



At a glance

Company

Gild

Industry

Talent Acquisition Software

Location

San Francisco, USA.

Web

www.gild.com

Highlights

- Millions of Developers with Dozens of Multi-Dimensional Attributes
- Elasticsearch used in a polyglot database environment alongside Cassandra and MongoDB
- Transitioned from a self-hosted Elasticsearch infrastructure to Found.no. Saw uptime improvements and was able to narrow operations focus to components that could only be handled in-house.

Searching Millions of Developers with Minimal Infrastructure

About Gild

Gild is a true innovator in the software talent acquisition market, helping recruiters at companies like Rackspace, Square, and Redhat find the best tech talent out there through its suite of automated tools. Gild has developed tools to sort through millions of developers' public data and score them based on expertise and demand. This capability is a crucial part of segmenting and identifying talent, making Gild a leader in the talent acquisition market.

The Challenge: Making Millions of Developers Quickly Discoverable

Gild's technical team found itself in the market for a search product in 2012 after a pivot from a previous business model. The business model they moved to, which has since proven to be a great success, revolves around automating previously manual tasks. By combining machine learning techniques with data scraped from multiple data sources on the internet, Gild has been able to build up a giant database of developers and their skills.

Calculating a developer's ability is no simple task. Gild tackled this problem using a number of proprietary tools which evaluate source code repositories. These tools evaluate a developer using a number of metrics, such as the quality of open-source code to determine the expertise of the developer. In addition to expertise, their tools help ascertain how in-demand the developer is. By measuring developers along these two axes, Gild helps recruiters better target developers.

The system Gild built to manage this, while intricate, is fairly straightforward in its general approach. Public developer sites like GitHub are crawled looking for data on developers. The data is then processed using proprietary algorithms, the output of which is then fed into various databases. The Cassandra NoSQL store, is used for their offline processing and machine learning tasks, Redis is used as a communications bus between services, and MongoDB is used for their OLTP needs. A proper search engine was needed to make much of this data discoverable by users, and that's where Elasticsearch entered the picture.

"There were three major technologies we were considering, Amazon Cloud Search, SOLR, and Elasticsearch. Our engineers felt that Elasticsearch was very promising."

Luca Bonmassar, Co-founder, Chief Product and Technology Officer

The Gild technical team had a short list of very specific requirements for a search engine. First, they needed something that was featureful and reliable. Second, they needed data to be searchable in near real-time. Finally, they needed a product with a solid scalability story. Gild wound up going with Elasticsearch due to its combination of a solid scaling story, real-time capabilities, and vibrant user community.

About Found

Found Elasticsearch is a hosted and fully managed search service powered by Elasticsearch.

We provide a fast, scalable and stable search platform that is developer friendly and easy to operate.

Found Elasticsearch suits a range of uses, from full text search on your web site to Big Data analytics.

Benefits using Found

- Hosted and managed by search experts.
- Your own dedicated Elasticsearch cluster.
- Reserved memory and storage.
- Start small, grow big. Scale your cluster anytime.
- HA with replication and automatic failovers.

Web

www.found.no

Twitter

[@fooundsays](https://twitter.com/fooundsays)

The Solution: Elasticsearch, From In-House to Cloud

There was only one question remaining: how to host Elasticsearch?

Initially, Gild made the decision to host their Elasticsearch clusters in-house. While they were able to quickly get going from a functional perspective, setting up the servers, writing queries for their app and executing them on the cluster; they quickly ran into operational issues, among them the all too famous split brain problem. While a distributed database such as Elasticsearch can be far more reliable than traditional databases, the complexity behind that reliability can actually pose serious problems when not understanding the subtleties of an Elasticsearch deployment.

"After going through the pain of restarting systems at 2am, we realized that it made perfect sense to outsource this to someone else."

Luca Bonmassar, Co-founder, Chief Product and Technology Officer

Having made the decision to outsource, Gild set out to find a hosted Elasticsearch provider. Describing the process they went through, Mr. Bonmassar recalled: "I remember the Director of Engineering finding Found, which was still in private beta, and he was very comfortable with the answers he got from the Found team". As a trial, Gild decided to continue running their own Elasticsearch servers and to duplicate all incoming data between Found and their own clusters, just in case. They've since found that they've never had to switch over to their own backup clusters. "Found hosted Elasticsearch has been rock solid since day one."

Business Benefits: Keeping Focus

The key benefit for Gild migrating to Found has been the ability to retain focus on the operational concerns specific to their core business. Describing the strategic reasoning behind their migration to a hosted Elasticsearch service, Mr. Bonmassar noted: "The main goal of having someone in operations is to have someone available to take care of non-standard components."

Costs must also be considered when choosing a cloud service provider. It's a calculation CTOs and CIOs everywhere are performing more often with the rise of SaaS infrastructure offerings. Describing the costs involved with using a cloud service, versus training and hiring in-house administrators, Mr. Bonmassar once again related the advantages of cloud services: "Regardless of the cost difference, it was worth the time we got back."

In Gild's case, moving to a cloud provider was a solid win. Tricky technology needs expert attention, and when it comes to managing business critical data, expert ability is what's needed to power serious services. When asked whether Gild will continue using Found's services in the future Mr. Bonmassar happily remarked that they "have been approached by many competitors to give them a try, including free Elasticsearch hosting, but we'd rather pay for quality service."