About   ┊   Advertise   ┊   Contact

Search this website...   `Searc`

insideBIGDATA
Your Source for AI, Data Science, Deep Learning & Machine Learning Strategies

AI Deep Learning   Machine Learning   Special Sections   White Papers   Special Reports   Topics   Industry Segments   Resources

Sign up for our newsletter and get the latest big data news and analysis.   Your email address   Subscribe

Don't Call It A "Data Product" Unless It Meets These 5 Requirements

June 9, 2022 by Editorial Team   2 Comments

tweet   share   share   share   email

In this special guest feature, Barr Moses, Co-founder and CEO of Monte Carlo, believes data products can transform an organization's ability to be data-driven as long as they meet 5 key requirements. Monte Carlo, the data reliability company, is creator of the industry's first end-to-end data observability platform. Monte Carlo works with such data-driven companies as Fox, Affirm, Vimeo, ThredUp, PagerDuty, and other leading enterprises to help them achieve trust in data. Prior to starting Monte Carlo, Barr was a Vice President at Gainsight where she helped grow the company 10x in revenue and built the customer data and analytics team. Barr has also consulted at Bain & Company, worked at the Statistics Department at Stanford, and served in the Israeli Air Force as a commander of an intelligence data analyst unit. She graduated from Stanford with a B.Sc. in Mathematical and Computational Science.

One of the buzziest movements in the data world currently is treating "data as a product." Almost overnight, data platforms were relabeled as data products in the presentation decks and architecture diagrams of data leaders far and wide.

Data products are an extremely important, and frankly overdue, concept that states data teams need to treat how they surface data within the business with the same rigor that software teams use to develop modern applications.

This rigor is essential given the elevated role of data as it has been unleashed from dashboards to powering machine learning modules, becoming part of customer facing applications, and much more. But as data has become more valuable, the cost of bad data has risen as well. Bad data costs the average company $12.9 million per year according to Gartner.

Not only does the concept of data product raise the bar for the availability, reliability, and usability of enterprise data, but it also represents a profound shift in perspective.

Instead of inside-out thinking, with data teams virtually huddling and tinkering with data before occasionally popping their heads up to relay nuggets of useful insights, data products promote outside-in thinking, with data teams focusing on how to maximize the usability of data for internal customers so those stakeholders can make data-driven decisions.

It's powerful and scalable. It's also a lot of work, and like many trends, not as well defined as it could be.

So in the spirit of product development, here are 5 essential requirements for data products that can help organizations move from buzz to production:

1. **Documentation is readily available**. No, the DAG in your Apache Spark management console doesn't count (although that's important too). User focused documentation detailing what the product is, how to use it, and why it works is an effective first step data teams can take in their data product journey. It forces you to answer basic but

---

INDUSTRY PERSPECTIVES

The Most Common Bias in Data Is Gender Bias — Here's How to Prevent It

In this special guest feature, Michaël Pilaeten, learning and development manager for CTG's European operations, explores why data bias exists, and the impact it creates. Then, he identifies major areas in which the data bias impacts women on daily basis to showcase the level of severity. Lastly, he provides key solutions and strategies combat data bias.

[READ MORE...]

WHITE PAPERS

Three Ways AI Can Boost Conversational Intelligence Across Your Enterprise

Artificial Intelligence can both help and scale human effort in building the conversational intelligence across the enterprise that drives successful business outcomes. In this eBook from Veritone, we will explore three common u cases of AI applied to customer conversations: (contact center insights – to drive improved customer interactions and call center efficiencie (ii) [...]

Download

See More White Papers »

FIND US ON:

elusive questions, "what exactly is the product, is this dashboard a part of it?" It also helps start the key shift your team will need to undergo from, "this makes sense to me" to "will this be useful for my customer?"

2. **Feedback and iteration take center stage**. Don't show up to your next business review and say, "we developed this new data product for you, what do you think?" Start with the very basic, but tragically overlooked step of ensuring there is a problem that your data product needs to solve. Gather customer requirements by asking a lot of questions like, "When do you report on this data?" "How frequently does it need to be updated?" "What business process is this data used to inform?" Find out if the team operates in a system that can display your data within their workflow or if it has the ability to automatically optimize based on your data (reverse ETL can be your friend here). Ship minimum viable products and continue to act on feedback. Don't forget about your internal customers once the product has been successfully deployed! NPS surveys can be a helpful formal feedback mechanism for roadmaps (your data product *does* have a roadmap right?) in addition to less formal meetings on at least a monthly cadence.

3. **There is a specialized support and ownership**. Virtually every application has a product manager, the person responsible for setting the vision and coordinated development. Data products should be no different. While data products and a decentralized data mesh architecture go together like peanut butter and jelly, a decentralized organizational structure is not necessarily a requirement. However, the data product model should be supported by dedicated engineers and analysts with specialized knowledge of the product and the business needs it's addressing.

4. **Service level agreements (SLAs) have been set**. Just as SLAs play a crucial role in setting expectations for application performance, it's becoming a more common practice to set expectations for the performance of data products within a data contract or SLAs. Data observability solutions that proactively monitor, alert, and measure data quality end-to-end play an important role in helping data teams set, report, and meet these SLAs.

5. **Internal customers are actually using it**. Products without customers are just ideas. Adoption is the true measure of your data product's value. Be sure to track your monthly active users and find creative ways to evangelize within your organization. If you are wildly successful, make sure your product is built to scale! While a data warehouse backbone isn't necessarily a requirement for every use case, it will take any scaling concerns off the table.

Congratulations, It's A Data Product

Not every data system or initiative has to, or should be, a data product.

With a common understanding and standards for data products, the label can be reserved and revered as the impressive, high impact accomplishment it is.

Sign up for the free insideBIGDATA newsletter.

Join us on Twitter: @InsideBigData1 – https://twitter.com/InsideBigData1

tweet   share   share   share   email

Filed Under: Big Data, Featured, Google News Feed, Industry Perspectives, Infrastructure, inside SPARK, News / Analysis, Uncategorized

Comments

John Wingate says:
June 10, 2022 at 2:10 pm

Great read. I love the focus on how products without customers are just ideas. With any type of user one track daily user adoption to meet the needs of those users.

Reply

Barr Moses says:
July 21, 2022 at 11:47 am

I agree–there are a lot of ways data teams measure the success of their data products, but perhaps none are more commensurate with value than how many active users they have.

Leave a Comment

Name *

Email *

Website

☐ Notify me of follow-up comments by email.

☐ Notify me of new posts by email.

Post Comment

insideBIGDATA

Your Source for AI, Data Science, Deep Learning & Machine Learning Strategies

About insideBIGDATA

Contact

Advertise with insideBIGDATA

Visit Our Other Site – insideHPC

Terms of Service & Copyright

Privacy Policy

Copyright © 2023