fusionbase

Fast & Agile Prototyping of Data-Driven Applications in Innovation Department

Challenge

The innovation lab of a German technology giant searches for new ways to radically reduce the development time of data-driven products. Their two biggest challenges:

- Being agile in assessing the quality and relevance of hundreds of data sources without starting expensive and engineering extensive sub-projects. These sub-projects result in costs so high that potential business cases are destroyed before evaluation.

- Finding, acquiring, and integrating third-party data to fill missing features of their internal data. The lack of features usually results in lower-quality products.

Previous Approaches

Before Fusionbase, they were guessing which of the enterprise's data could be useful for a potentially new data-driven product. Once they identified such a relevant dataset, they were required to start a process requesting this data from the respective data owner. After several days, the department's engineers or the central IT usually sent them an export of the requested data. Then, they had to assess the quality and relevance of the dataset – and repeat it over and over again.

Solution with Fusionbase Data Hub

To address the first challenge, a dedicated Fusionbase deployment team quickly set up Fusionbase Data Hub on the innovation lab's cloud environment. After the initial setup, their data scientists immediately started to connect to databases and data lakes of the other departments.
With Fusionbase's virtualized data access model, the data owners now can quickly and securely share access to their data sources. The data scientists can instantly scan the metadata of a dataset and run analytical queries to assess its quality.
The Fusionbase data marketplace addresses the second challenge by providing access to thousands of open and third-party datasets - cleaned, queryable, and ready for analytics. They found the data they were looking for and even more they had not anticipated before.
With tools like Python, Jupyter, and plotly they connect to the Fusionbase real-time API and explore the data further.

Results

- ✓ **Average data access** in 5 hours instead of 12 days
- ✓ 3648 **engineering hours** saved
- ✓ **Iteration cycle** of 20 days reduced to 5 days

Sounds like a familiar challenge? Get in touch.

sales@fusionbase.io | www.fusionbase.io