

Research Insights

user research from cloud data services

Damien Solomon | Design Manager | CDS

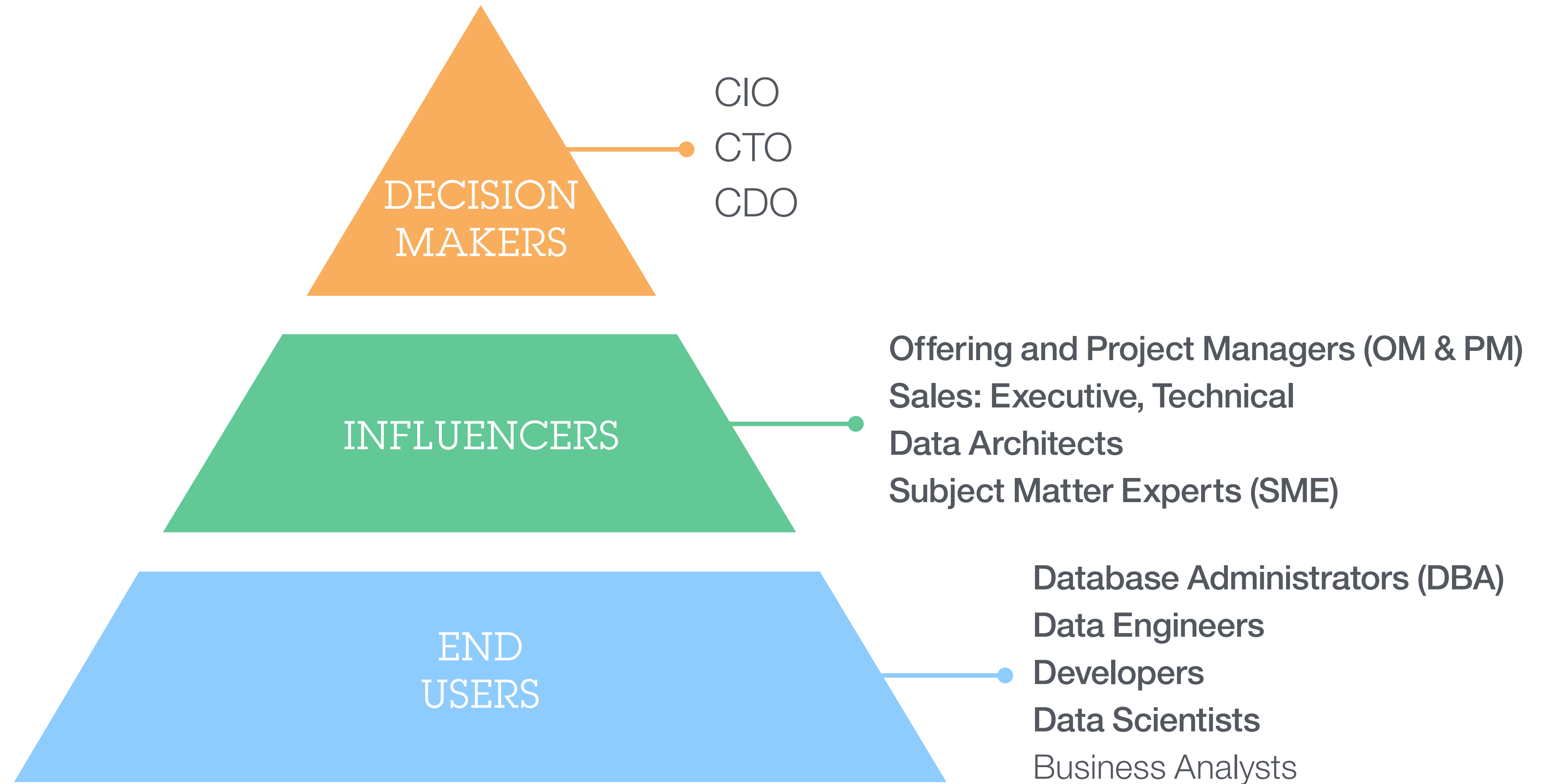
Bhavika Shah | Offering Manager | CDS

Virginia Honig | Researcher | CDS

Our CDS user groups

Stakeholder Map

Breakdown

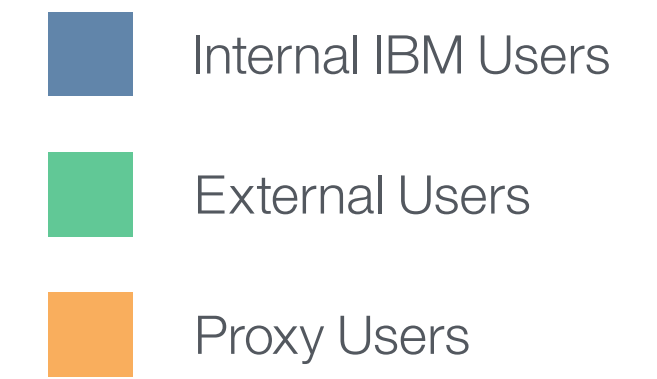


People interviewed
Need to reach out to

who we have
talked to

Interviewees

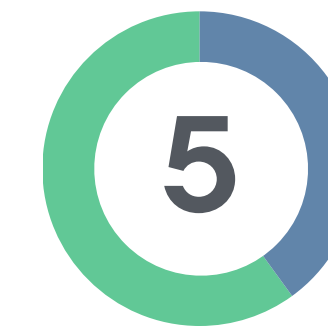
Breakdown



Developers



Sales (Executive and Technical)



Data Scientists



Database Administrators



Subject Matter Experts



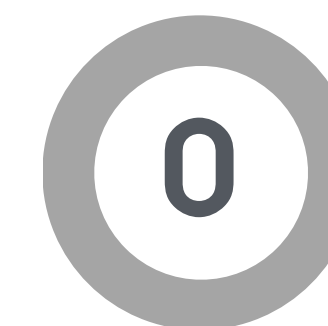
Offering and Product Managers



Data Architects



Data Engineers

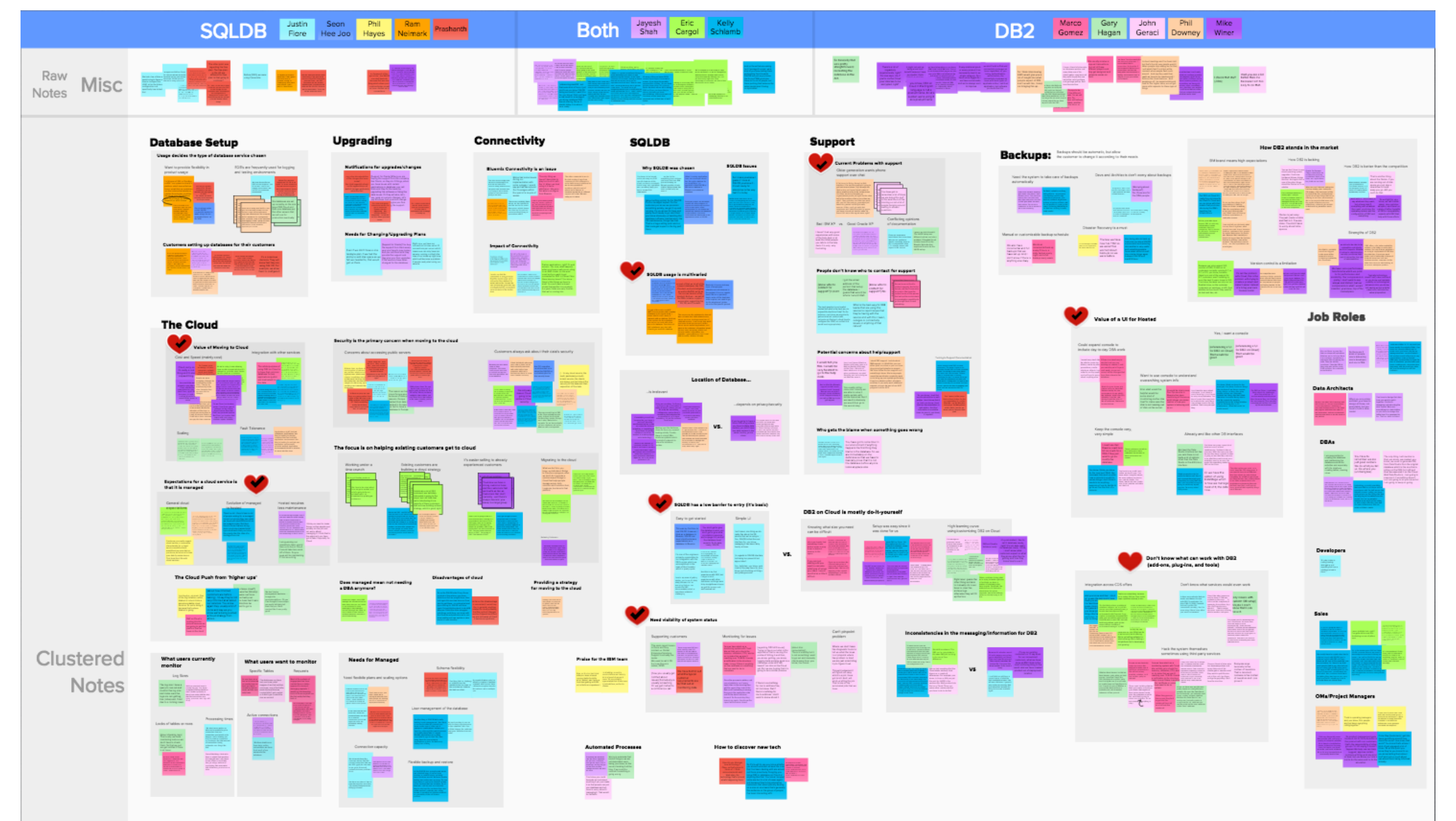


Business Analyst

our methods of research

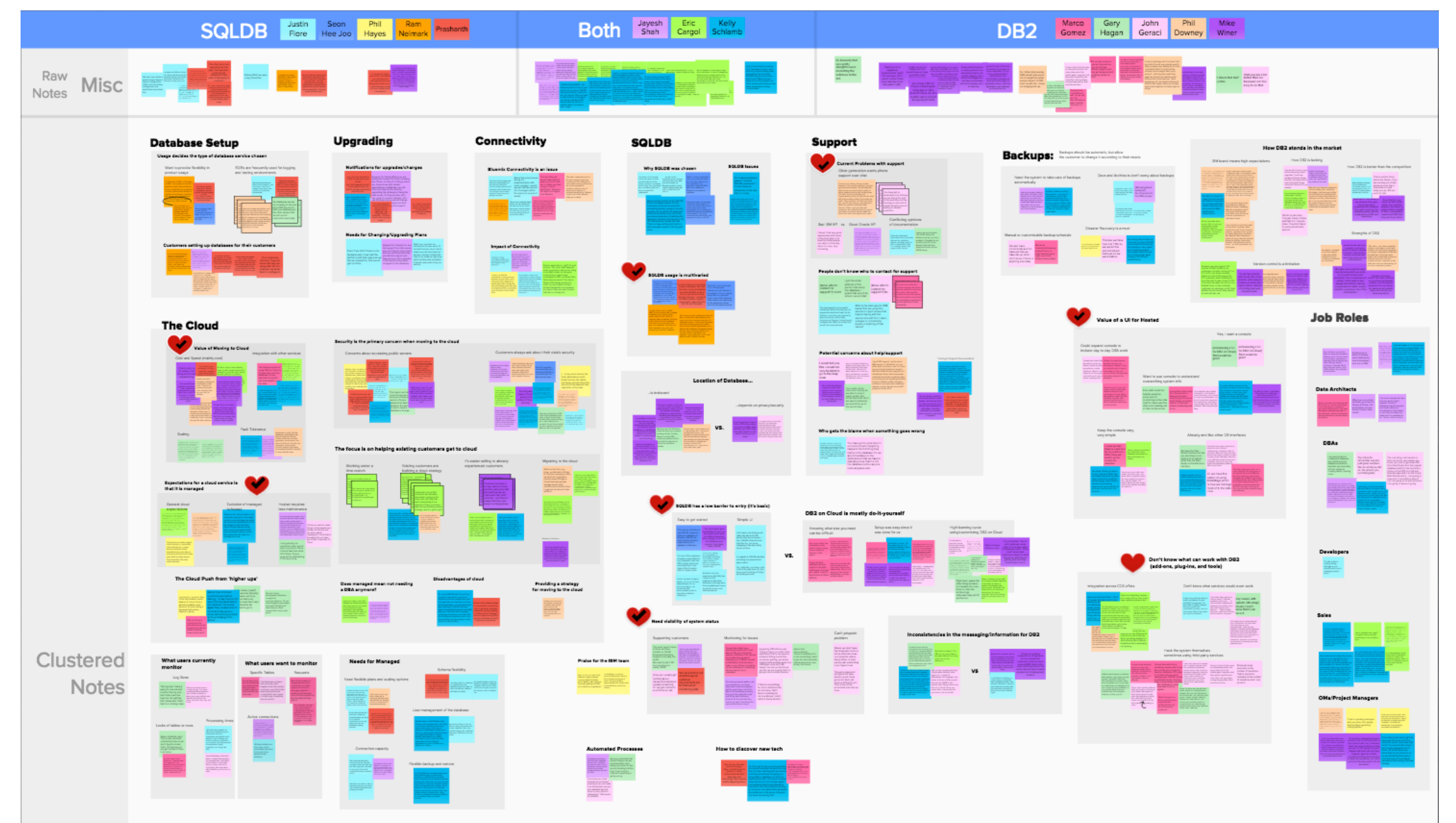
Research Conducted

- 35+ informational interviews
- 2 CDSX user tests < 3 users
- 1 Apache Spark user test < 4 users
- Competitive analysis
 - AWS, Azure, Heroku
- 2 conferences
 - SciPy, SDGC



Research To-Do

- On-site visits to understand how our users work together
- Team-wide user recruitment to secure sponsor users
- Usability tests with total representations of our users
- User testing of the entire CDS experience



research findings for CDS

User Group Flow

Skill set breakdown

**TRANSACTIONAL
(Database)**

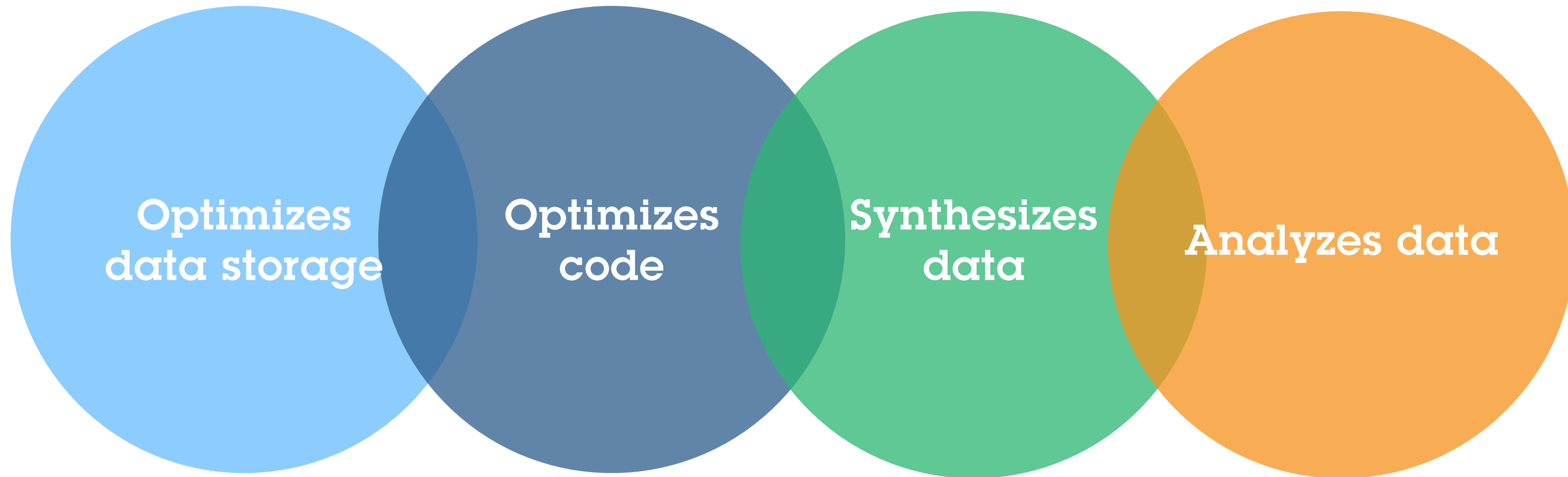
Database
Administrator



Developer



Business Analyst



**ANALYTICAL
(Data Warehouse)**

Data Engineer



Developer



Data Scientist



Business Analyst



Database Administrators

I am responsible for creating the database and maintaining the database and all its activities, especially creating tables and creating views.

What they do

- Sets up database
- Manages: backups, storage, logs, locks, updates
- Troubleshoots issues

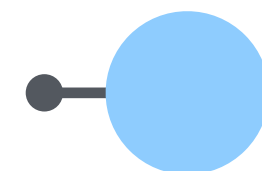
Their concerns

- Scalability
- Reliability
- Performance

Their expectation

A tool that helps them pinpoint issues instead of having to monitor everything

STORAGE



ANALYTICS



Data Engineers

I care enough about the data and have the skills to transform the data into a model that makes sense from a data perspective and do so in a way that is optimized from a systems perspective.

What they do

- Designs, builds and manages analytical data warehouses
- Ensures system performance
- Troubleshoots issues

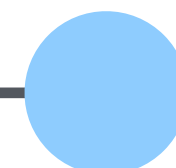
Their concerns

- Scalability
- Performance
- Data quality and structure

Their expectation

Figuring out the best way to structure data based on ever-changing technology, business needs, and data science curiosities

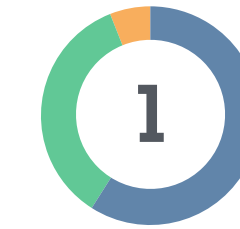
STORAGE



ANALYTICS

Developers

User Group



Interviewed
Developers



Developers

My day to day is coding, testing, debugging, and repeating with some meetings mixed in there.

What they do

- Writes/reviews code
- Fixes bug issues
- Manages/maintains code repository
- Improves application performance

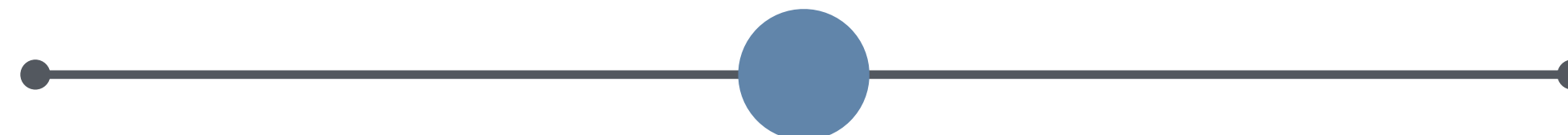
Their concerns

Building and maintaining apps, not dealing with data or analytics

Their expectation

Using products that have a current and active community of users because it saves time in troubleshooting

STORAGE



ANALYTICS



Data Scientist

For data science and analytics the results can be inclusive. Assuming data is correct and you [can] analyze it.

What they do

- Draws insights from big data to help inform business strategy
- Build/deploy algorithms for predictive/prescriptive analytics

Their concerns

- Defining the questions
- Getting good data
- Building the best model

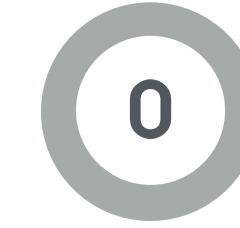
Their expectation

- Fast and reliable access to data
- Less time cleaning data
- Catalog of go-to models and docs

STORAGE



ANALYTICS



Business Analyst

70% of our time is spent cleaning and collecting [data], and only 30% is spent doing analysis and generating insight. In an ideal world, I'd rather have that be 30% collection and 70% analysis.

What they do

- Supports business decisions through data analytics
- Assesses business risks

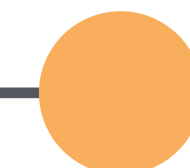
Their concerns

- Workflow issues with IT
- Access to production systems
- Collecting rather than analyzing

Their expectation

- Quick results for analyzing data
- Use of several tools to analyze data
- Provide informed insights from data

STORAGE



ANALYTICS



The reliance on ever-evolving tools and skill-sets can compromise working optimally with data.

Due to

1

Data quality is often uncertain, but the show must go on.

2

Organizing data storage is complex and reliant on ever-changing business goals.

3

Getting the best information out of data can be subjective and dependent on the person not the tool.



Keeping up with the pace of change is overwhelming.

Due to

1

Learning and adopting new tools and practices is risky and time consuming.

2

There is too much information to keep track.

3

Knowing which tools will deliver the most value for the job amidst all of the offerings is challenging.



In the evolving capabilities of big data, users are often functioning as generalists in expert's clothing.

Due to

1

There is a high demand for anyone who can work with data.

2

Learning while doing is the cultural norm.

3

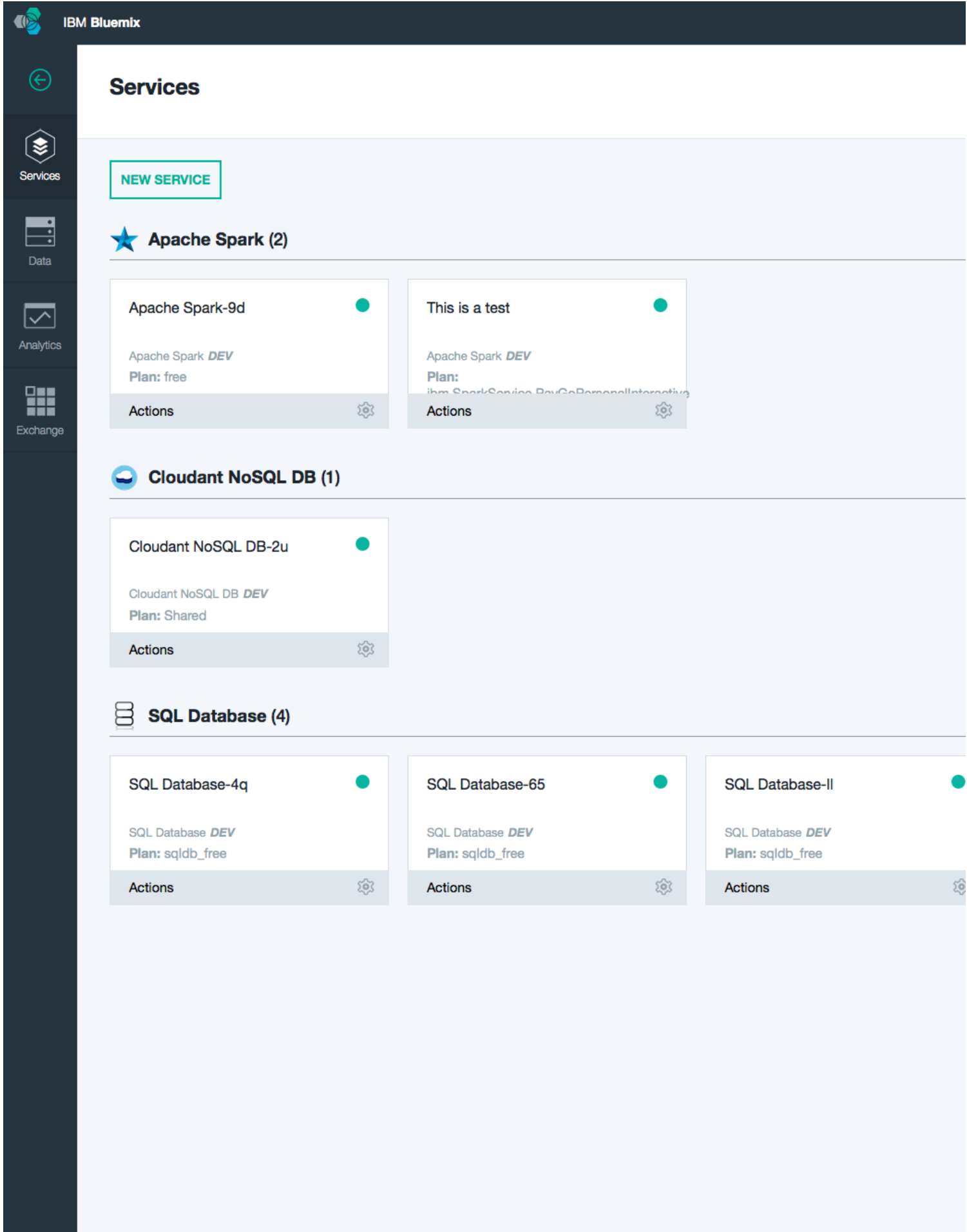
Users come from varying backgrounds so the level of experience is all over the map.

how research
informs design

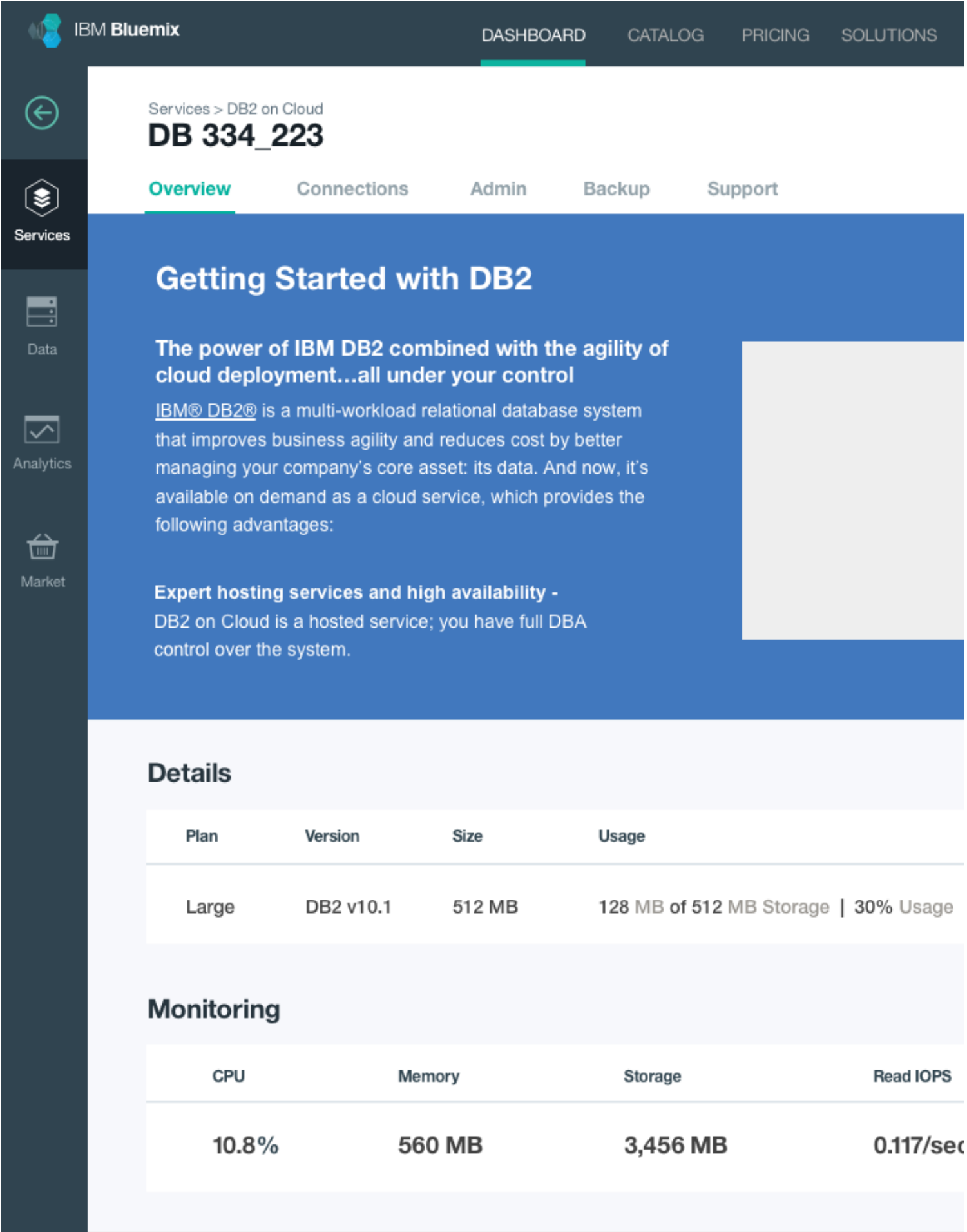
CDS Solutions

using research to address user issues

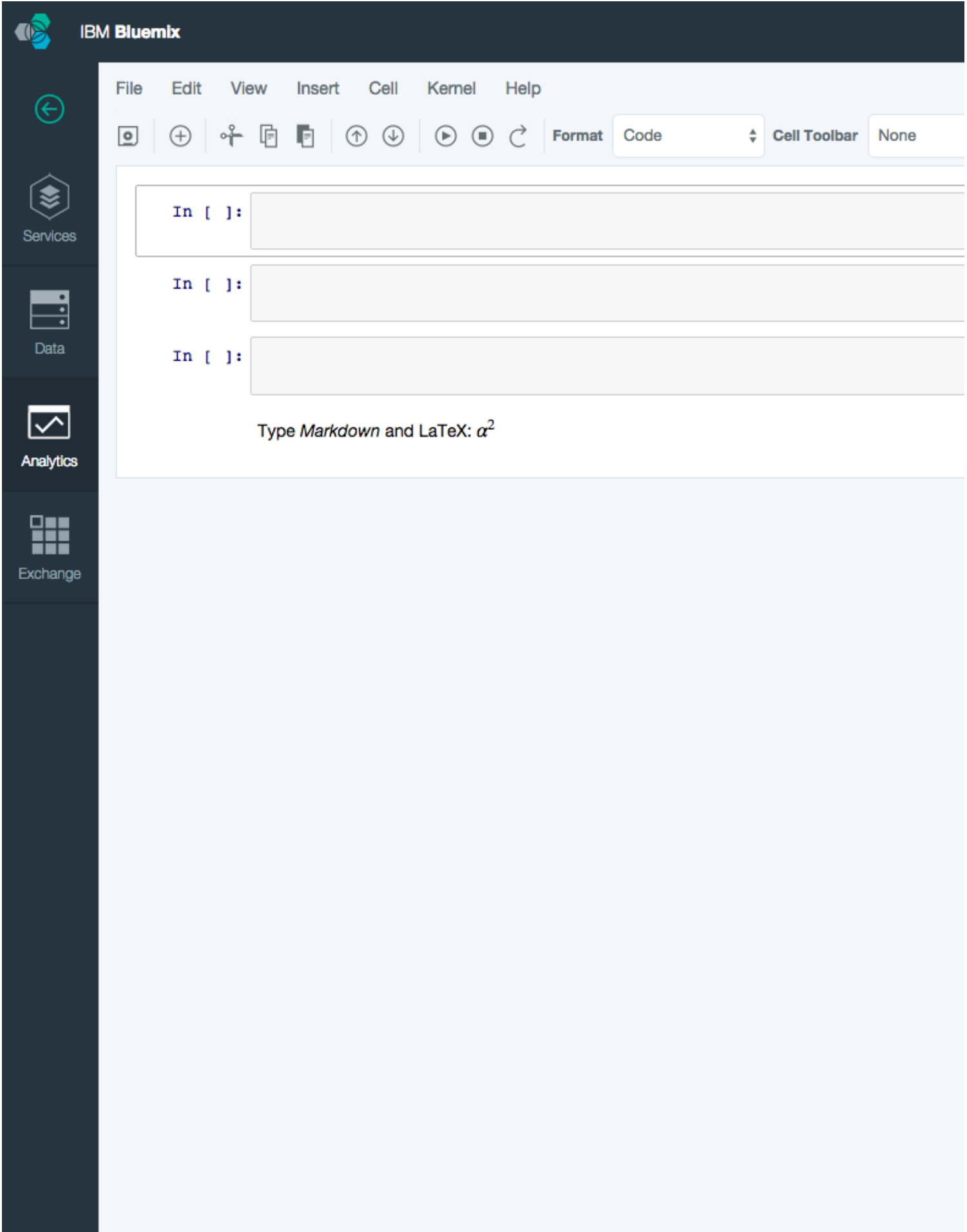
CDSX



DB2 on Cloud



Apache Spark

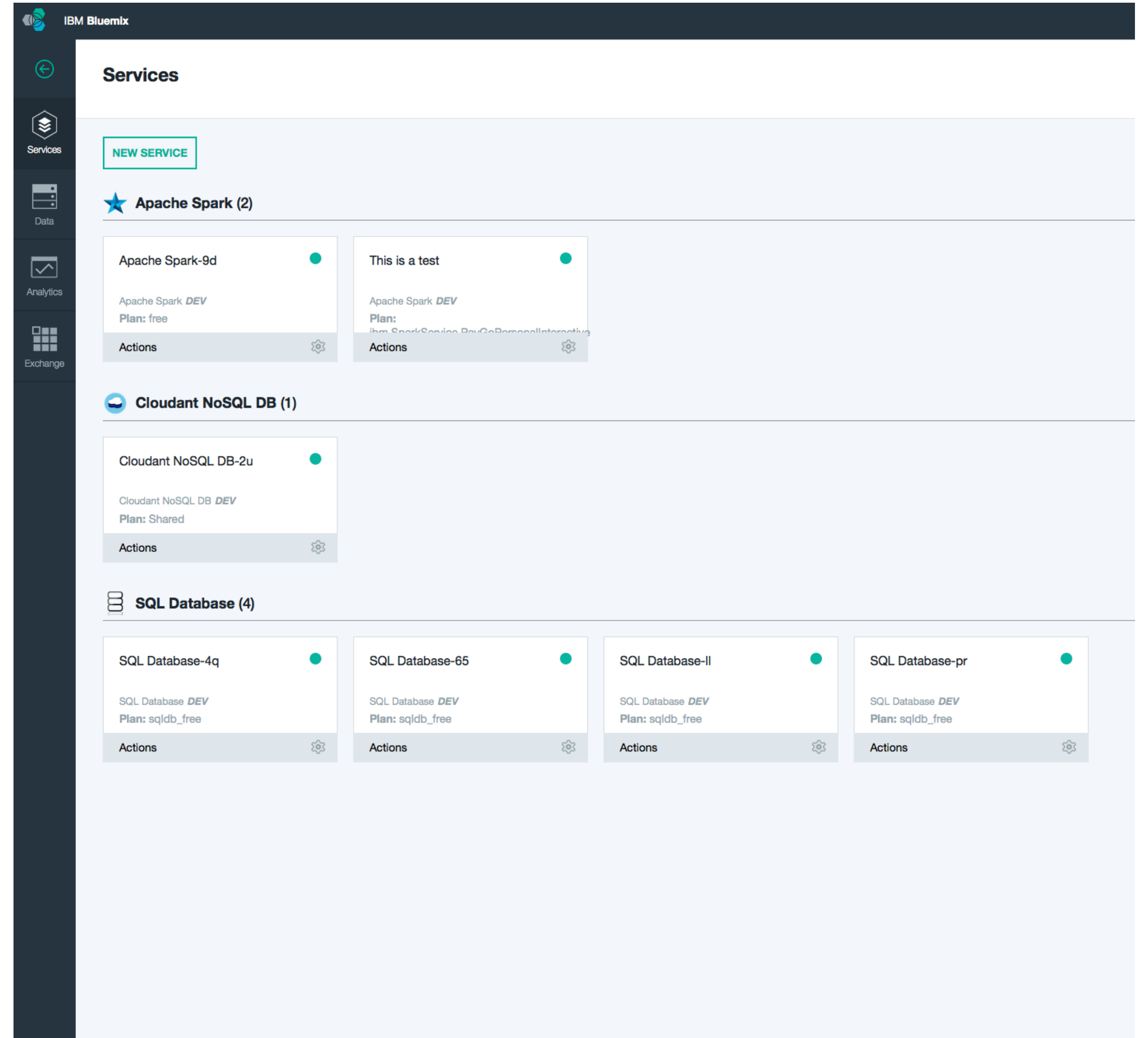


CDS Solutions

using research to address user issues

CDSX

Bringing our users all of the tools they need to work with data and analytics in one location.



CDS Solutions

using research to address user issues

CDSX

Important messages are buried

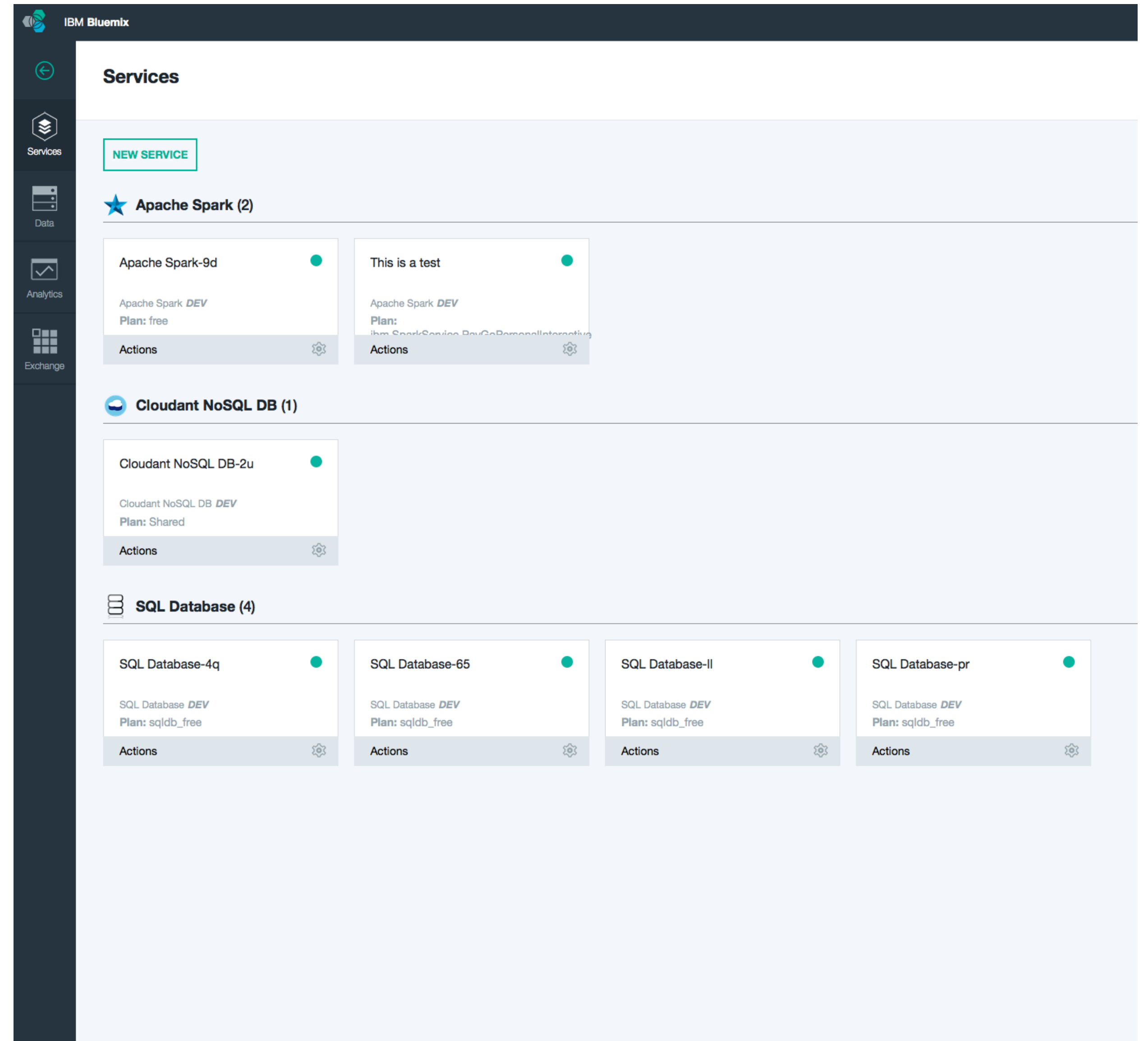
"If there is an error, I need to see it at an alert level. If there is a problem I need to know: where it is, what it is, how to fix it, and how to prevent it in the future."

-Developer

Help Me Pick = A reorganized catalog view

"But what this did is take me to a different page and totally ignored what I was searching for."

-Developer

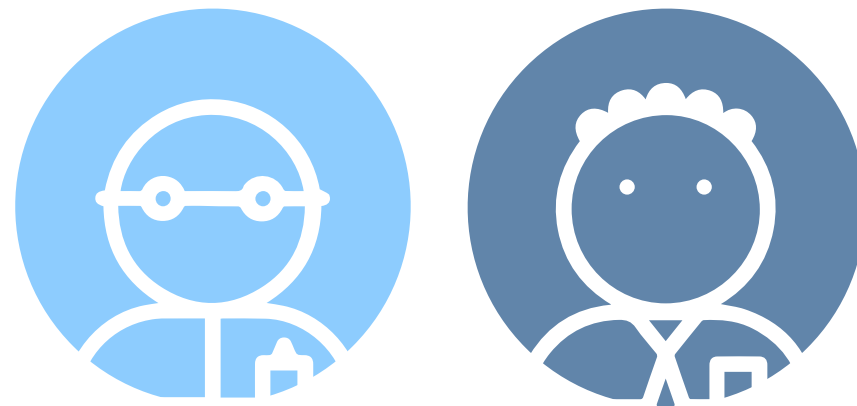


CDS Solutions

using research to address user issues

DB2 on Cloud

Providing our users a way to work with their database without having to deal with all of details.



The screenshot shows the IBM Bluemix interface for managing a DB2 on Cloud instance. The top navigation bar includes 'DASHBOARD', 'CATALOG', 'PRICING', 'SOLUTIONS', 'DOCS', and 'COMMUNITY'. The breadcrumb trail is 'Services > DB2 on Cloud'. The instance name is 'DB 334_223'. A secondary navigation bar offers 'Overview' (selected), 'Connections', 'Admin', 'Backup', and 'Support'. The main content area features a 'Getting Started with DB2' section with a video player and introductory text. Below this is a 'Details' table showing instance specifications and a 'Monitoring' table with real-time performance metrics.

Getting Started with DB2

The power of IBM DB2 combined with the agility of cloud deployment...all under your control

IBM® DB2® is a multi-workload relational database system that improves business agility and reduces cost by better managing your company's core asset: its data. And now, it's available on demand as a cloud service, which provides the following advantages:

Expert hosting services and high availability - DB2 on Cloud is a hosted service; you have full DBA control over the system.

Details

Plan	Version	Size	Usage	Status	Actions
Large	DB2 v10.1	512 MB	128 MB of 512 MB Storage 30% Usage	Available	

Monitoring

CPU	Memory	Storage	Read IOPS	Write IOPS
10.8%	560 MB	3,456 MB	0.117/sec	1.38/sec

DB2 on Cloud

Integration with other services is unknown

We have learned going along as best we can. I went and found this third party tool [but] it would be nice if IBM had a product like that.
-Database Administrator

Support is both obscure and unreliable

You have got to remember a lot of times we are doing this stuff in the wee hours of the morning so not a lot of people hanging around waiting to answer questions.
-Database Administrator

The screenshot shows the IBM Bluemix interface for a DB2 on Cloud instance. The top navigation bar includes 'DASHBOARD', 'CATALOG', 'PRICING', 'SOLUTIONS', 'DOCS', and 'COMMUNITY'. The instance name is 'DB 334_223'. The main content area features a 'Getting Started with DB2' section with a video player and text describing the service. Below this is a 'Details' table and a 'Monitoring' table.

Plan	Version	Size	Usage	Status	Actions
Large	DB2 v10.1	512 MB	128 MB of 512 MB Storage 30% Usage	Available	Settings

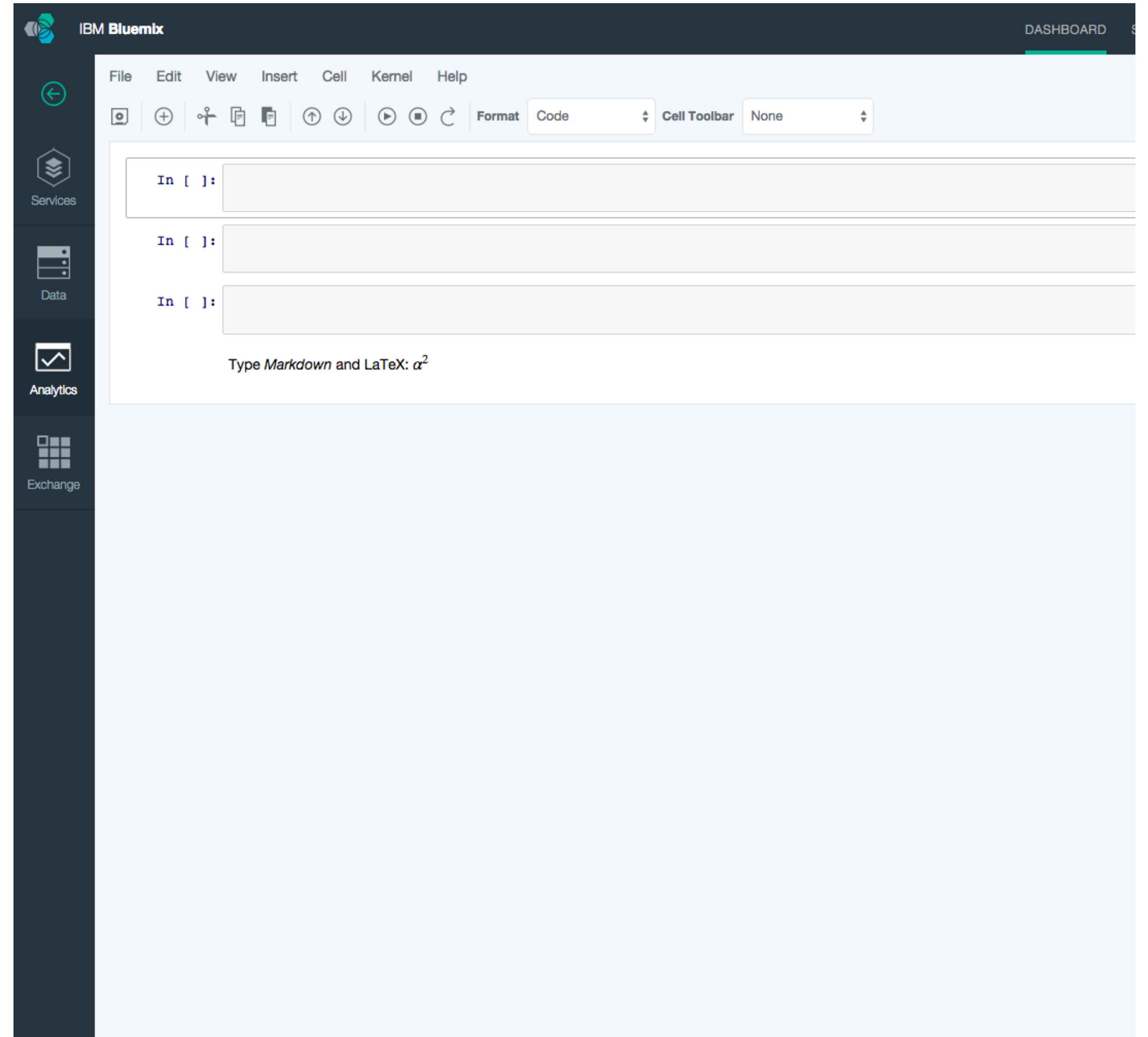
CPU	Memory	Storage	Read IOPS	Write IOPS
10.8%	560 MB	3,456 MB	0.117/sec	1.38/sec

CDS Solutions

using research to address user issues

Apache Spark

Providing a platform for all of our users to use the latest in analytics, while working together across skills.



Apache Spark

Relationships are fuzzy within the promising platform

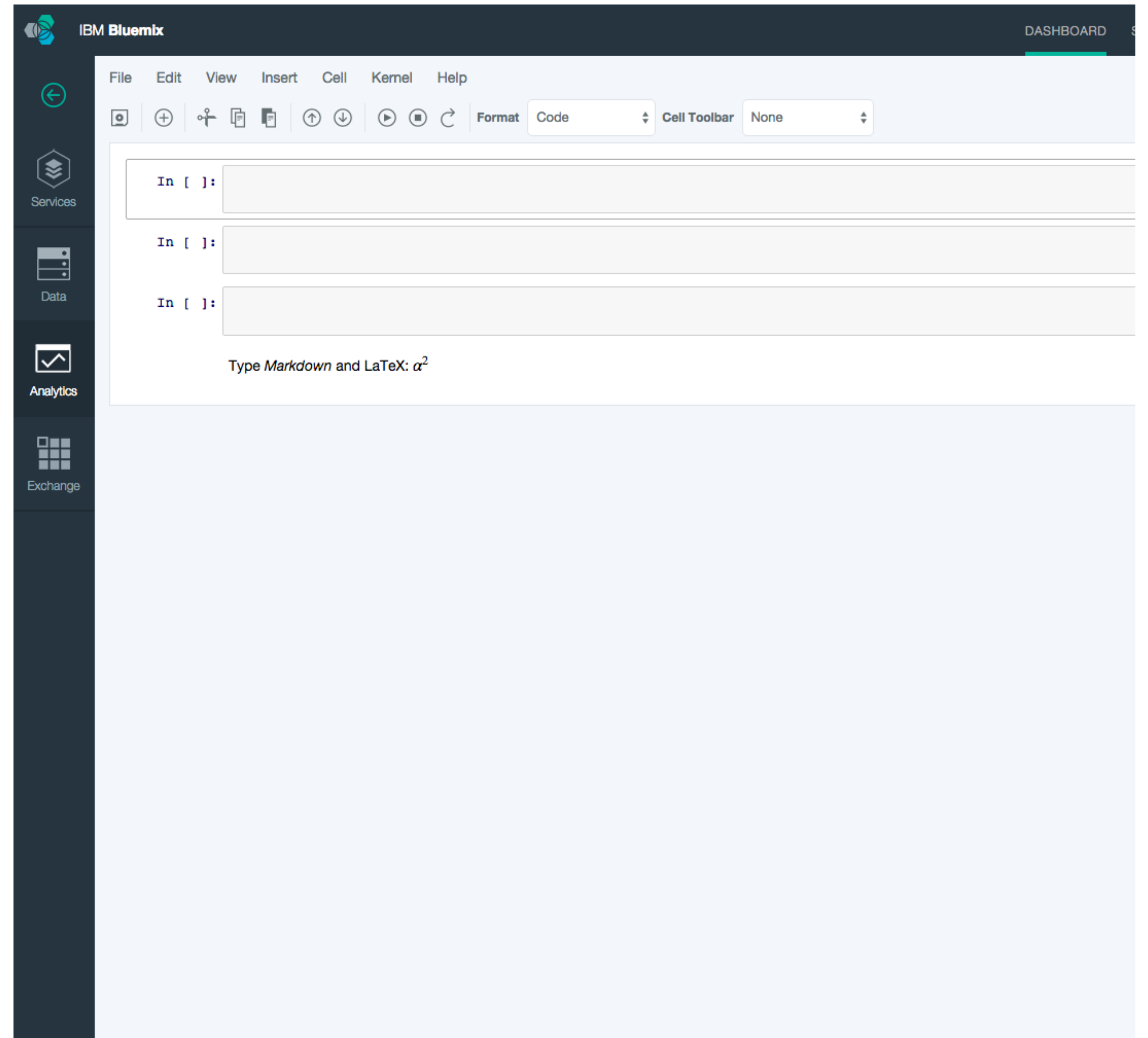
“If I’ve got my Spark hat on, I came here to learn about Spark, but then I see this Notebooks thing, and what does that mean? As a new user, I would be confused about what Notebooks had to do with Spark.”

-Data Scientist

Users get lost very quickly

Now I am in Analytics, but oh wait, I clicked Spark. But nowhere on this page does it say ‘Spark.’”
Data Scientist at IBM

-Developer/Data Scientist



Design Principles

1

Only show they need to know

“If there is something for me to address fine, let me know. But if there is nothing for me to address I don’t want to know about it.”

2

Manage the important parts

“I think you need to make things so that people can actually do the job and not worry about managing it. I think that is the approach you have to take. Especially for cloud.”

3

Provide bird’s eye and bullseye view

“When user goes into the tool they have a picture of what is going on and then they can zero in on what they want to work on or what they need to address.”

4

Get to demo as soon as possible

“My goal is to get a demo as fast as possible. My goal is to make it work and then make it work well later.”

Design Principles

5

Design discover/try in one moment

“I have to do very many clicks and open very many tabs to understand really how I would host all of these services. how I would actually start using them.”

7

Provide in-context learning

“If I had an interactive panel on the left hand side and I could select items there, and then it would redraw what is in the primary panel. That would be very useful and intuitive to me.”

6

Provide well designed sdk/docs

“We have looked for the documentation on the Blue Mix docs page. So that gets us the details.”

current challenges

Research Challenges

- Small user sample size limits the validity of our personas
(ie, one interview with a data engineer does not a personas make)
- Feedback is based on mostly internal user reviews which limits our perspective
- Effectiveness of research is watered down because of limited manpower.

Thanks!