TECTON + **HELLO FRESH**

HelloFresh Adopts Tecton as a Foundational Component of its MLOps Platform

> *"At HelloFresh, our philosophy is to avoid reinventing the wheel. When it came to feature stores, building one in-house was never a viable option. We needed to move quickly, and we wanted the best solution available. That's what led us to Tecton."*
>
> — Erik Widman, Ph.D., AI Product Management Director at HelloFresh

OVERVIEW

About:

HelloFresh SE is a global food solutions group based in Berlin, Germany. It is the largest meal-kit provider in the world and operates in 18 countries, including the US, UK, and Australia.

Challenge:

Building the Foundations of a Data Layer

In 2019, HelloFresh set out to build the foundations of a data layer on its path to automating decision-making across the value chain. From the project's inception, ML and engineering teams already knew that the feature store was a central piece of infrastructure that would help solve the standardization problem involved in building and deploying data products at scale.

Solution:

Benchmarking ML Feature Solutions

HelloFresh engaged in an agile process to test the best feature store solution. They picked an existing machine learning model, broke apart the features, uploaded them to the competing vendor solutions, tested different scenarios, and assessed scores for each. This quantitative approach removed the possibility of personal bias and led to objective decision-making across teams.

Results:

Tecton as a Foundational Part of the MLOps Platform

After adding up all the weighted scores into a total category score for the vendors, the platform team came to the following conclusions:

- Tecton 90%
- Vendor 2: 63%
- Vendor 3: 62%

In under six months, HelloFresh successfully onboarded eight teams that now autonomously build data pipelines and materialize feature values to power models with Tecton. One of their biggest projects is Morpheus.

A Deep Dive into HelloFresh's Quantitative Benchmarking Approach

Since its launch in 2011, HelloFresh has grown exponentially. The company is now publicly traded, counts more than 8 million active customers, employs over 21,000+ people worldwide, and is the world's biggest meal-kit provider. To support its growth, HelloFresh invests heavily in creating machine learning models and delivering trustworthy data products to the rest of the organization.

Increasing Tech Debt Tied to a Centralized Model

Global Success and Rapid Growth Comes With...

By 2015, HelloFresh had organically grown a classical centralized data management setup. A handful of people, including a centralized BI team, used a small number of internal and external resources to produce executive reports to steer the company's performance. However, at this stage of their growth, HelloFresh's data management systems remained siloed even as the centralized data warehouse team provided the flexibility for the company to grow quickly.

When the company went public in 2017 and expanded to new countries, launched new brands (EveryPlate), and acquired other brands like Green Chef (2018) and Factor (2020), they also increased engineering headcount from approximately 40 to almost 300 engineers, and substantially grew other business units, from marketing to sales, as well. This explosive growth led to reprioritizing business objectives across the company and resulted in an increase in the number and complexity of data requests, to the point that data engineers suffered from context switching. In short, the centralized model became a bottleneck and slowed down innovation.

...Growing Pains: Siloed Data and Tech Debt

The bottleneck led analytics teams in marketing and supply chain management to build their own data solutions and select their own tools. Unfortunately, these home-grown, bespoke data pipelines did not meet engineering standards and increased the central team's maintenance and support. Over time, the knowledge of those bespoke pipelines faded or disappeared. Data assets turned into debt, and data quality decreased.

Simultaneously, data science teams formed organically in different parts of the organization, creating machine learning models to tackle problems from marketing applications to logistics optimization. MLOps tooling was selected ad hoc by the individual teams to try to push models to production as quickly as possible, creating a fragmented approach to model productionalization that incurred significant tech debt.

Pursuing the Foundations of a Data Layer

In 2019, HelloFresh set out to ensure that its data scientists and engineers could quickly discover, understand, and securely access high-quality data to prototype and deliver large-scale predictive data products. The ultimate goal was to build the foundations of a data layer that the company as a whole could trust on its path to automating decision-making across the value chain.

For data scientists and machine learning engineers, the first step to standardizing the use of feature data across the organization was figuring out what was important to the data science teams. That's when Erik Widman, HelloFresh's AI Product Management Director, set out on a mission to benchmark the solutions that would, in time, provide infrastructure and tools to allow domain teams to focus on building data products, reduce domain-agnostic complexity, and minimize time to insights. The first part of this process was figuring out which parts of the existing infrastructure worked well, which ones needed to be replaced, and above all, which ones were completely missing.

The Feature Store Component, a Fundamental Part of Standardization

When the project began, although the concept was relatively new, ML and engineering teams already knew that feature stores were a central piece of infrastructure that would help solve the standardization problem involved in building and deploying data products at scale. But who—particularly in a large organization with organically grown data science and engineering teams—would have ownership of a feature store solution?

To begin the process of choosing a feature store, Erik and his team gathered some of the data science teams to better understand their needs and pain points. From February to March 2022, the groups engaged in an agile process designed to not only understand what part of the puzzle a feature store could solve, but also to create buy-in of a solution across the organization.

To Build or To Buy?

From there, the obvious next question was whether the team should build a feature store from scratch, investigate open-source feature stores, or look into buying an all-in-one platform.

HelloFresh wanted to move quickly and, for this reason, decided against building an in-house feature store from scratch because it would be a lengthy and costly process. Opting for an open-source solution for the feature store component was almost immediately eliminated—most available solutions didn't have all the product features HelloFresh needed to solve the standardization problem that was starting to negatively impact productivity.

> *"We didn't want to be in the business of creating MLOps tools when there are so many companies out there who are ahead of us in the game. Instead, we want to reduce the time to push models to production, creating more business value for HelloFresh"*
>
> — Erik Widman, Ph.D., AI Product Management Director at HelloFresh

Go Wide or Go Deep?

As for going down the end-to-end platform route, leadership had already learned from past experiences that the political buy-in for a "one tool fits all" approach would be very difficult. After all, many of their existing tools worked well, and replacing them would be more costly than beneficial to the teams depending on them. Furthermore, many end-to-end platforms that claimed to have the "feature store" component simply didn't have the bandwidth or the dedicated teams to work on making that part of the puzzle the best it could be. With this in mind, the team narrowed down the number of vendors and benchmarked Tecton against other leading feature platform providers.

## Benchmarking the Leading ML Feature Solutions

Developing a Data-Driven Framework

To quickly identify the topics that really mattered to the data scientists, Erik's team held some agile brainstorming sessions where data teams jotted ideas on Post-it notes and clustered them into categories. Erik then translated the topics into a list of questions to which he assigned a weighted priority level (High = 5, Medium = 3, Low = 1) depending on technology, cost, user experience, and long-term vs. short-term strategic considerations.

For instance, topics such as "What data formats can be stored as features (important formats: Parquet, data lake, JSON)?" or "Can you define feature transformations with SQL / PySpark?" were HIGH priority, whereas "Does the solution support embedding as features" or "Is automation possible via the CLI?" were LOW, and thus weighted differently. "We identified many questions, but it's important to realize that all considerations are not equally important, and the key is to pick out what matters," explained Erik.

Picking a Winner: POCs and Vendor Scoring

Define Pre-screening Criteria Define team's needs in a feature store and differentiate nice-to-have vs. must-have features

Vendor Website Review Review vendors and 3rd party websites to compare high level platform features

Whitepaper Review Dive deeper into the platforms of interest to understand details of a technical implementation

Vendor Demos Conducted demos with our preferred top vendors

Evaluate Vendors Vendors were evaluated against the pre-screening criteria and Tecton emerged as the preferred option
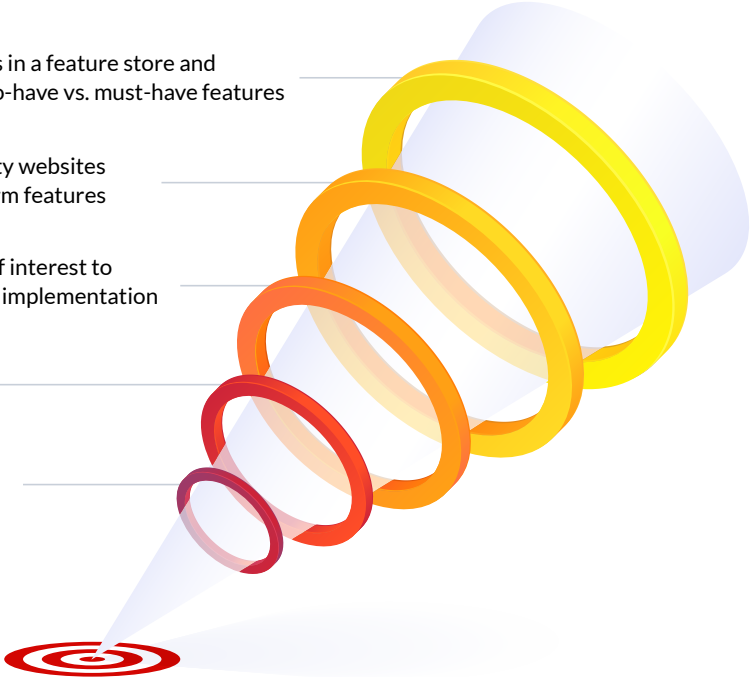
Erik's team wanted to conduct the POCs and benchmark other providers in the feature store space, as well as compare leading vendors (Vendor 3, Vendor 2), and Tecton against each other. To do this, they picked one of HelloFresh's machine learning models, broke apart the features, and uploaded them to the different vendor solutions so they could test different scenarios and better assess the scores for each listed category.

The team then presented the analysis to data leadership, explained how the platforms scored against each other, and walked through the associated business risks. This quantitative approach removed the possibility of personal bias and led to objective decision-making across teams.

After adding up all the weighted scores into a total category score for each tool, with the highest possible score of 430, the platform team came to the following conclusions:

- Tecton 388/430 (90.23%)
- Vendor 2: 273.5/430 (63%)
- Vendor 3: 269.5/430 (62%)

Tecton fulfilled 90% of the feature store requirements established by HelloFresh compared to 63% and 62%, respectively, for Vendor 2 and Vendor 3.

Feature Platforms Compared	TECTON	Vendor 2	Vendor 3
Vendor Quality	Small but most mature	Good	Good
Feature Tagging	✔	◗	◗
Search	✔	◗	◗
Version Control	✔	✘	✘
Compatible Formats	✔	✔	✔
Online + Offline Store	✔	✔	✔
RBAC	◗	◗	◗
Databricks Integration	✔	◗	✔
Spark Integration	✔	✔	✔
AWS Integration	✔	✔	✘
Technical Scores — Max Score: 430	388 (90.2% of max)	274 (63.7% of max)	270 (62.8% of max)

Key: ✔ Fully Covered ◗ Partially Covered ✘ Missing

Change Management & Feature Store Roll-Out

After selecting Tecton as the best possible feature store component on their path to standardizing the use of data to build reliable predictive products, HelloFresh now had to roll out the feature store to the many engineering and data science teams across the organization.

Early in the process, leadership was aware that simultaneously opening the feature store to everyone could create feature overlap. To avoid this problem, the team responsible for rolling out the feature store inter-

viewed the different data science teams, mapped out what models were in production, quantified business impact and feature reusability for other teams, and discussed which model features would be suitable for the feature store. Next, they considered individual team roadmaps to assess their availability to start transitioning their existing features and feature roadmap to Tecton.

From there, they gave read access to all the teams that would eventually start using Tecton. Finally, they prioritized which teams would go through a quarterly training session before unlocking the ability to upload and build features into Tecton.

In under six months, HelloFresh successfully onboarded eight teams that can now autonomously build data pipelines and materialize feature values that models can ingest to make predictions. For new model development, Tecton is now the interface between the data and the models.

The Feature Store as a Foundational Part of the MLOps Platform

Now that HelloFresh is rolling out the adoption of Tecton across teams, HelloFresh is also building an MLOps platform as part of their continued goal of better understanding and addressing consumer needs with high-quality machine learning.

This platform is intended to make it easier for data scientists and ML engineers to create robust pipelines, standardize the underlying tools, simplify scaling models across geographies, and reduce the time to put models in production.

Tecton is the foundation at the feature and ETL level and is helping standardize the other components involved. The MLOps platform consists of different products for different end users, and the main product is a low-level API library called Spice Rack. Spice Rack is an API wrapper that controls an MLOps core layer, which interacts with the different tooling used for the pipeline. Spice Rack consists of various components that control each part of the pipeline creation process. End users can quickly learn and load these individual services into their notebooks to control various parts of pipelines or work from pipeline templates to accelerate pipeline creation. The democratization of widely available and more powerful tools will enable more teams to productionize models faster and focus on creating models rather than building — and maintaining — pipelines.

Tecton and Hellofresh, in practice

Morpheus: Personalized Experiences & Customer-level Predictions

Morpheus, HelloFresh's customer prediction model, is a cornerstone of the company's ongoing program to democratize access to high-quality machine-learning models. Born out of HelloFresh's drive to offer consumers personalized experiences, Morpheus is HelloFresh's newest algorithm that uses the latest ML techniques to offer weekly customer-level predictions.

Challenge: The Limits of a Cohort Perspective

Before Morpheus, HelloFresh would forecast the future average order rate of its customers by grouping together customers who signed up the same week. Stored in standard formats, these predictions fueled financial reporting, marketing optimization, and digital product development.

But this cohort perspective was limited because the relative importance of a feature set available to make predictions could vary depending on a particular customer segment. For example, a feature set regarding a customer's previous actions can, in some cases, be handy to predict the value of a reactivation. But in cases where there was little or no data for a new customer, that feature set was small, and its predictive power was non-existent.

Solution: Reframing Value with Customer-level Predictions

Today, HelloFresh has deployed Morpheus, a system that reframes customer profitability with 1,360 different gradient boosting models. Each model is trained in parallel on multiple geographies and time horizons for specific customer segments.

But models are only as good as the data that powers them, and Morpheus depends on the engineering of informative features from the collection and transformation of as much high-quality data as possible. That's why the teams behind Morpheus are in the process of migrating all of the Morpheus features into Tecton. When the full migration is complete, Tecton will run the hundreds of Morpheus-specific predictors from various data sources. For example, features developed from customer acquisition funnels and CRM data—such as how customers interact with the company's emails—can be strong indicators of a customer's future profitability.

Tecton will not only help the teams streamline features from all types of different data sources to power the thousands of models in Morpheus, but it will also help the teams ensure data quality and consistency across environments, and supervise training pipelines' progress to avoid common ML pitfalls like model drift or training/serving skew. By ensuring good pipelines that transform raw data into predictive signals and by providing those fresh features to the models automatically, the teams behind Morpheus will be able to focus their time on improving model development and scaling the models out to new territories.

Results: 1,360 Models Up and Running

Today, Morpheus has 1,360 models up and running, making predictions that teams across finance, marketing, product, and operations use to make business-critical decisions. The data scientists and engineers behind Morpheus are migrating their features to Tecton to enable continuously computed accurate, fresh features on demand. This central source of truth for feature materializations will enable the teams to scale Morpheus to other geographies. Additionally, it will allow teams to use subsets of these features to power even more models across teams and use cases.