



MARKET REPORT

UPDATED

11/10/2022

Fivetran: the \$200M/yr Zapier of ETL

TEAM

Jan-Erik Asplund

Co-Founder

jan@sacra.com

DISCLAIMERS

This report is for information purposes only and is not to be used or considered as an offer or the solicitation of an offer to sell or to buy or subscribe for securities or other financial instruments. Nothing in this report constitutes investment, legal, accounting or tax advice or a representation that any investment or strategy is suitable or appropriate to your individual circumstances or otherwise constitutes a personal trade recommendation to you.

This research report has been prepared solely by Sacra and should not be considered a product of any person or entity that makes such report available, if any.

Information and opinions presented in the sections of the report were obtained or derived from sources Sacra believes are reliable, but Sacra makes no representation as to their accuracy or completeness. Past performance should not be taken as an indication or guarantee of future performance, and no representation or warranty, express or implied, is made regarding future performance. Information, opinions and estimates contained in this report reflect a determination at its original date of publication by Sacra and are subject to change without notice.

Sacra accepts no liability for loss arising from the use of the material presented in this report, except that this exclusion of liability does not apply to the extent that liability arises under specific statutes or regulations applicable to Sacra. Sacra may have issued, and may in the future issue, other reports that are inconsistent with, and reach different conclusions from, the information presented in this report. Those reports reflect different assumptions, views and analytical methods of the analysts who prepared them and Sacra is under no obligation to ensure that such other reports are brought to the attention of any recipient of this report.

All rights reserved. All material presented in this report, unless specifically indicated otherwise is under copyright to Sacra. Sacra reserves any and all intellectual property rights in the report. All trademarks, service marks and logos used in this report are trademarks or service marks or registered trademarks or service marks of Sacra. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any report is strictly prohibited. None of the material, nor its content, nor any copy of it, may be altered in any way, transmitted to, copied or distributed to any other party, without the prior express written permission of Sacra. Any unauthorized duplication, redistribution or disclosure of this report will result in prosecution.



Published on Nov 10th, 2022

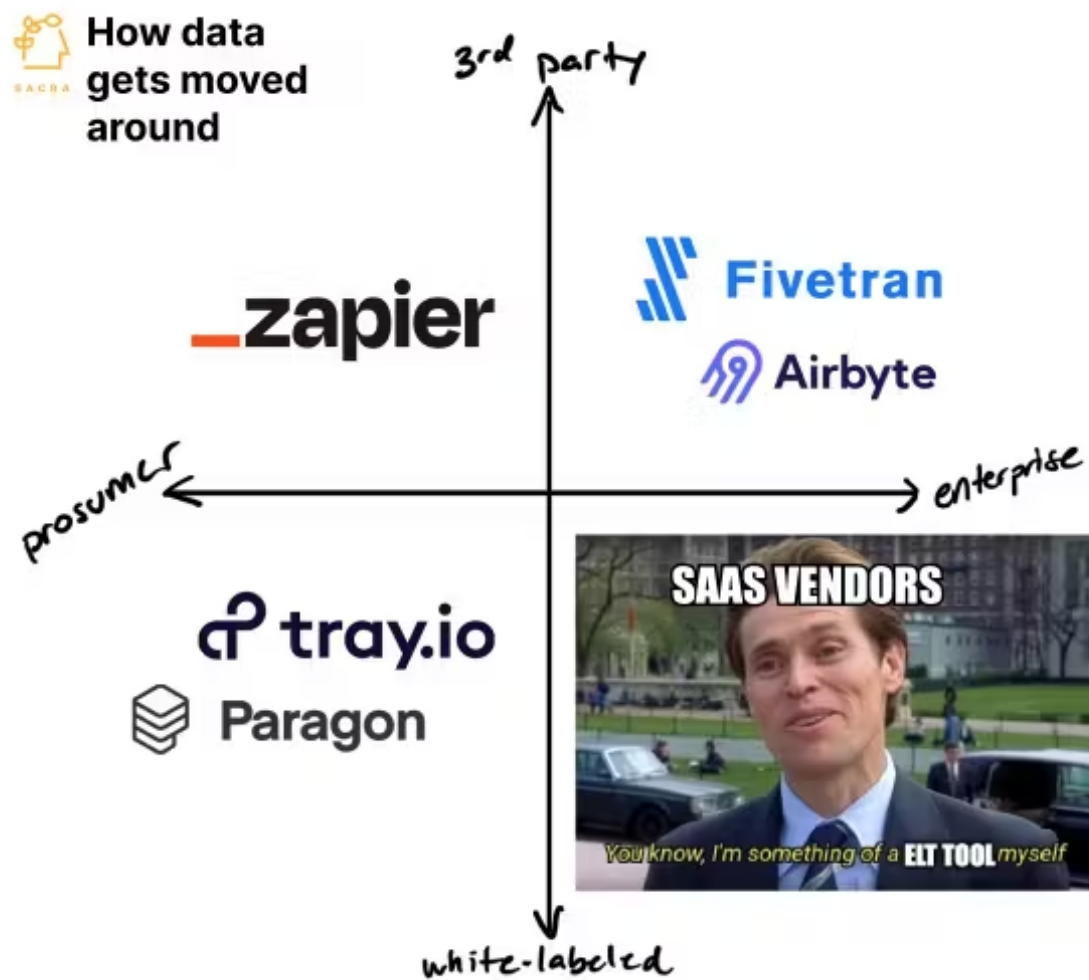
Fivetran: the \$200M/yr Zapier of ETL

By Jan-Erik Asplund

TL;DR: Fivetran (\$5.6B) has built a ~\$190M a year business as the Zapier of ETL, but like Zapier, they face an existential threat as SaaS companies like Salesforce and Stripe begin to offer native integrations—in Fivetran’s case, to the cloud data warehouse. Check out our coverage of Fivetran (dataset) and read our interview with Conor McCarter, co-founder of Prequel, for more.

Key points from our research:

- **With the rise of the cloud data warehouse and proliferation of SaaS, the need arose for connectors that could bring data into the cloud data warehouse from all the SaaS apps that a team uses.** Many SaaS applications built their own BI tools for data analysis, but these were generally inflexible and couldn’t analyze data joined across disparate datasets, limiting teams’ ability to make sense of the data. ([link](#))
- **Fivetran (\$5.6B) found initial product-market fit as the “Zapier” of ETL, saving data engineers time piping data from SaaS apps into a data warehouse by building and maintaining ~200 connectors for the most popular applications.** Fivetran’s model is built on charging companies for the time-consuming work of maintaining their fat head of connectors built on top of SaaS vendors’ ever-changing APIs. ([link](#))

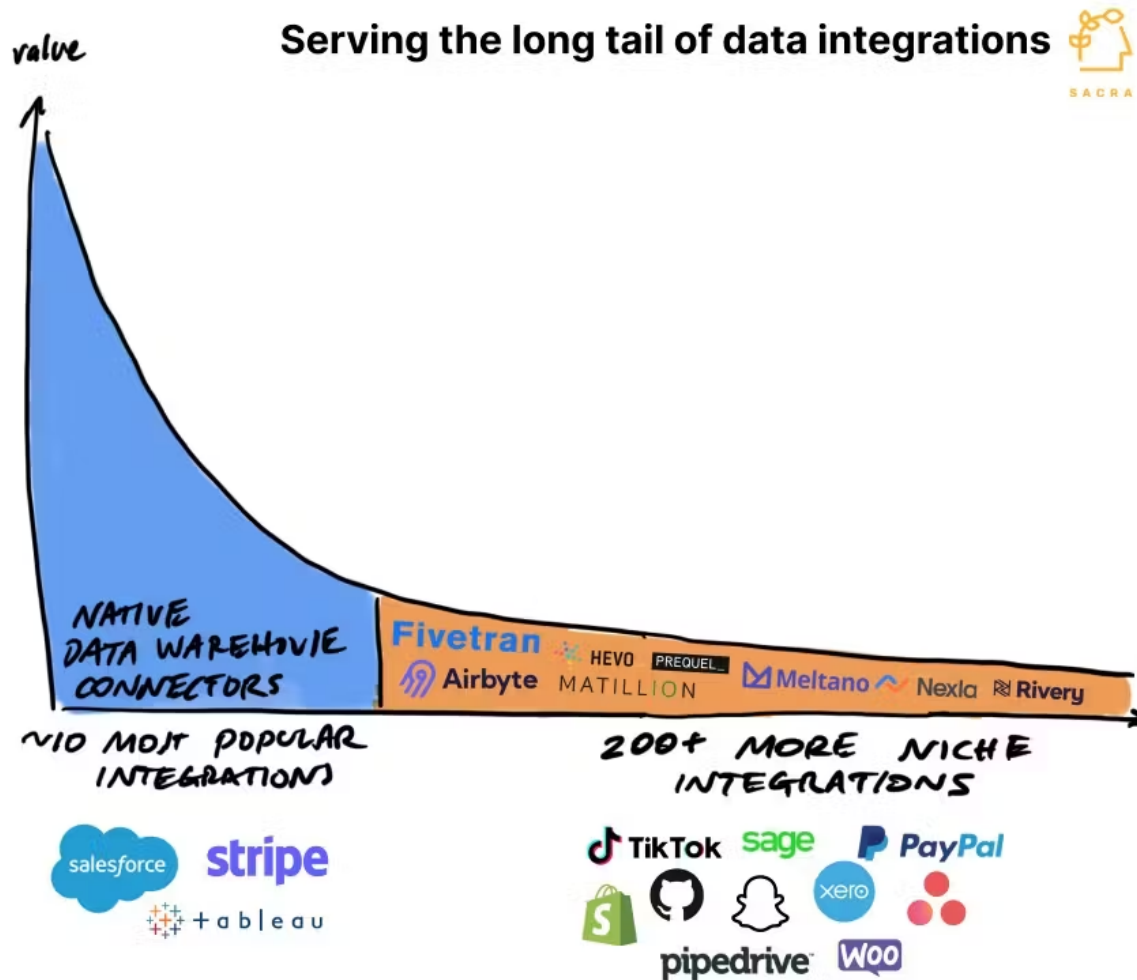


- Fivetran made \$95M in 2021, growing revenue by \$30M after acquiring enterprise competitor HVR to move upmarket—they're now projected to hit \$190M by the end of 2022. Fivetran charges customers on a per-usage basis (how many rows of data they're syncing) indexing them on the rise of companies syncing data from their apps into their cloud data warehouse. ([link](#))
- Zapier grew in a capital efficient way by using their 6M+ uniques per month to incentivize companies to build their own zaps as a distribution channel—Fivetran's model, where they build each connector themselves, comes with expensive ongoing engineering maintenance costs. Fivetran's model indexes on quality of connectors and reducing work for companies' data engineers due to the high criticality of the analysis jobs running through it, while Zapier can take a few more liberties on quality as a primarily prosumer, non-mission-critical tool. ([link](#))
- Similar to how Plaid used screen scraping to build its universal API into banking data, Fivetran relies on API scraping to build its universal API into SaaS vendor data, which means Fivetran's connectors break (requiring 3-7 days to fix) when those APIs change. Fivetran's reliance on vendor APIs also opens up the risk, as with Plaid, that those



vendors will want to reclaim the integration function for themselves and own the data that flows through their pipes.
([link](#))

- **As with Zapier, the data connector is getting commoditized, with Fivetran now competing with an increasing number of ETL companies like Matillion (\$1.5B), Rivery (raised \$48M), Meltano (raised \$12.4M), Hevo Data (raised \$43M), Nexla (\$15.5M) and Airbyte (\$1.5B) that traverse the spectrum of connector quantity/quality and managed/custom.** Some of these companies expose API templates to companies so their engineers can manually tweak their off-the-shelf connectors, while others aim to use community-contributed connectors to cover more of the space.
([link](#))
- **Open source Airbyte (\$1.5B) made around \$1M last year and has become a prime competitor attacking the long-tail of SaaS-data warehouse integrations that Fivetran won't build by shifting the responsibility of building connectors to their customer base (9,000+ strong as of 2021).** The downside of Airbyte's model is that while it already offers 200+ connectors, it relies on the companies that use their data warehouse connectors to maintain and update those connectors—with inconsistent quality and results—over time.
([link](#))



- Today, there's a growing trend of SaaS companies like Stripe, Salesforce and Customer.io offering native data warehouse integrations that increase retention/activation and revenue (Stripe charges \$0.03 per transaction). SaaS vendors can design and support their customers' export needs end-to-end rather than send them to a 3rd party tool like Fivetran. ([link](#))
- These native integrations are rising first across high-volume vendors like Stripe, which cost Fivetran customers the most to sync because of Fivetran's per-row pricing model. Customers don't churn from all their Fivetran connectors at once, but instead begin to shift usage of high-value/high-frequency connectors first, ala ecommerce business that have to run frequent syncs on all their transactional data moving to Stripe's native data warehouse integration. ([link](#))
- Where Fivetran sold connectors and Airbyte sold their platform/community for building connectors, new tools like Prequel (\$125K raised) sell the ability for vendors to build their own data warehouse connectors. Instead of ceding the revenue customers will pay a Fivetran/Airbyte for a warehouse integration—companies can sell their own native connector, charging anywhere from 10-30% of ACV for it. ([link](#))
- With the continuing growth of both data warehouses and the long tail of SaaS, there's an opportunity for Fivetran



and similar companies to build a platform that provides teams with visibility across all their data pipelines for observability and debugging. Today, engineers have to spend time monitoring a variety of different transformations, sources, and destinations to have a comprehensive view across all their data pipelines. ([link](#))

For more, check out our [full interview with Conor McCarter, co-founder at Prequel](#), and this other research from our platform:

- [Fivetran company_profile \(dataset\)](#)
- [Databricks company_profile \(dataset\)](#)
- [Julia Schottenstein, Product Manager at dbt Labs, on the business model of open source](#)
- [Sean Lynch, co-founder of Census, on reverse ETL's role in the modern data stack](#)