



Data Hub

Reviews, tips, and advice from real users



Powered by  **PeerSpot**

Contents

Product Recap..... 3 - 4

Valuable Features..... 5 - 10

Other Solutions Considered..... 11 - 13

ROI..... 14 - 15

Use Case..... 16 - 19

Setup..... 20 - 22

Customer Service and Support..... 23 - 25

Other Advice..... 26 - 29

Trends..... 30 - 31

About PeerSpot..... 32 - 33

Product Recap



Data Hub

Data Hub Recap

Data Hub is an advanced platform designed to streamline data management processes, enhance data accessibility, and provide comprehensive analytics capabilities for informed decision-making.

Data Hub offers a unified approach to handling large-scale datasets, empowering organizations to effectively manage, analyze, and extract insights from their data infrastructure. It provides robust features for data integration, storage, and visualization, supporting diverse business needs and driving data-driven strategies.

What are the key features of Data Hub?

- **Data Integration:** Seamlessly combines data from multiple sources for cohesive analysis.
- **Scalability:** Supports growing data needs without performance degradation.
- **Analytics Tools:** Enables advanced data processing and insightful reporting.
- **Security:** Robust security protocols to ensure data compliance and protection.
- **Custom Dashboards:** Offers customizable interfaces for tailored analytics displays.

What benefits do users experience with Data Hub?

- **Increased Efficiency:** Streamlines data processes, reducing time and resource spent.
- **Cost-Effectiveness:** Minimizes need for multiple data management tools.
- **Improved Decision Making:** Provides actionable insights through data analytics.
- **Enhanced Collaboration:** Facilitates data sharing across departments, encouraging teamwork.

Data Hub is implemented across industries such as finance, healthcare, and retail, providing tailored solutions that meet specific demands in areas like customer data analysis, patient record management, and inventory tracking. Its ability to adapt to sector-specific requirements makes it a versatile choice for businesses seeking enhanced data capabilities.

Valuable Features

Excerpts from real customer reviews on PeerSpot:



“Data Hub has positively impacted my organization by significantly reducing manual work that was previously needed to identify upstream and downstream data relationships, as well as recognizing duplicate datasets.”



Akashkhurana Hirana

Senior Software Engineer 2 at Porch



“Data Hub has impacted my organization positively by helping us build a data governance environment and share the knowledge about the data for the entire company.”



Matheus Costa

Data Engineer at CI&T



“One of the biggest advantages of Data Hub is the very good integration, for example, a department focused on development made the integrations between Data Hub and BigQuery.”



Henrique dos Anjos

Data Quality Engineer at truelogic



“Data Hub positively impacts my organization by enhancing collaboration as previously, we had to ask the team to provide the schema information.”



PrashantGupta2

Software Engineer L2 at a tech vendor with 5,001-10,000 employees



“Data Hub has positively impacted our organization by reducing the knowledge transition period from three months to one month for new team members, enabling them to refer to the complete lineage without depending heavily on others, which is a substantial improvement.”



Shubham-Agarwal

Manager - Projects at Cognizant



“Data Hub proved to be a robust, scalable, enterprise-ready data catalog that is well-suited for AWS-based architecture and complex organizational environments.”



Verified user

Software Engineer at a tech vendor with 10,001+ employees



“Data Hub has saved approximately two million every four years, which is one of the major savings.”



Verified user

Lead Business Analyst at a tech vendor with 10,001+ employees

What users had to say about valuable features:

The best features that Data Hub offers and stand out to me are its ability to invest in metadata from a variety of databases including Oracle, MySQL, AWS database, and MongoDB. Data Hub can also work with files, embedding dataset structures. It is easy to comprehend and facilitates our teams, such as BI, EI, and data governance, to understand data and IT structures. Using Data Hub, we can build and maintain a live data dictionary, providing company-wide clarity on data.

Anhnq Ho

Finance Feedback Committee at MB Shinsei Finance Limited Liability Company

[Read full review](#) 

“The best feature that Data Hub offers is its data catalog. The level of detail in the data catalog stands out to me, as it represents the column details, description, and tags, which are easily accessible whenever I want to access a particular column. Data Hub has positively impacted my organization by helping me with the data governance aspect. It has helped me reduce errors in data governance, save time, and improve compliance..”

Verified user

Associate Staff Engineer at a tech vendor with 10,001+ employees

[Read full review](#) 

“The data injection feature was valuable to me. If comments were inserted in the tables, the data would automatically gather and enter all necessary data into Data Hub. Additionally, the data lineage graphs were really helpful in showing how data flowed to data marts and which columns were used for creating other columns. Those were helpful but somewhat hard to manage features..”

Verified user

[Read full review](#) 

Data Engineer at a consultancy with 11-50 employees

“The best features that Data Hub offers include metadata propagation and lineage propagation.

These features have specifically helped my team and our workflows by enabling people to find the right data. We have different sets of data that include business data and application data. People who are new, including data analysts, machine learning scientists, or data scientists, can easily find the specific data they are looking for because it is all centralized in one place.

Data Hub has positively impacted our organization by centralizing and co-locating all data through metadata, and we have made this our enterprise metadata catalog rather than having disorganized information across different teams. It has saved time for many data analysts and data scientists to find the right data..”

Somashekar Venkataramaiah

[Read full review](#) 

Senior Technical Program Manager at a financial services firm with 10,001+ employees

“I appreciate two things about Data Hub. One thing I appreciate a great deal is the ability that it has to automatically integrate with some of your tools and work as a receptor for receiving data information from your pipeline. For example, if the pipeline starts reading this table and writing these other tables, all this lineage information is sent to Data Hub automatically for some sources, while others only allow a pull-based approach. In those cases, the software needs to push the metrics from time to time to Data Hub, but it allows us to stream all these metrics to Data Hub, which is excellent and is almost seamlessly integrated with some tools that we had.

The second one is the lineage itself, because Data Hub offers a tremendous amount of data lineage capabilities, even column-level lineage, which was very useful for us and is still very useful.

I find Data Hub to be a very important data catalog tool for a company that values data..”

Igor Uchoa

Senior Data Engineer at a tech services company with 1-10 employees

[Read full review](#) 

“The best features are that we do not need to use multiple tables to create our own flow. With one enriched data set that we are creating by combining the main data set with referentials, consumers can cater to their needs through one specific data set rather than having various tables at our disposal.

The second best feature is that security is built in here, whereas on Oracle, we had to manage all security-related aspects by maintaining the RDLs, restricting users based on their access levels. On Azure, security is based on the Azure RDLs that are already in place for a public Azure data lake account.

The third feature, which I consider one of the best, is OPAL, which I have already explained as Open Platform Access Library. This helps users retrieve data much faster.

The fourth advantage for the application creating this data set is that one of the new versions of OPAL, called OPAL++, helps the application audit which particular user is using this data and what kind of queries they pass throughout the day. With this, statistics can be built to see which user is utilizing the most resources and how they are utilizing them.

Data Hub has saved approximately two million every four years, which is one of the major savings. In terms of a second savings metric, more consumers have onboarded to the Azure Data Lake compared to those onboarded before, as the process is more streamlined and structured. The third savings is in terms of data on tables. Previously, tables were added as per user request, but on Azure, we have established an Enterprise Data Model. Any feeder that publishes data into Azure Data Lake must adhere to this particular model. This way we have ensured that normalization and standardization are properly taken care of..”

Verified user[Read full review](#) 

Lead Business Analyst at a tech vendor with 10,001+ employees

Other Solutions Considered

Before choosing Data Hub, we evaluated other options including Apigee and IBM. I do not remember the exact names, but we chose Data Hub because we had time to pilot it and found it to be the best choice at that moment.

Anhnq Ho

Finance Feedback Committee at MB Shinsei Finance Limited Liability Company

[Read full review](#) 

“I was not previously using a different solution. The implementation was already directly part of a data governance initiative and it was done directly with Data Hub, meaning there was no previous solution..”

Henrique dos Anjos

Data Quality Engineer at truelogic

[Read full review](#) 

“Previously, we used the Snowflake inbuilt lineage graph to identify data flow, but we switched to Data Hub for its centralized governance capabilities across multiple databases..”

Shubham-Agarwal

Manager - Projects at Cognizant

[Read full review](#) 

“We searched for other use cases and how other companies were developing their own solutions for using data, but we have not been able to use other products similar to Data Hub itself..”

Verified user

[Read full review](#) 

Data Engineer at a consultancy with 11-50 employees

“We have significant contracts with Google, Amazon, and Microsoft, all of them and our platform, which you can visit at peerspot.com, offers a comprehensive search bar where you can put in almost any product and probably find it, even if there are not a lot of reviews for it. So, Data Hub, as I remember, is a sort of metadata management and AI observability tool..”

Igor Uchoa

[Read full review](#) 

Senior Data Engineer at a tech services company with 1-10 employees

“I'm not using Atlan anymore because the company that I was working with, I'm no longer there. I went to another consultancy group and now I'm working with other platforms. Atlan is not the one that I'm working with at the moment.

“I am working with a different platform that is also regarding data governance and metadata management. The platform itself, the back end, is Data Hub. But the user interface is customized for this client. I'm currently working for Uber, the Uber company..”

Henrique dos Anjos

Data Quality Engineer at truelogic

[Read full review](#) 

ROI

Real user quotes about their ROI:

We have seen a return on investment from using Data Hub, as it saves our data governance team time by collating metadata and viewing the live data dictionary. It is very helpful in building data quality for the company, leading to approximately thirty percent improvement in efficiency.

Anhnq Ho

Finance Feedback Committee at MB Shinsei Finance Limited Liability Company

[Read full review](#) 

“I have seen a return on investment with Data Hub, notably in reducing the knowledge transition period and improving our ability to troubleshoot production issues in Power BI, thus saving time..”

Shubham-Agarwal

Manager - Projects at Cognizant

[Read full review](#) 

“I have seen a return on investment with Data Hub. For instance, I have noticed time savings during incidents and while looking up schemas. In terms of resources, Data Hub centralizes data cataloging and classification, saving us from having to disclose PII column information to teams not utilizing it. Regarding financial metrics, I do not have specific metrics available..”

PrashantGupta2

Software Engineer L2 at a tech vendor with 5,001-10,000 employees

[Read full review](#) 

“In terms of ROI, I would say that Atlan is better. I had a very good experience using Atlan, and I believe it's faster. Velocity in organizations today is very important; people want to see things very fast. I believe Atlan has a better approach compared to Data Hub.

“The way Data Hub is implemented at the moment, Atlan is much better. It's much, much faster..”

Henrique dos Anjos

Data Quality Engineer at truelogic

[Read full review](#) 

Use Case

The main use case for Data Hub is using HubSpot for investing from our data by utilizing AWS service. Each year, we invest based on the data files. A specific example of how my team uses Data Hub in our daily work is that our data engineer team leverages Data Hub to check and verify data limits. The PIT tool helps in understanding data mass via metadata and integrates business philosophy to enhance dashboard comprehension. Regularly, Data Hub is utilized to monitor various aspects.

Anhnq Ho

Finance Feedback Committee at MB Shinsei Finance Limited Liability Company

[Read full review](#) 

“My main use case for Data Hub is for data governance, specifically the use of data lineage and data catalog. I use Data Hub for data governance and data lineage in my day-to-day work by checking the schema and the detailed data metadata of a particular database using the data catalog and to spot the hierarchy of the data from one historical level to the present..”

Verified user

Associate Staff Engineer at a tech vendor with 10,001+ employees

[Read full review](#) 

“Since many developers were developing and creating data models, and there were no DDLs left in the company, we had to recreate all the descriptions of tables and clarify which columns meant what. We made it a place where all stakeholders in our company could log in and see which data were used for which data marts, which column values meant for which definitions, and how they were measured. We primarily used Data Hub for sharing information and increasing the data literacy of our company..”

Verified user

[Read full review](#) 

Data Engineer at a consultancy with 11-50 employees

“Our main use case for Data Hub is to build our enterprise data catalog within Visa using the open-source version.

We use Data Hub to build pipelines and to construct our enterprise data catalog to see where the data is coming from, how the lineage flows, where the lineage of the data originates from, and how the metadata propagation occurs. With this metadata information and the description of all the fields, we have built a layer on top of this that performs natural language querying for people to find where and how the data comes from..”

Somashekar Venkataramaiah

[Read full review](#) 

Senior Technical Program Manager at a financial services firm with 10,001+ employees

“My main use case for Data Hub is one of the programs at Société Générale to migrate from the existing Oracle database to Azure cloud architecture, migrating the entire agency business from legacy to the target architecture.

In terms of data flow on the old system, there were many constraints with respect to the new processes that had to be onboarded due to system constraints. There was a four-year fee in the millions just to maintain the services of that particular database. To eliminate all these limitations, a project team was created to migrate all the processes and flows from the existing Oracle database to the Azure cloud architecture. The advantages of Azure are that you only pay for what you use. Additionally, we can streamline the resources and virtual machines that need to be used for a particular process, and security comes built in. This is why we moved to the Azure architecture.

Our primary objective was to ensure that Data Hub would be useful for multiple consumers, meaning the data we produce can be used by multiple consumers who can build their own process flows and architectures. The basic and main data is fed by the respective upstream applications, which in Société Générale terminology we call deal processing applications. These particular applications publish their data, and along with this main data set, we enrich it with other referentials. Consumers receive the required data along with the main data through these referentials, allowing them to build their processes with this one particular data set rather than subscribing to multiple applications. This is how Data Hub use case is currently maintained..”

Verified user[Read full review](#) 

Lead Business Analyst at a tech vendor with 10,001+ employees

“I provide services related to Data Hub and data management. In total, I can consider my career using Data Hub to be approximately two to two and a half years. I tried to use it on several occasions in previous experiences, and I even started an initiative in my team to have it installed and used broadly in the company because everyone was always lost when it comes to data. They didn't know where the data comes from, which transformation it suffered, and where that data lands. These types of questions were always being asked, and I had the idea to use Data Hub, the open-source version, to give people these analytical capabilities and also integrate with Data Catalog and glossary. I tried to lead that initiative in the past, and then I left the company, so I couldn't see what the end result was.

Right now, Data Hub is already in place, and I use it sometimes, not too frequently because it's maintained by another team, but I use it as an end-user. My company is migrating off the open-source version and is moving to the Atlan version, and I think they use the same tool, but Atlan has more functionalities because they contribute to their private repository.

I used some built-in catalogs on Snowflake, for example, and Databricks, which serve well for that purpose inside their tools. In those cases, you could really see the lineage in Databricks or Snowflake, but at that time, there was no global capability. When we were evaluating Data Hub, there were other companies we were also checking, such as proprietary ones, including Atlan. Although I don't remember all of them, I haven't used those other private companies enough to be confident in my statements about the competitors..”

Igor Uchoa

Senior Data Engineer at a tech services company with 1-10 employees

[Read full review](#) 

Setup

The setup process involves configuring and preparing the product or service for use, which may include tasks such as installation, account creation, initial configuration, and troubleshooting any issues that may arise. Below you can find real user quotes about the setup process.

“I installed Data Hub on an EC2 instance and, proceeding with the database installation, accessed the complete metadata using data catalog and data lineage..”

Verified user

[Read full review](#) 

Associate Staff Engineer at a tech vendor with 10,001+ employees

“I think the initial deployment was straightforward. When I used it, there were Helm charts already available, so we just had to tweak some minor things, and that was it; it was quite easy..”

Igor Uchoa

[Read full review](#) 

Senior Data Engineer at a tech services company with 1-10 employees

“During implementation, the documentation and support resources from Data Hub were very helpful. I followed the guidelines, accessed each section, and understood the platform effectively, which made the initial setup easy..”

Gbytyqi Gbytyqi

[Read full review](#) 

Software Engineer at Reveal data cooperation

“Data Hub exceeds expectations in user-friendliness and functionality. It features a great user interface, an available SDK, APIs, and GraphQL previews, all complemented by a responsive Slack community and helpful customer support. The ease of documentation, website usability, and setup contributes to its overall effectiveness..”

Akashkhurana Hirana

[Read full review](#) 

Senior Software Engineer 2 at Porch

“The initial deployment of Data Hub itself was easy because we could get the Docker Compose YAML file and run the Docker command to get the service up. However, since Data Hub uses various types of components, we had trouble assigning each server for each component and connecting them via network. We overcame that problem..”

Verified user

[Read full review](#) 

Data Engineer at a consultancy with 11-50 employees

“The setup of Data Hub was quite straightforward. One aspect of the architecture I appreciated is that Data Hub relies heavily on Cron jobs and jobs in Kubernetes. Whenever it needs to fix something, it initiates a job to repair its MySQL or its Elasticsearch. Operationally, I find it to be an excellent service, as they worked well on that aspect with the open-source version. However, the lack of support for KRAFT out of the box was somewhat problematic..”

Alireza Khorami

[Read full review](#) 

Data Engineer at Snappfood

Customer Service and Support

The customer support for Data Hub is robust. I had full support and did not use it extensively, relying primarily on Slack for questions and the documentation, which was sufficient since I utilized the open-source version.

Chakib Bekhouche

Consultant at jems groupe

[Read full review](#) 

“We were fortunate not to need customer service, but we did not know that technical support for Data Hub was possible. If we had been aware of that, we could have asked for support or guidance..”

Verified user

Data Engineer at a consultancy with 11-50 employees

[Read full review](#) 

“We have not gone that far with customer support; as far as the POC is concerned, we received good support from the team and the sales team that helped us evaluate the tool..”

Azhagarasan Annadorai

Director at a university with 1-10 employees

[Read full review](#) 

“Customer support for Data Hub is very genuine, and they are responsive and attentive. If I raise a ticket today, they usually respond by the next day. Additionally, they host webinars monthly to discuss new features and updates. They also have an open Slack community where responses tend to be immediate..”

Akashkhurana Hirana

Senior Software Engineer 2 at Porch

[Read full review](#) 

“My support is internal when I have any questions or requests, so I direct it to a support team from Uber and not from the provider. When I was working with Atlan, and needed support, they were very good at attending to my requests directly. I had contact with the provider, so it was very fast. At the moment, I don't have that; I direct my requests to an internal department of Uber..”

Henrique dos Anjos

Data Quality Engineer at truelogic

[Read full review](#) 

“I have had to reach out to Data Hub customer support multiple times. For example, when we were setting up a private link to connect to Data Hub GraphQL APIs, we required our account to be whitelisted. I have also requested some future features for our use cases. For instance, when working with a metadata test scenario, I needed to have a range date column, which was not available. I requested the Data Hub team to make it public so we could use it..”

PrashantGupta2

[Read full review](#) 

Software Engineer L2 at a tech vendor with 5,001-10,000 employees

Other Advice

“I do not have any advice to give to others looking into using Data Hub. I found this interview valuable and do not think anything needs to change for the future. My overall review rating for Data Hub is 8..”

Verified user

Associate Staff Engineer at a tech vendor with 10,001+ employees

[Read full review](#) 

“Currently, I am an end user of Data Hub. I'm not sure about my company's relationship with Data Hub, but personally, I've used it as an enthusiastic user of the Data Catalog subject without any official partnerships. I would rate this product a nine out of ten..”

Igor Uchoa

Senior Data Engineer at a tech services company with 1-10 employees

[Read full review](#) 

My advice for others looking into using Data Hub is that it is a simple way to ingest technical metadata, and it is an open-source solution worth trying out to optimize compared to other enterprise solutions. Regarding Data Hub's governance and security, I think it is a tool that supports the Data Hub team in improving their skills and work. I would rate this product a seven out of ten.

Anhnq Ho

Finance Feedback Committee at MB Shinsei Finance Limited Liability Company

[Read full review](#) 

“I am not familiar with how the data is priced, so I cannot answer that question. Since we were running all the components on one server, there were issues when data injection was occurring too frequently, and sometimes the server could run out of memory or the disk storage could become full. We had maintenance issues that we had to handle in terms of memory and disk storage. We had to alter some injection strategies and timelines so that the data would not grow from too many tables at the same time. These maintenance jobs took place regularly. I would rate this product a 7..”

Verified user

Data Engineer at a consultancy with 11-50 employees

[Read full review](#) 

“For considering the use case of Data Hub, I would specifically say that for constructing a data hub, you should have a variety of consumers. The first requirement is that there should be multiple consumers who would want to access this particular data hub because a data hub is a collection of entire data which can be used by various consumers. The use case is built to help consumers get whatever data they need in one specific place, rather than going to individual applications, subscribing to them, and getting the data they need. That is one of the advantages. The second advantage of Data Hub is that it is more secure and reliable compared to other database architectures in place. I would rate this product an eight out of ten..”

Verified user

Lead Business Analyst at a tech vendor with 10,001+ employees

[Read full review](#) 

“On a scale of one to ten, I rate Data Hub a nine.

I chose nine out of ten because Data Hub is a single solution that we could adopt easily and build our platform on top of it. It provided all the features that we needed, which is why I gave that rating.

Regarding Data Hub's AI capabilities, I have not explored its governance and security features, but I would like to explore them. I have not gone through the AI features of Data Hub concerning the accuracy and reliability of output.

My advice to others looking into using Data Hub is that it is a fantastic tool for people who want to centralize and keep all the data discoverable in one single place. I would highly recommend using it. I give this review an overall rating of nine out of ten..”

Somashekar Venkataramaiah

Senior Technical Program Manager at a financial services firm with 10,001+ employees

[Read full review](#) 

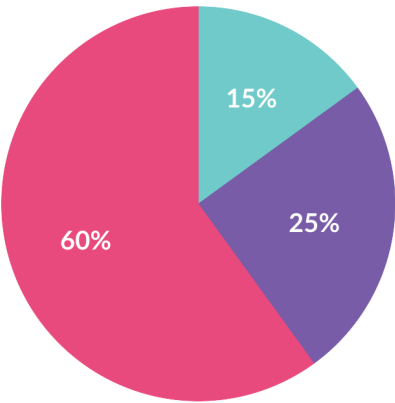
Top Industries

by visitors reading reviews

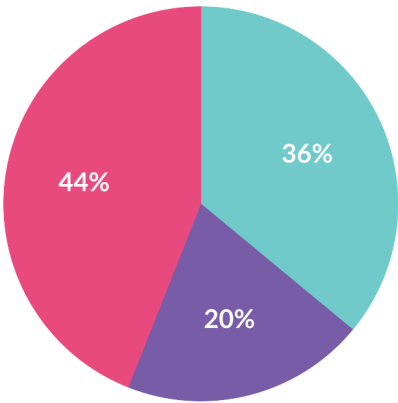


Company Size

by reviewers



by visitors reading reviews



 Large Enterprise  Midsize Enterprise  Small Business

About this buyer's guide

Thanks for downloading this PeerSpot report.

The summaries, overviews and recaps in this report are all based on real user feedback and reviews collected by PeerSpot's team. Every reviewer on PeerSpot has been authenticated with our triple authentication process. This is done to ensure that every review provided is an unbiased review from a real user.

Get a custom version of this report... Personalized for you!

Please note that this is a generic report based on reviews and opinions from the collective PeerSpot community. We offer a [customized report](#) of solutions recommended for you based on:

- Your industry
- Company size
- Which solutions you're already considering

The customized report will include recommendations for you based on what other people like you are using and researching.

Answer a few questions in our short wizard to get your customized report.

[Get your personalized report here](#)

About PeerSpot

PeerSpot is the leading review site for cloud, AI, and business software. We created PeerSpot to provide a trusted platform to share information about software, applications, and services. Since 2012, over 22 million people have used PeerSpot to choose the right software for their business.

PeerSpot helps tech professionals by providing:

- A list of products recommended by real users
- In-depth reviews, including pros and cons
- Specific information to help you choose the best vendor for your needs

Use PeerSpot to:

- Read and post reviews of products
- Access over 30,000 buyer's guides and comparison reports
- Request or share information about functionality, quality, and pricing

Join PeerSpot to connect with peers to help you:

- Get immediate answers to questions
- Validate vendor claims
- Exchange tips for getting the best deals with vendor

Visit PeerSpot: www.peerspot.com

PeerSpot

244 5th Avenue, Suite R-230 • New York, NY 10001

reports@peerspot.com

+1 646.328.1944