They didn't want to have any personally identifiable information but wanted to know what sites their users are visiting so that they can help them through their remote process automation (RPA).
2. Anonymized RPA interaction error data for troubleshooting: They also realized that it would be useful to have error handling data for troubleshooting, whereby the service provider gets errors snapshotted to help resolve them.

Strivve first chose Elasticsearch due to its popularity, put a Logstash server up into Linux, and created an EC2 instance on AWS. Their technical journey progressed as follows:

1. Started by streaming everything from their experimental TopWallet application - they called that Massive Event Logger - MEL.
2. Analyzed MEL data - As they started to analyze the data, they realized they need to stream only the necessary data (anonymized site navigation for machine learning and anonymized RPA interaction error data for troubleshooting), so they reduced the data volume accordingly.
3. Performed load testing - they ran load testing on Logstash based on projected user estimates and quickly crushed their first Logstash instance.

Lessons learned from using Logstash were that it:

- Was difficult to get the machine learning and troubleshooting artifacts out of Logstash
- Wouldn't scale for Strivve's needs
- Didn't fit their strategic model of leveraging non-intellectual property hosted/cloud services.

As a small team, Strivve had to be creative and efficient in gathering data and extracting value from it. They decided to invest their energy into their core value proposition, which consisted in machine learning and remote process automation – not building a large Elasticsearch infrastructure on Amazon. That put an end to using Elasticsearch and set them on a search for an alternative.

The solution

“We saw some of the other folks using InfluxDB, and some of the scale points they were using it at. Could we generate a big enough load to bother Influx? It was far beyond what we were doing with Logstash at that time...We knew it would work for what we were doing.”

Gary Tomlinson, Director of R&D

Why InfluxDB?

Having been aware of InfluxDB, Strivve decided to test it against Logstash at the same transaction rates that had pushed Logstash over. InfluxDB easily handled that load, so Strivve converted the test to a production system and today have both a production and a test system.

Strivve chose InfluxDB because it is:

- Purpose-built for time series data (which facilitates the ingestion of site navigation events)
- Scalable (able to handle massive workloads)
- Tag system based (which provided context for data)
- Available in a cloud version (which fit Strivve’s strategy to leverage the cloud)
- Widely adopted by the community (numerous customer use cases to learn from)

Strivve modified their machine learning system to ingest time-based site navigation events stored in InfluxDB. And since InfluxDB is time-based to begin with, it was easy to use InfluxQL to query the data on time for analysis in their machine learning system. This allowed them to structure their data nicely so when problems arise, they are able to navigate the records to determine cause during any specified time period. The InfluxData platform is well-suited for Machine Learning because it can easily send time series data to an ML solution for training. The trained model can then be brought back into InfluxDB to provide real-time dashboarding as well as trigger the appropriate workflow.

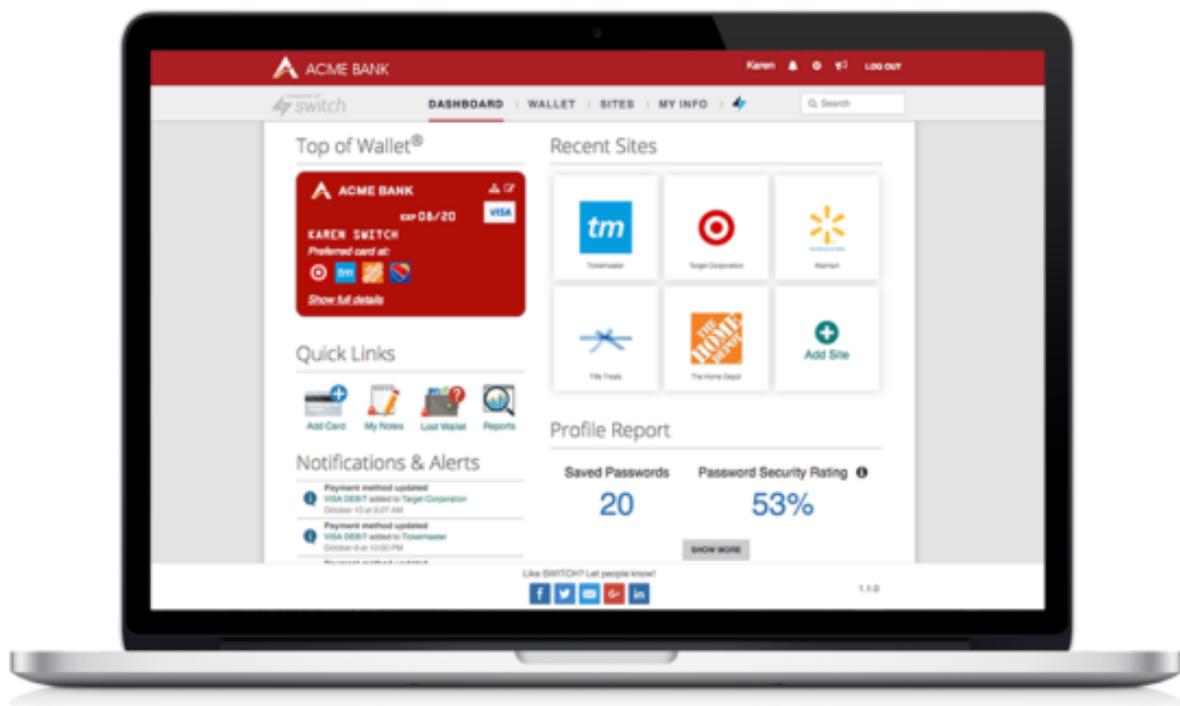
Strivve modified TopWallet app to stream anonymized navigation events and capture clickstream of interesting forms – records of navigation route from hovering to clicking to transitioning across form fields. For troubleshooting, they also capture screenshots to have a record of errors encountered, then tag and index them.

Strivve products

Strivve has developed two products, TopWallet platform and CardSavr API.

TopWallet

TopWallet is a consumer beta product that can be accessed on the company's website at switchme.com. It's Strivve's full-featured, card-issuer-branded web-based Chrome extension and mobile app for cardholders. It gives access for users to get all the benefits of Strivve's platform. TopWallet apps promote the card issuer brand to help cardholders manage their entire payments and password ecosystem from a single place.



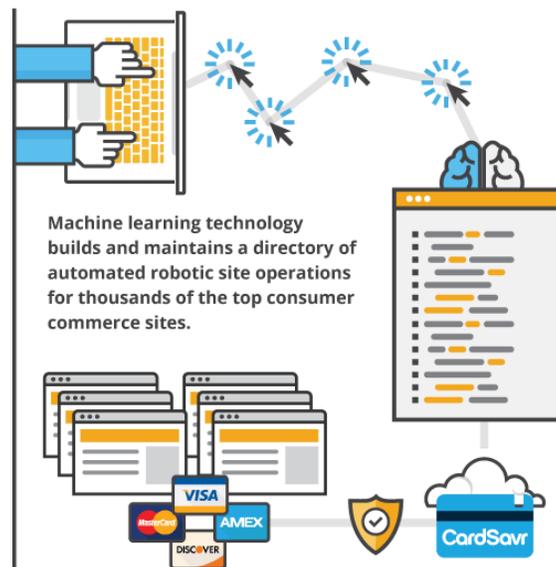
Upon interest from its financial institution customers, Strivve has packaged TopWallet not only with an API but with an SDK, whereby some TopWallet capabilities can be built into financial institution apps. Strivve are working with some of the large card issuers in the United States to integrate their systems, and CardSavr is an outgrowth of that.

Upon interest from large merchants, Strivve also built TopWallet capabilities into merchant apps, especially since some of the larger merchants in the U.S. have their own co-branded cards and since there is a large number of card switches occurring between card brands.

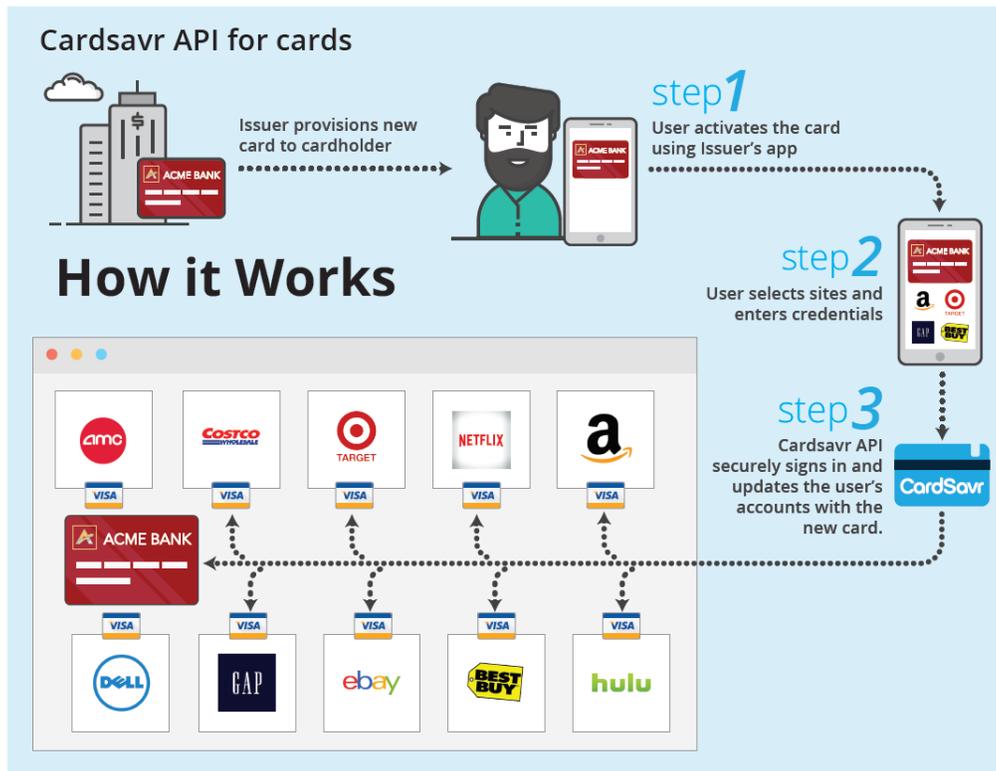
CardSavr™

CardSavr is a card-on-file API built to help cardholders seamlessly update their new or reissued cards anywhere they pay online. Strivve use machine learning technology to build a crowdsourced database of automated robotic processes that can be performed on existing merchant sites.

CardSavr is a cardholder service integrated into a card issuer's branded applications. Via the CardSavr API, existing card-branded applications can embed Strivve card updating technology for 1000's of sites, into their existing mobile and web banking applications while maintaining 100% control of the user experience. The Strivve platform already supports thousands of online merchants without individual opt-in integrations by using anonymized crowdsourcing techniques coupled with a machine learning engine.



CardSavr does not require any opt-in integration with individual merchants. Virtual browser technology performs the card updates on the user's merchant site accounts and saves it on their behalf. The entire process of updating payment information online is carried out in seconds. When a cardholder updates their card's information in Strivve, Strivve logs into each site that uses the card and updates the information automatically.



Additional Features

Strive does more than update credit cards:

- **Autofill:** Strive can fill in the user's name, address, phone numbers, and email address – even credit card numbers (sparing users repetitive, error-prone typing)
- **Login Management:** Auto-saves usernames and passwords as users visit sites, so that they can log in to them with one click, and also generates new secure passwords.
- **Free to Use:** The app is free for an issuer's cardholders, and issuers pay a monthly fee based on the number of cardholders actively using Strive.

Best-in-Class Security

Strive safeguards user data via several security measures:

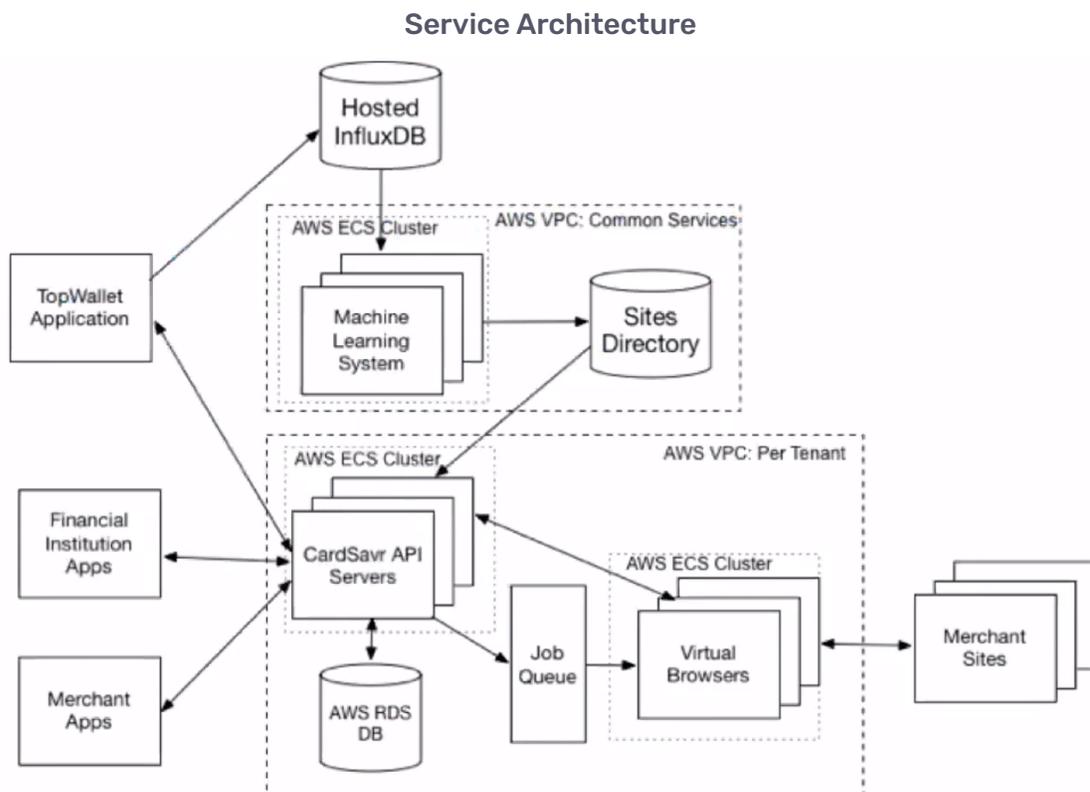
- Master passwords meet industry leading strong password requirements, and are never stored.
- End-to-end encryption means no one can access a user's except the user.
- StriveSafe™ technology provides an additional level of encryption to protect a user's most sensitive information

- Two-factor authentication provides users additional control of their account and more peace of mind.
- Strivve allows users to control additional security features like device tracking, inactivity timeouts, single session login, sign-in failure lockout, and security alerts.

Technical architecture

“The trick in machine learning is to gain as much data as you can. But instead of just trying to sift through massive amounts of unstructured data looking for a nugget, we have a lot of context running with our app. That helps us a lot, paring it down so it's a manageable amount of data. InfluxDB makes it pretty easy with the structure that we can apply.”

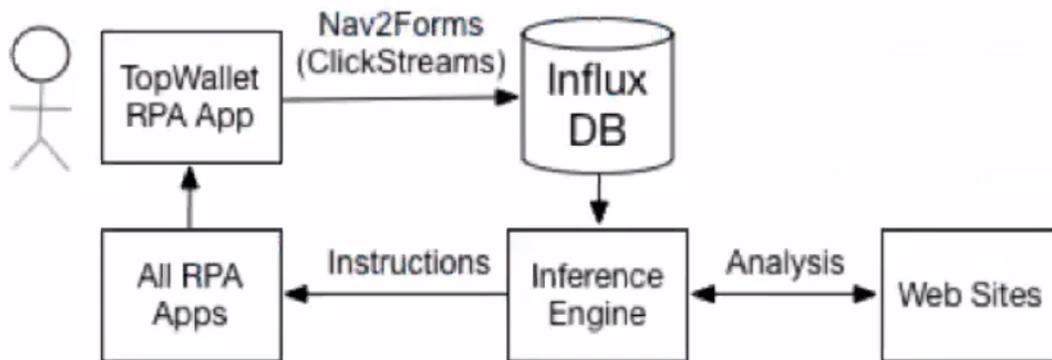
Gary Tomlinson, Director of R&D



Early on, Strivve aimed to learn which sites their users visit. They realized that if they anonymize their data, they could crowdsource site artifacts and learn about visited sites that they have commercial relationships with:

- When users on the TopWallet application visit sites, Strivve tracks that activity and starts to stream that event information into InfluxDB Cloud.
- Strivve’s machine learning system analyzes the data that’s coming in from InfluxDB. They learn about the systems, push that into a Sites Directory, which then feeds Strivve’s main service.
- The CardSavr service itself runs behind the scenes.
- There is an AWS per tenant: one for TopWallet itself, where Strivve manage the service for end users while financial institutions or merchants have their own virtual private cloud tenants. So they’re logically all isolated from each other to maintain PCI compliance.
- All the apps – the TopWallet App, financial institution apps, and merchant apps – communicate with the CardSavr backend, which then has virtual browsers in support of financial apps and merchant apps. In TopWallet, there’s no virtual browser necessary because the browsing application is built into the TopWallet app.
- Then the Remote Process Automation, which runs in the virtual or real browser, interfaces with merchant sites on behalf of an end user. So they can log in, impersonate users, navigate forms, update billing information and card information, and replace a card.

Machine Learning Flow



This machine learning flow is as follows:

Crowdsource anonymized navigation to forms of interest:

- The TopWallet RPA App tracks navigations to forms of interest, in the form of clickstreams.
- Strivve learned very quickly that they could not rely only on a URL to track navigation and that a URL often has to be used in conjunction with clickstreams, which is first driven by users (whose identity is anonymized).

Analyze forms using navigation information:

- Trigger conditions, to determine whether a given site is a site of interest, include how frequently it is visited and if visits are region-based (such as for utility companies). This is especially important to some of Strivve's financial institution customers and where the crowdsourced information becomes highly useful.
- Crowdsourced information is also useful with regard to large sites because such sites are constantly changing. Strivve re-scans them to pick up clickstream events and relearn very quickly. All that data goes into InfluxDB and is then parsed down into an inference engine.

Produce interaction instructions for Remote Process Agents (RPA):

- Strivve's Harvester grabs that data, analyzes the merchant sites, then generates inferencing information that Strivve can send back to feed their RPA agents (RPA apps in TopWallet, or their own virtual browsers in CardSavr). The machine learning system simulates how humans process information: We learn, we click, and we recognize paradigms. The larger the learning sets from crowdsourcing, the greater the inferencing.
- A timestamp is automatically added to the data, and the data is sent over clickstream for current form. As users navigate a site, the JavaScript knows where they are on the site.

Feed RPA's:

- Inside the schema, Strivve have record types distinguishing between urls, clickstreams, and forms. When an element is unidentified in real time, the Harvester grabs it to perform a much deeper analysis and identify its type.
- Since available data available is often not clearly tagged, Strivve send in information about geometric and spatial relationships between elements, which supports deeper analysis in their inferencing engine. If an automatic tag can't be given to some elements, they occasionally have a human tag it and send that to their information system to recognize it as part of a pattern. When machine learning recognizes a new element as belonging to a previous group, this helps establish confidence intervals. The larger the data sets in any machine learning system, the more accurate they become over time. This is why Strivve crowdsource site artifacts: in order to set confidence thresholds.
- CardSavr API takes the streamed data and then uses the InfluxDB library to store it in.
- Through crowdsourcing, Strivve have mechanisms to feed the sites that are obvious up front. Since they can't always get all sites crowdsourced, they also use Amazon's Alexa site ranking to gain insight into the biggest sites and dredge them using the Harvester.

What's next for Strivve?

As Strivve grow, they plan to utilize InfluxDB and Kapacitor for TopWallet and Cardsavr operational support systems:

- Augment AWS tools for capacity planning
- Capture failures (lack of coherence between their machine learning system and actual site data streamed) and process them through Kapacitor to trigger on and analyze them
- Analyze infrastructure-related problems and issues such as having their autoscaling at the right policies in Amazon

Results

“We modified our machine learning system to ingest site navigation events. And so this was a lot easier to do. So that was one of the nice things we realized with the Influx system is that, being time-series-based, it became super easy.”

Gary Tomlinson, Director of R&D

Strivve’s automation technology provides visibility to track and increase usage of newly-issued cards. It increases revenue by empowering cardholders with the fastest and easiest way to securely update credit and debit cards across tens of thousands of merchant sites. Card issuers want to know where they rank in TopWallet and who they are competing against – something they have absolutely no other way in the industry to gain visibility into. The settlement statements that card issuers receive are often unhelpful since merchants are somewhat obfuscated. Strivve has machinery that can dredge through these settlement statements and uses machine learning to map statements to their site’s directory to give issuers more clarity.

Through Strivve, issuers can learn how often cardholders use their card compared to their other payment cards and which sites a cardholder uses or does not use their card on. Strivve is delivering multi-faceted value for card issuers:

Gain insight:

- Immediate validation of card use
- Know where cardholders store their card online

Recapture and increase revenue:

- Cards at Top of Wallet
- Interchange revenue
- Interest Income
- Customer Lifetime Value

Increase new card activations:

- Make it simple for cardholders to use their cards online and increase activations

Increase circulation:

- Reissued cards are back in use faster
- Cardholders easily keep cards up-to-date when replaced

Issuers can also leverage Strivve to show cardholders offers that promote use of their payment card as the default payment card on additional sites. Cardholders can see where an issuer's card is used, encouraging them to use that card everywhere they pay.

Strivve's technology also enhances cardholder experience through:

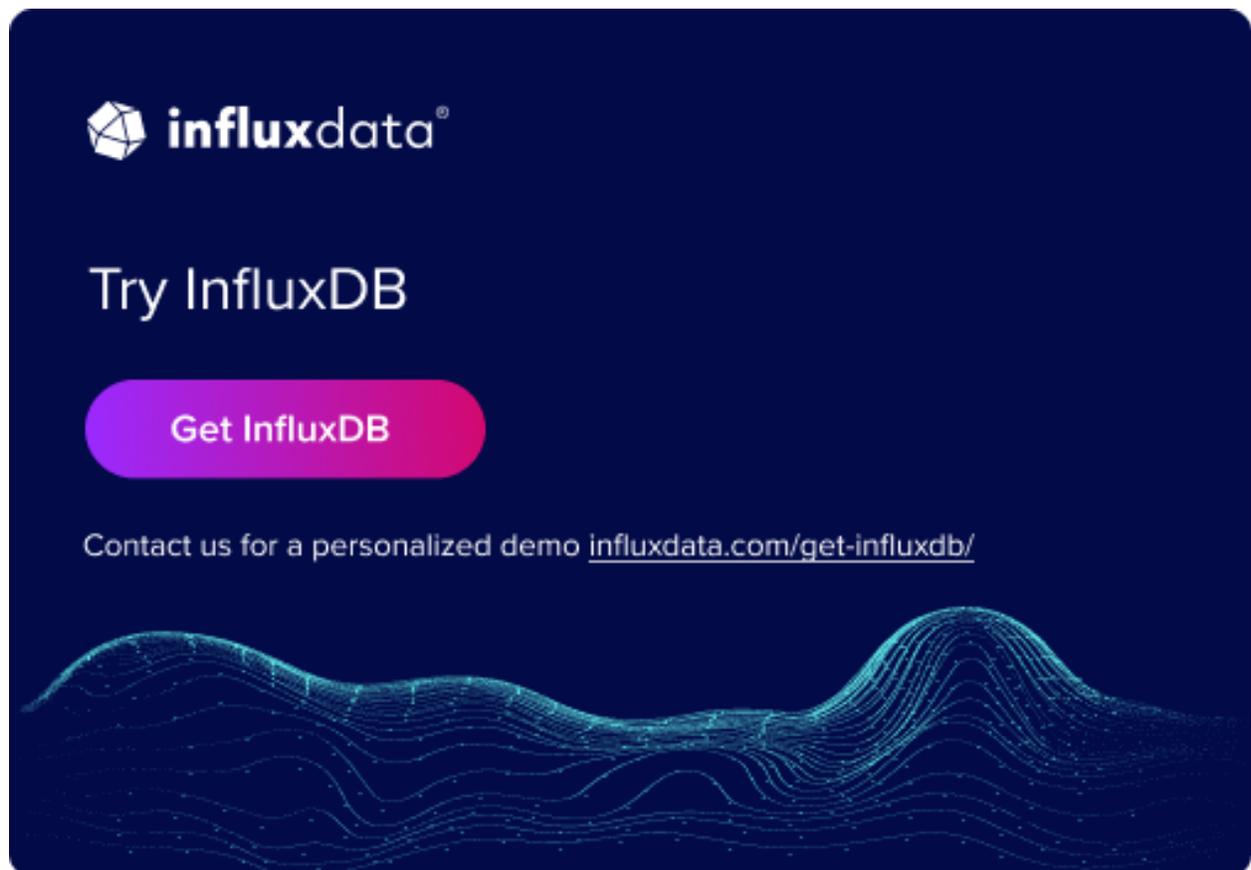
- Building trust - Cardholders no longer associate an issuer's brand with the problem but with the solution
- Strengthening loyalty - Providing cardholders a better card replacement and overall e-commerce experience makes the card sticky
- Customer Service Impact - Cardholders can directly manage card replacement fallout
- Differentiated and Innovative Service Offering - Issuers get 100% control of the user experience for their cardholders

Since one of the greatest card issuer challenges is losing cardholders to competitors, some of the card issuing banks that Strivve is working with today want to use TopWallet app for their own employees to cruise those sites and then crowdsource the information to learn how to interact with them. Strivve's system picks them up and can support those sites. Strivve are also currently launching a production service for a partner with 3 million cardholders switching from Visa to Mastercard.

By choosing InfluxDB Cloud, Strivve have successfully built their groundbreaking credit card updater service to card issuers, merchants, and cardholders alike. It provided them with the scale they needed to bring in lots of time-based site navigation events for analysis into their machine learning frameworks. These events are inherently time-based so querying the data was simple and the hosted solution meant that they could spend their precious engineering time on working on features that will continue to make their solution competitive.

About InfluxData

InfluxData is the creator of InfluxDB, the leading time series platform. We empower developers and organizations, such as Cisco, IBM, Lego, Siemens, and Tesla, to build transformative IoT, analytics and monitoring applications. Our technology is purpose-built to handle the massive volumes of time-stamped data produced by sensors, applications and computer infrastructure. Easy to start and scale, InfluxDB gives developers time to focus on the features and functionalities that give their apps a competitive edge. InfluxData is headquartered in San Francisco, with a workforce distributed throughout the U.S. and across Europe. For more information, visit influxdata.com and follow us [@InfluxDB](https://twitter.com/InfluxDB).

A dark blue promotional banner for InfluxData. At the top left is the InfluxData logo, which consists of a white geometric cube icon followed by the text 'influxdata' in a white sans-serif font. Below the logo, the text 'Try InfluxDB' is written in a large, white, sans-serif font. Underneath this text is a prominent, rounded rectangular button with a gradient from purple to pink, containing the white text 'Get InfluxDB'. Below the button, the text 'Contact us for a personalized demo influxdata.com/get-influxdb/' is displayed in a smaller white font. The bottom of the banner features a decorative graphic of glowing, wavy lines in shades of blue and green, resembling a data visualization or signal waveform.

548 Market St. PMB 77953, San Francisco, CA 94104