How We Improved Data Discovery for Data Scientists at Spotify

Posted on February 27, 2020 by Andrew Maher (https://labs.spotify.com/author/andrewmaher02ac847b39/)

At Spotify, we believe strongly in data-informed decision making. Whether we're considering a big shift in our product strategy or we're making a relatively quick decision about which track to add to one of our editorially-programmed playlists, data provides a foundation for sound decision making. An insight is a conclusion drawn from data that can help influence decisions and drive change. To enable Spotifiers to make faster, smarter decisions, we've developed a suite of internal products to accelerate the production and consumption of insights. One of these products is Lexikon, a library of data and insights that help employees find and understand the data and knowledge generated by members of our insights community.

We've learned a lot since we first launched this product. In this blog post, we want to share the story of how we iterated on Lexikon to better support data discovery.

Improving the data discovery experience

Diagnosing the problem

In 2016, as we started migrating to the Google Cloud Platform (https://labs.spotify.com/2019/12/09/views-from-the-cloud-a-history-of-spotifys-journey-to-the-cloud-part-1/), we saw an explosion of dataset creation in BigQuery. At this time, we also drastically increased our hiring of insights specialists (data scientists, analysts, user researchers, etc.) at Spotify, resulting in more research and insights being produced across the company. However, research would often only have a localized impact in certain parts of the business, going unseen by others that might find it useful to influence their decision making. Datasets lacked clear ownership or documentation making it difficult for data scientists to find them. We believed that the crux of the problem was that we lacked a centralized catalog of these data and insights resources.

In early 2017, we released Lexikon, a library for data and insights, as the solution to this problem. The first release allowed users to search and browse available BigQuery tables (i.e. datasets)— as well as discover knowledge generated through past research and analysis. The insights community at Spotify was quite excited to have this new tool and it quickly became one of the most widely used tools amongst data scientists, with ~75% of data scientists using it regularly, and ~550 monthly active users.

However, months after the initial launch, we surveyed the insights community and learned that data scientists still reported data discovery as a major pain point, reporting significant time spent on finding the right dataset. The typical data scientist at Spotify works with ~25-30 different datasets in a month. If data discovery is time-consuming, it significantly increases the time it takes to produce insights, which means either it might take longer to make a decision informed by those insights, or worse, we won't have enough data and insights to inform a decision.

Recent Posts

How We Gave Superpowers to Our macOSCI (https://labs.spotify.com/2020/05/01/how-we-gave-superpowers-to-our-macos-ci/)

How We Use Backstage atSpotify (https://labs.spotify.com/2020/04/21/how-we-use-backstage-at-spotify/)

My Beat: AnnClifton (https://labs.spotify.com/2020/04/17/my-beat-ann-clifton/)

Introducing the Spotify Podcast Dataset and TREC Challenge2020 (https://labs.spotify.com/2020/04/16/introducing-the-spotify-podcast-dataset-and-trec-challenge-2020/)

Reach for the Top: How Spotify Built Shortcuts in Just SixMonths (https://labs.spotify.com/2020/04/15/reach-for-the-top-how-spotify-built-shortcuts-in-just-six-months/)

When should I write an Architecture DecisionRecord? (https://labs.spotify.com/2020/04/14/when-should-i-write-an-architecture-decision-record/)

My Beat: StefanÅlund (https://labs.spotify.com/2020/04/01/my-beat-stefan-alund/)

Spotify's New Podcast API: From Design toLaunch (https://labs.spotify.com/2020/03/20/spotifys-new-podcast-api-from-design-to-launch/)

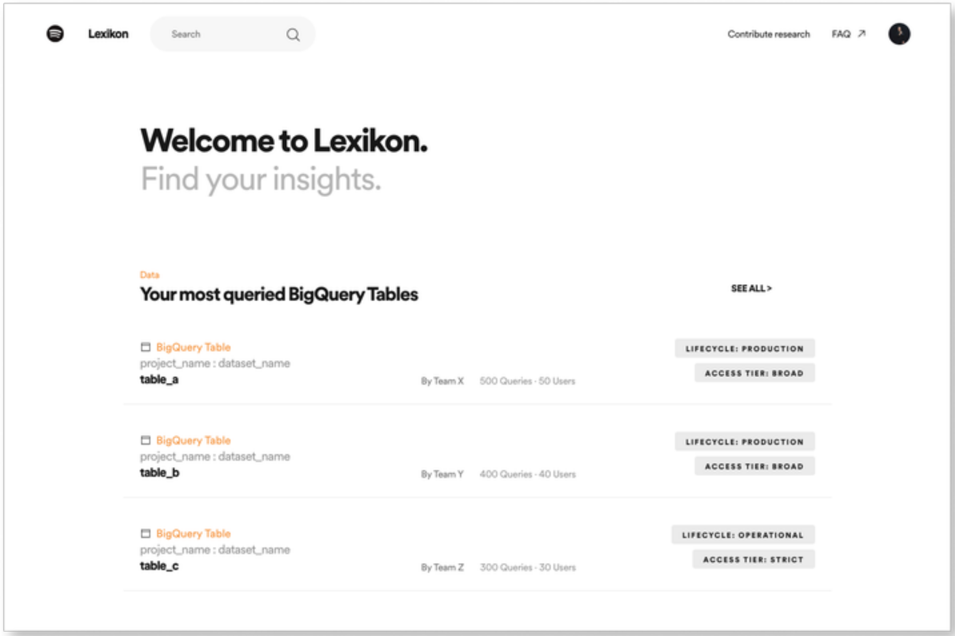
My Beat: TonyJebara (https://labs.spotify.com/2020/03/18/my-beat-tony-jebara/)

What the Heck is BackstageAnyway? (https://labs.spotify.com/2020/03/17/what-the-heck-is-backstage-anyway/)

Our team decided to focus on this specific issue by iterating on Lexikon, with the goal to improve the data discovery experience for data scientists and ultimately accelerate insights production. We were able to significantly improve the data discovery experience by (1) gaining a better understanding of our users intent, (2) enabling knowledge exchange through people, and (3) helping users get started with a dataset they've discovered.

Understanding Intent

To kick things off, we spent time conducting user research to learn more about our users, their needs, and their specific pain points regarding data discovery. In doing so, we were able to gain a better understanding of our users intent within the context of data discovery, and use this understanding to drive product development.



Welcome to Lexikon.
Find your insights.

Data
Your most queried BigQuery Tables
SEE ALL >

BigQuery Table
project_name : dataset_name
table_a
By Team X 500 Queries · 50 Users
LIFECYCLE: PRODUCTION
ACCESS TIER: BROAD

BigQuery Table
project_name : dataset_name
table_b
By Team Y 400 Queries · 40 Users
LIFECYCLE: PRODUCTION
ACCESS TIER: BROAD

BigQuery Table
project_name : dataset_name
table_c
By Team Z 300 Queries · 30 Users
LIFECYCLE: OPERATIONAL
ACCESS TIER: STRICT

(https://spotifylabscom.files.wordpress.com/2020/02/am_1-2.png)

Low-intent data discovery

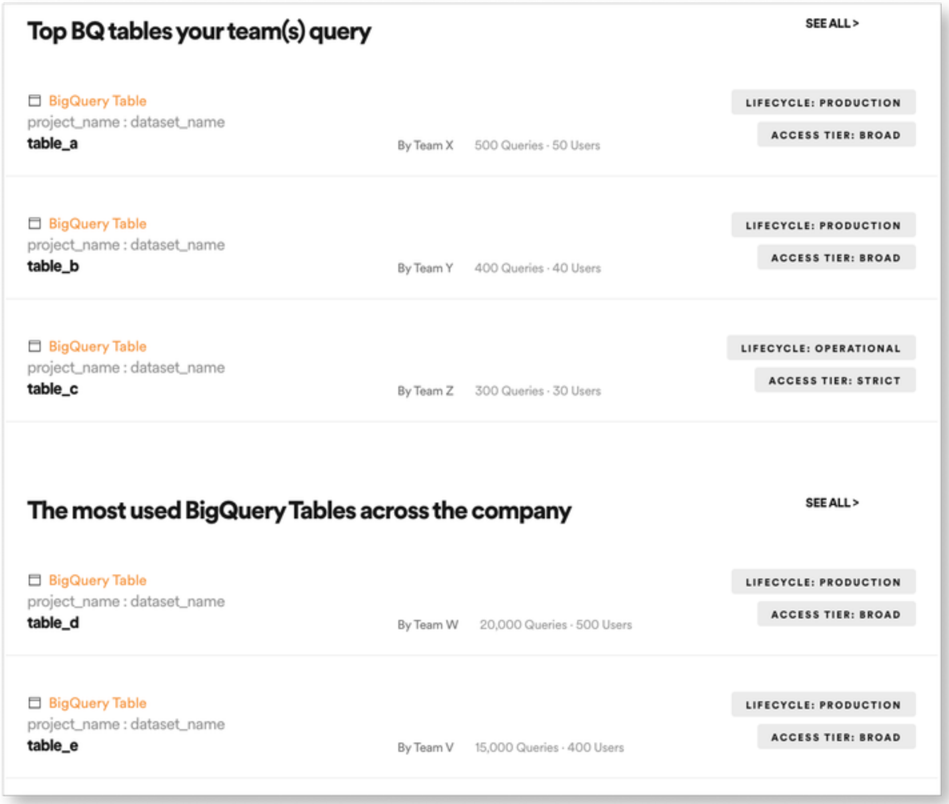
Let's say you're having a rough day and you want to listen to some music to lift your spirit. So, you open up Spotify, browse some of the mood playlists, and put on the Mood Booster (https://open.spotify.com/playlist/37i9dQZF1DX3rxVfibe1L0?si=Y2IPVkn5Qm6Ptb2qI6_OAg) playlist. You've just had a low-intent discovery experience! You had some broad goal to lift your mood and you didn't have extremely strict requirements on what you wanted to listen to.

@SpotifyEng on Twitter

Tweets by @SpotifyEng

Spotify Engineering
@SpotifyEng
Developer Advocate, @joshubrown, talks about the Spotify Web API, including new improvements, developer apps, and the most requested feature by our developer community

Embed View on Twitter



Top BQ tables your team(s) query

SEE ALL >

□ BigQuery Table
project_name : dataset_name
table_a
By Team X 500 Queries · 50 Users
LIFECYCLE: PRODUCTION
ACCESS TIER: BROAD

□ BigQuery Table
project_name : dataset_name
table_b
By Team Y 400 Queries · 40 Users
LIFECYCLE: PRODUCTION
ACCESS TIER: BROAD

□ BigQuery Table
project_name : dataset_name
table_c
By Team Z 300 Queries · 30 Users
LIFECYCLE: OPERATIONAL
ACCESS TIER: STRICT

The most used BigQuery Tables across the company

SEE ALL >

□ BigQuery Table
project_name : dataset_name
table_d
By Team W 20,000 Queries · 500 Users
LIFECYCLE: PRODUCTION
ACCESS TIER: BROAD

□ BigQuery Table
project_name : dataset_name
table_e
By Team V 15,000 Queries · 400 Users
LIFECYCLE: PRODUCTION
ACCESS TIER: BROAD

(https://spotifylabscom.files.wordpress.com/2020/02/am_2-1.png)

Within the context of data discovery, a data scientist with low-intent has a broad set of goals and might not be able to identify exactly what it is they're looking for. This mode of discovery is particularly important for new employees or for people who are starting on a new project or team. For example, as a data scientist, I may want to:

- find popular datasets used widely across the company,
- find datasets that are relevant to the work my team is doing, and/or
- find datasets that I might not be using, but I should know about.

In order to satisfy the needs of low-intent data discovery, we revamped the homepage of Lexikon to serve personalized dataset recommendations to users. The homepage provides users with a number of potentially relevant, algorithmically generated suggestions for datasets including:

- popular datasets used widely across the company,
- datasets you've recently used,
- datasets used widely by the teams to which you belong, and
- recommendations for datasets you haven't used, but might find useful.

While we did experiment with more advanced methods for serving recommendations, including using natural language processing and topic modeling on the dataset metadata to provide content-based recommendations, we determined through user feedback that relatively simple heuristics leveraging data consumption statistics worked quite well. In the first version of Lexikon, most traffic to BigQuery table pages was driven by search. After making these changes, we now see that 20% of monthly active users navigate to BigQuery tables through personalized recommendations on the homepage.

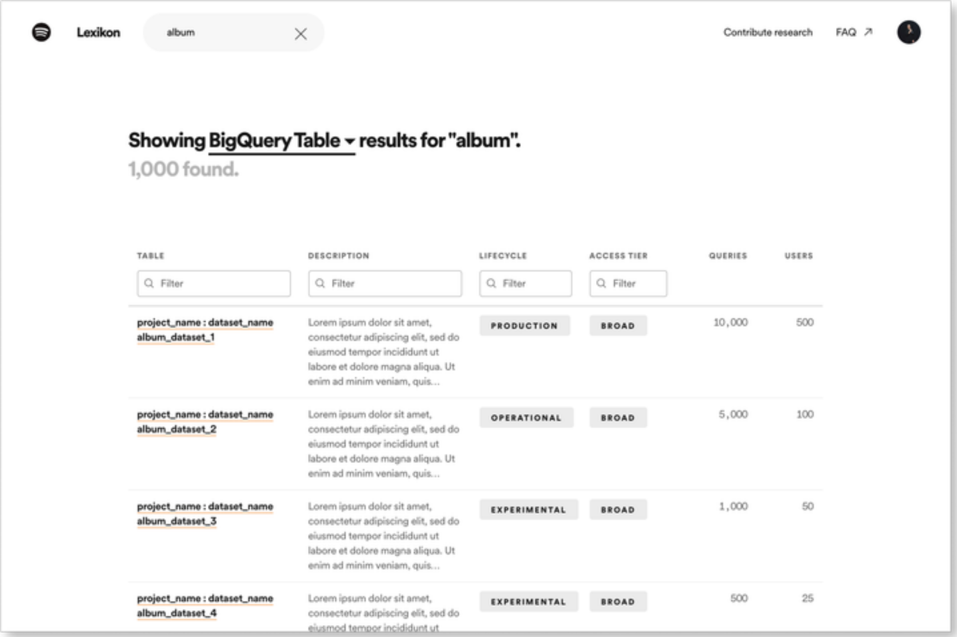
High-intent data discovery

You're walking down the street and hear a passing car blasting a great song you haven't heard in a while. You can't get the song out of your head and need to listen to it immediately. So you pull up Spotify on your phone, search for the track, and play it (on repeat). You've just had a high-intent discovery!

A data scientist with high-intent has a specific set of goals and can likely articulate exactly what they're looking for. This mode of discovery is often more important to more tenured data scientists who may be familiar with some datasets, but may be looking for something they haven't used before that meets a certain set of criteria. For example, as a data scientist with high-intent, I may want to:

- find a dataset by its name,
- find a dataset that contains a specific schema field,
- find a dataset related to a particular topic,
- find a relevant dataset located in a particular BigQuery project,
- find a dataset that my colleague has used of which I can't remember the name, and/or
- find the top datasets that a team has used because I'm collaborating on a new project with them.

To better serve the use case of high-intent data discovery, we iterated on the search experience. First, we focused on the search ranking algorithm. We learned through data analysis that although we have tens of thousands of datasets on BigQuery, the majority of consumption occurred on a relatively small share of top datasets. Data scientists in a high-intent mode of discovery were often looking for one of these top used datasets that met their needs. So, we adjusted our search algorithm to weight search results more heavily based on popularity. Following this change, in user feedback sessions, data scientists reported that the search results not only seemed more relevant, but they were also more confident in the datasets they discovered because they were able to see the dataset they found was used widely by others across the company.

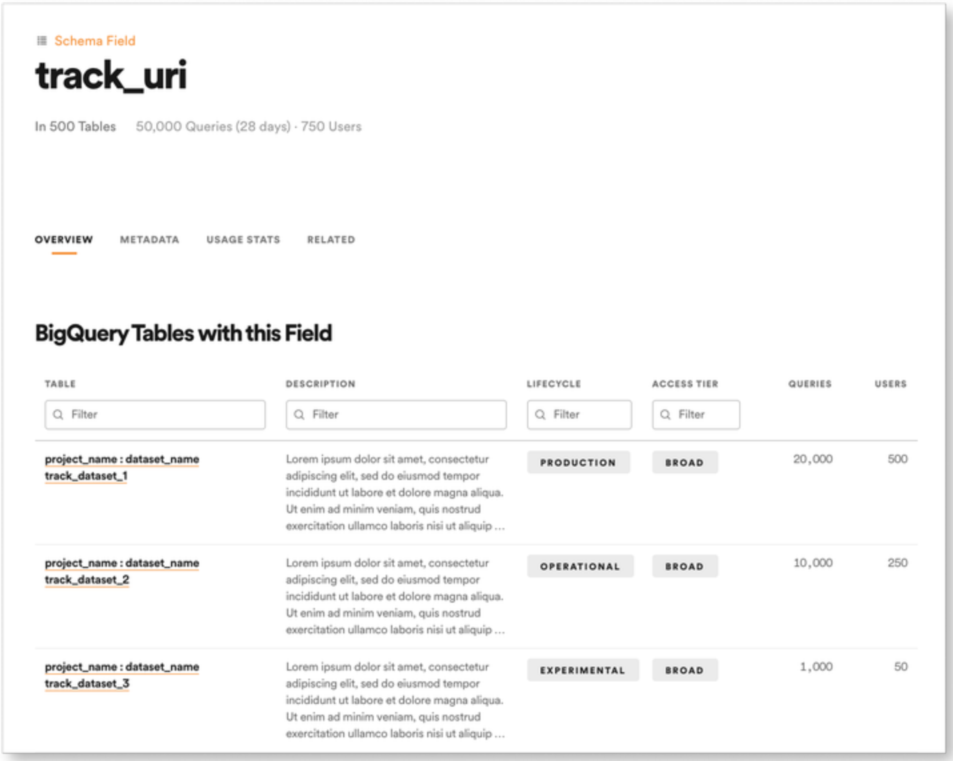


Showing BigQuery Table ▾ results for "album".
1,000 found.

TABLE	DESCRIPTION	LIFECYCLE	ACCESS TIER	QUERIES	USERS
Filter	Filter	Filter	Filter		
project_name : dataset_name album_dataset_1	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis...	PRODUCTION	BROAD	10,000	500
project_name : dataset_name album_dataset_2	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis...	OPERATIONAL	BROAD	5,000	100
project_name : dataset_name album_dataset_3	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis...	EXPERIMENTAL	BROAD	1,000	50
project_name : dataset_name album_dataset_4	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut	EXPERIMENTAL	BROAD	500	25

(https://spotifylabscom.files.wordpress.com/2020/02/am_3.png)

In addition to improving the search rank, we also introduced new types of entities (e.g. schema field, BigQuery project, person, team, etc.) to Lexikon to better represent the landscape of insights production. Our belief was that by making these types of entities more explorable, we would open up new pathways for data discovery. For example, a data scientist might be looking for the best dataset to use that contains a track's URI (https://community.spotify.com/t5/Spotify-Answers/What-s-a-Spotify-URI/ta-p/919201) track_uri. In this case, a user can search for "track uri", navigate to the "track_uri" schema field page, see the top BigQuery tables that contain the schema field, and navigate to the dataset page. In addition to the schema field page, we've added BigQuery Project, people, and team pages, which can serve as a similar stepping stone on the pathway to data discovery. Since launching these new entity pages, we've seen that they've proven to be a critical pathway for discovery, with 44% of Lexikon's monthly active users visiting these types of pages.



track_uri

In 500 Tables 50,000 Queries (28 days) · 750 Users

BigQuery Tables with this Field

TABLE	DESCRIPTION	LIFECYCLE	ACCESS TIER	QUERIES	USERS
Filter	Filter	Filter	Filter		
project_name : dataset_name track_dataset_1	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip …	PRODUCTION	BROAD	20,000	500
project_name : dataset_name track_dataset_2	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip …	OPERATIONAL	BROAD	10,000	250
project_name : dataset_name track_dataset_3	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip …	EXPERIMENTAL	BROAD	1,000	50

(https://spotifylabscom.files.wordpress.com/2020/02/am_4.png)

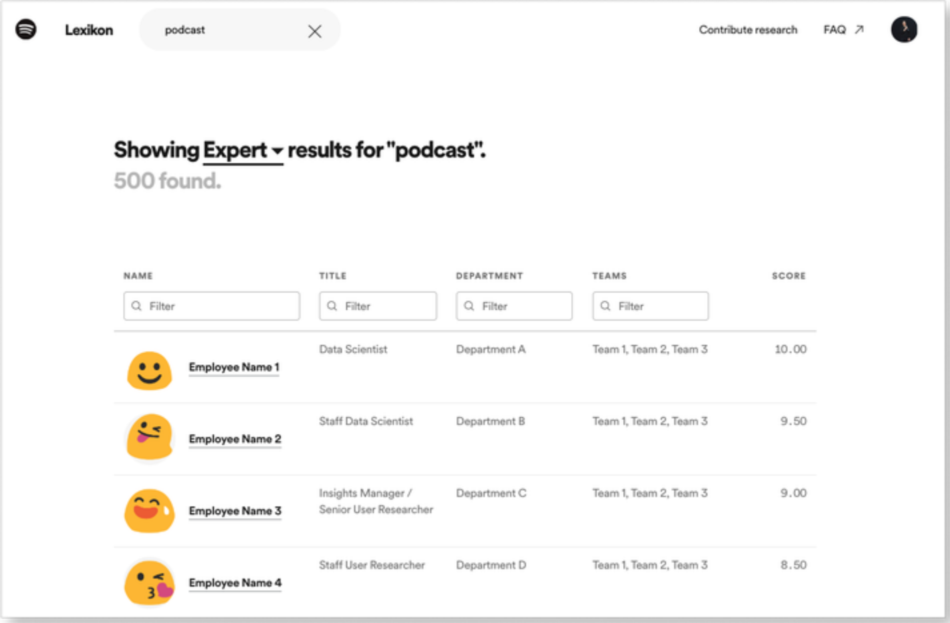
Enabling knowledge exchange through people

Imagine you're starting to explore the genre of jazz. You happen to notice that your coworker has a jazz album on Spotify pulled up on her desktop screen. You strike up a conversation and learn that she is a jazz aficionado. She has become your new genre guide. We've found that there are similar opportunities for people-to-person knowledge exchange with data discovery.

In addition to using learnings from user surveys, feedback sessions, and exploratory analysis to drive product development, we also conducted research on knowledge management theory to better understand how we might adjust our approach (recommended reading: Knowledge Management in Organizations: a critical introduction (https://www.amazon.com/Knowledge-Management-Organizations-critical-introduction/dp/0198724012) by Hislop, Bosua, and Helms).

With the first iteration of Lexikon, we used the knowledge management strategy of codification, which is based on the objectivist perspective of knowledge. This perspective assumes that knowledge can take the form of a discrete entity and can be separated from the people who understand and use it. In the case of Lexikon, we initially believed that if data producers did a great job describing their datasets there would be little-to-no need for person-to-person knowledge exchange. However, in reality, while the first iteration of Lexikon reduced the need for person-to-person knowledge exchange in discovery contexts, there were still instances in which people found it useful to connect with others to find the right data. Rather than fight this, we decided to embrace the idea by (1) mapping expertise within the insights community and (2) providing supplemental information in collaboration tools.

Mapping Expertise



Showing Expert ▾ results for "podcast".
500 found.

NAME	TITLE	DEPARTMENT	TEAMS	SCORE
Filter	Filter	Filter	Filter	
Employee Name 1	Data Scientist	Department A	Team 1, Team 2, Team 3	10.00
Employee Name 2	Staff Data Scientist	Department B	Team 1, Team 2, Team 3	9.50
Employee Name 3	Insights Manager / Senior User Researcher	Department C	Team 1, Team 2, Team 3	9.00
Employee Name 4	Staff User Researcher	Department D	Team 1, Team 2, Team 3	8.50

(https://spotifylabscom.files.wordpress.com/2020/02/am_5.png)

Through user research, we learned that data scientists who failed to discover the data they were looking for would often fall back to finding an expert in the insights community on a given topic and connecting with them in person or online. However, in some cases, data scientists found it difficult to find the right person to talk to about a particular topic. This was especially true for new employees who hadn't yet built personal connections with members of the insights community. So, we introduced a feature in Lexikon that allows you to search for people working in the data and insights space related to a given keyword (i.e. "experts"). These results are powered by summarizing an employee's insight production and consumption activity related to the given keyword. For example, an employee who queries/owns datasets, views/owns dashboards, authors research reports, and/or runs A/B test experiments related to the given keyword will be returned in the list of results. More weight is given to actions related to insights production (e.g. owning a dashboard) rather than insights consumption (e.g. viewing a dashboard).

Providing supplemental information in collaboration tools

Following the release of the first version of Lexikon, we found that data scientists continued to talk with each other about datasets in Slack. Rather than discourage this discussion, we felt like we could help improve the person-to-person knowledge exchange by providing supplemental information. So, we built a Lexikon Slack Bot to improve discussions about datasets. When a user shares a link to a dataset in Lexikon, the Slack bot provides a brief summary of the dataset including:

- the name,
- owner,
- description,
- usage stats,
- data lifecycle information,
- access tier,
- an overview of the most used schema fields in the table, and
- links to view more information in Lexikon, request access, or open directly in BigQuery.

Not only does this provide useful information to users in the moment, but it has also helped raise awareness and increase the adoption of Lexikon. Since launching the Lexikon Slack Bot, we've seen a sustained 25% increase in the number of Lexikon links shared on Slack per week.

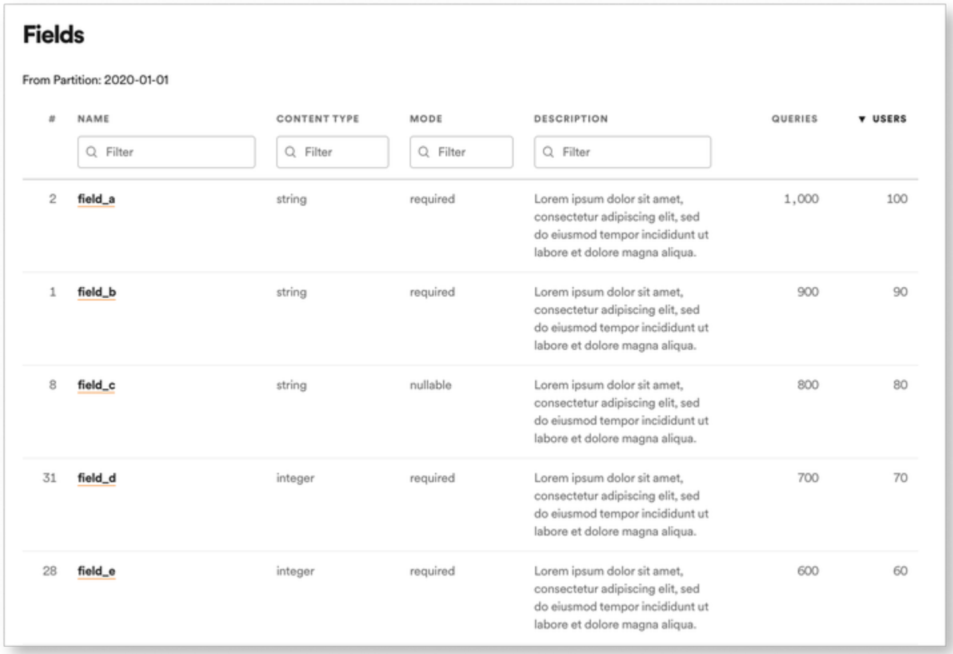
Helping people get started with a dataset they've discovered

You just listened to a track by a new artist on your Discover Weekly and you're hooked. You want to hear more and learn about the artist. So, you go to the artist page on Spotify where you can check out the most popular tracks across different albums, read an artist bio, check out playlists where people tend to discover the artist, and explore similar artists. Using these features on the artist page after your first listen allows you to truly discover and build a connection with the artist. Similar to artist discovery, one of the most critical steps in data discovery is the final step—starting to use the dataset you've discovered.

Through user research, we learned that data scientists would often have a lot of questions about how to start using a dataset, which slowed down their ability to start using the dataset they just discovered. So, we developed the features Schema-field consumption statistics, Queries, and Tables commonly joined to address this last mile of discovery.

Schema-field consumption statistics

Datasets often contain dozens or even hundreds of schema fields. Once you've determined that you've found the right dataset, it can be quite daunting to try to understand all of the available fields and determine which ones are actually relevant. In addition to basic metadata about the schema fields, we included consumption statistics at the schema field level. This shows the number of queries referencing the schema field and the number of unique people who have queried the schema field. This feature gives Lexikon users a way to sort the list of available fields by usage to easily find the ones that are likely to be the most relevant.



Fields

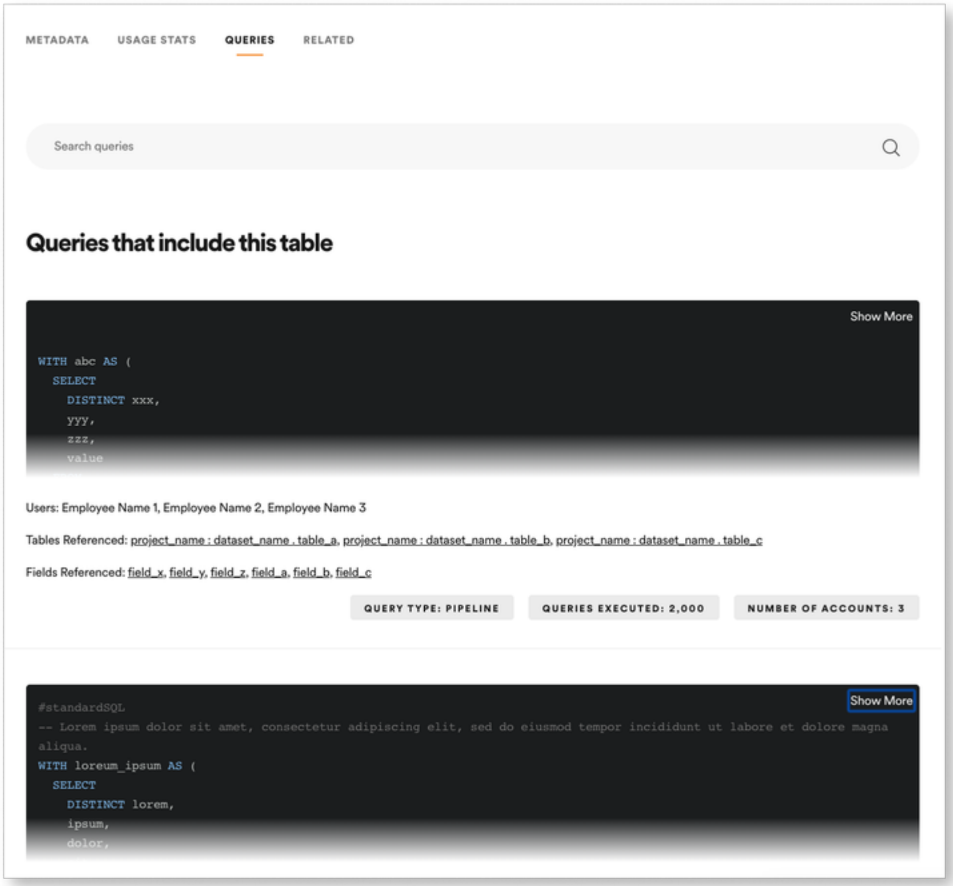
From Partition: 2020-01-01

#	NAME	CONTENT TYPE	MODE	DESCRIPTION	QUERIES	▼ USERS
	Filter	Filter	Filter	Filter		
2	field_a	string	required	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	1,000	100
1	field_b	string	required	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	900	90
8	field_c	string	nullable	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	800	80
31	field_d	integer	required	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	700	70
28	field_e	integer	required	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	600	60

(https://spotifylabscom.files.wordpress.com/2020/02/am_6.png)

Queries

Data scientists are often curious to see how a dataset is actually used in practice. In the first version of Lexikon, we introduced example queries that allowed data producers to submit example queries to give data scientists an idea of how they might use the available dataset. We found there were a few issues with this approach. First, we ran into challenges encouraging data producers to share example queries for all datasets. Second, of the example queries that were submitted, they often became outdated quickly given the ever-changing landscape of data. For example, an example query might be out-dated because it included a join to a deprecated table. So, we abandoned the curated example query and instead allow users to search through all recent queries made on the given dataset. This gives users the opportunity to see a variety of up-to-date queries that use the dataset, and the ability to search for specific queries on the dataset (e.g. "show me queries on this table that reference this specific field"). Since launching this feature, we've seen that 25% of users who visit a dataset page use the queries feature.



Queries that include this table

```
WITH abc AS (
  SELECT
    DISTINCT xxx,
    yyy,
    zzz,
    value
```

Users: Employee Name 1, Employee Name 2, Employee Name 3

Tables Referenced: project_name : dataset_name . table_a, project_name : dataset_name . table_b, project_name : dataset_name . table_c

Fields Referenced: field_x, field_y, field_z, field_a, field_b, field_c

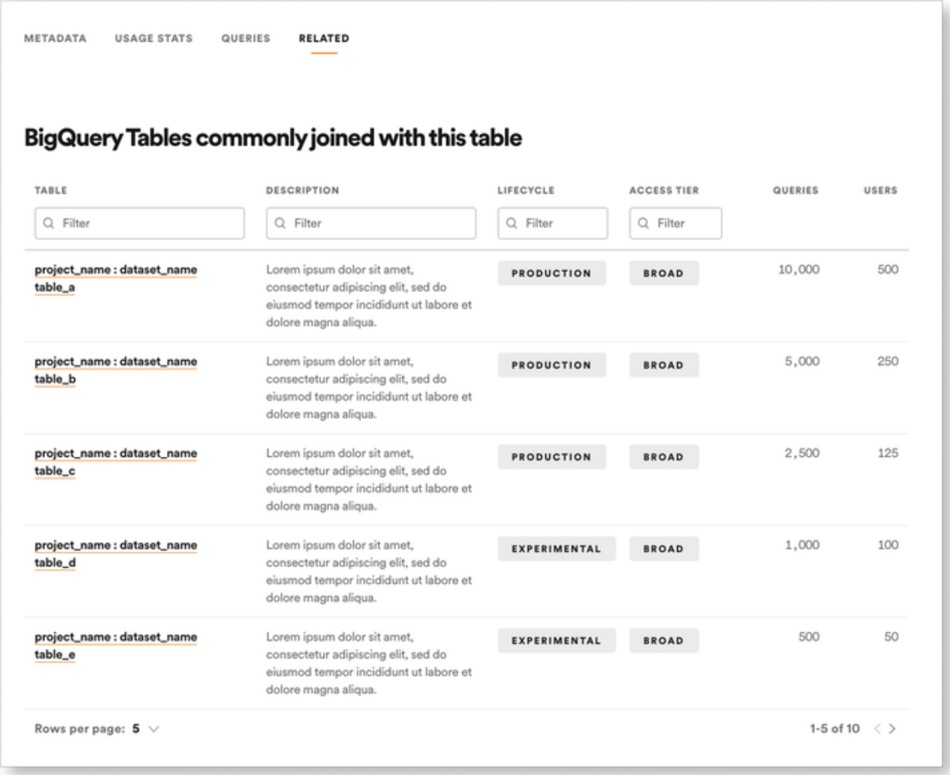
QUERY TYPE: PIPELINE
QUERIES EXECUTED: 2,000
NUMBER OF ACCOUNTS: 3

```
#standardSQL
-- Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna
aliqua.
WITH loreum_ipsum AS (
  SELECT
    DISTINCT lorem,
    ipsum,
    dolor,
```

(https://spotifylabscom.files.wordpress.com/2020/02/am_7.png)

Tables commonly joined

It's rare that a single dataset will contain all of the information for which a data scientist is looking. It's likely the case that they'll need to join a dataset with others in order to answer the question they have. So, we built a feature on a BigQuery table page that allows the user to see tables that are most commonly joined with the given dataset. While this isn't the most widely used feature, we've seen that it is consistently used by 15% of users who visit a dataset page.



BigQuery Tables commonly joined with this table

TABLE	DESCRIPTION	LIFECYCLE	ACCESS TIER	QUERIES	USERS
project_name : dataset_name table_a	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	PRODUCTION	BROAD	10,000	500
project_name : dataset_name table_b	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	PRODUCTION	BROAD	5,000	250
project_name : dataset_name table_c	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	PRODUCTION	BROAD	2,500	125
project_name : dataset_name table_d	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	EXPERIMENTAL	BROAD	1,000	100
project_name : dataset_name table_e	Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.	EXPERIMENTAL	BROAD	500	50

(https://spotifylabscom.files.wordpress.com/2020/02/am_8.png)

Final Thoughts

By understanding the user's intent, enabling knowledge exchange through people, and by helping people get started with a dataset they've discovered, we've been able to significantly improve the data discovery experience for data scientists at Spotify.

Since making these improvements to the data discovery experience in Lexikon we see that adoption of Lexikon amongst data scientists has increased from 75% to 95%, putting it in the top 5 tools used by data scientists. For comparison, more people report using Lexikon than BigQuery UI, Python, or Tableau at Spotify. Lexikon's user base has organically grown from ~550 to ~870 monthly active users as it has proven to be useful to data consumers in non-insights specialist roles (e.g. engineers, data-savvy product managers, etc.). We've also seen a significant increase in engagement with the average number of sessions per MAU increasing from ~3 to ~9 since our initial launch. In addition to these encouraging adoption and engagement metrics, we've learned from surveying data scientists that after making these improvements data discovery is no longer identified as a primary pain point in insights production.

Join the band

If you're interested in helping us tackle similar problems or you're a data scientist that's looking to work at a company where producing impactful insights is becoming easier every day, visit the Join the Band (https://www.spotifyjobs.com/) page to view open roles.

Acknowledgments

We've had a number of folks help get this product to where it is today. Shout out to our current team (Ambrish Misra, Bastian Kuberek, Beverly Mah, David Lau, Erik Fox, and Nithya Muralidharan) and others who have contributed to Lexikon (Adam Bly, Aliza Aufrichtig, Colleen McClowry, David Riordan, Edward Lee, Luca Masud, Mark Koh, Molly Simon, Mindy Yuan, Niko Stahl, and Tianyu Wu).

Andrew Maher (https://www.linkedin.com/in/andrewjohnmaher) is a Product Manager for Spotify's Insights Platform Product Area.

Like 107 people like this. Sign Up to see what your friends like.

Share this:

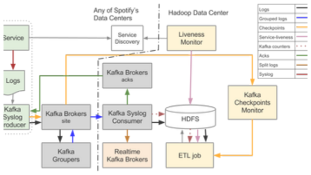
Twitter (https://labs.spotify.com/2020/02/27/how-we-improved-data-discovery-for-data-scientists-at-spotify/?share=twitter&nb=1)

Facebook (https://labs.spotify.com/2020/02/27/how-we-improved-data-discovery-for-data-scientists-at-spotify/?share=facebook&nb=1)

LinkedIn (https://labs.spotify.com/2020/02/27/how-we-improved-data-discovery-for-data-scientists-at-spotify/?share=linkedin&nb=1)

Email (https://labs.spotify.com/2020/02/27/how-we-improved-data-discovery-for-data-scientists-at-spotify/?share=email&nb=1)

Related



Spotify's Event Delivery - The Road to the Cloud (Part I) (https://labs.spotify.com/2016/02/25/spotifys-event-delivery-the-road-to-the-cloud-part-i/)
In "Labs"

How we use Python at Spotify (https://labs.spotify.com/2013/03/20/how-we-use-python-at-spotify/)

(https://labs.spotify.com/2013/03/20/how-we-use-python-at-spotify/)

The most frequent question we heard at PyCon this weekend, was how do we use Python at Spotify. Hopefully this post answers the question! At Spotify the main two
In "Labs"

Spotify
Welcome to the end of a decade
Ready for your 2019 Wrapped?
Relive and discover the artists, music, and podcasts you've loved the mo

Spotify Unwrapped: How we brought you a decade of data (https://labs.spotify.com/2020/02/18/wrapping-up-the-decade-a-data-story/)
In "Labs"

This entry was posted in LABS (HTTPS://LABS.SPOTIFY.COM/CATEGORY/LABS/). Bookmark the permalink (https://labs.spotify.com/2020/02/27/how-we-improved-data-discovery-for-data-scientists-at-spotify/).

← HOW SPOTIFY ALIGNED CDN SERVICES FOR A LIGHTNING FAST STREAMING EXPERIENCE (HTTPS://LABS.SPOTIFY.COM/2020/02/24/HOW-SPOTIFY-ALIGNED-CDN-SERVICES-FOR-A-LIGHTNING-FAST-STREAMING-EXPERIENCE/)

MY BEAT: PER WENDIN → (HTTPS://LABS.SPOTIFY.COM/2020/03/03/MY-BEAT-PER-WENDIN/)

Comments

1. **How We Improved Data Discovery for Data Scientists at Spotify - Ask sendai (https://asksendai.com/how-we-improved-data-discovery-for-data-scientists-at-spotify/)** says:
   February 29, 2020 at 12:44 am (https://labs.spotify.com/2020/02/27/how-we-improved-data-discovery-for-data-scientists-at-spotify/#comment-1624)
   […] Source: spotify […]
2. **Data Science newsletter – March 3, 2020 | Sports.BradStenger.com (https://sports.bradstenger.com/newsletters/data-science/2020/03/03/data-science-newsletter-march-3-2020/)** says:
   March 19, 2020 at 3:47 pm (https://labs.spotify.com/2020/02/27/how-we-improved-data-discovery-for-data-scientists-at-spotify/#comment-1665)
   […] How We Improved Data Discovery for Data Scientists at Spotify […]
3. **5 Things Business Leaders Need to Know About Data Strategy – SchoolOfPython**