aws INNOVATE
DATA AND AI/ML EDITION

Create centralized, smarter search capabilities across distributed data stores with Amazon Kendra

Samuel Gordon
Senior Cloud Architect
Amazon Web Services

Ed Fraga
Cloud Architect
Amazon Web Services

Agenda

- Why are we here?

- Challenges that we typically face

- Where have we seen these problems solved before?

- Solution to improve the tried and tested formula

- Use case

- Wrap up

Why are we here?

(1) What do we spend our time on?

- Waiting on other team/staff for answers to questions
- Undertaking ineffective tasks due to lack of info
- Duplicating efforts due to lack of awareness on other work
- Effective work
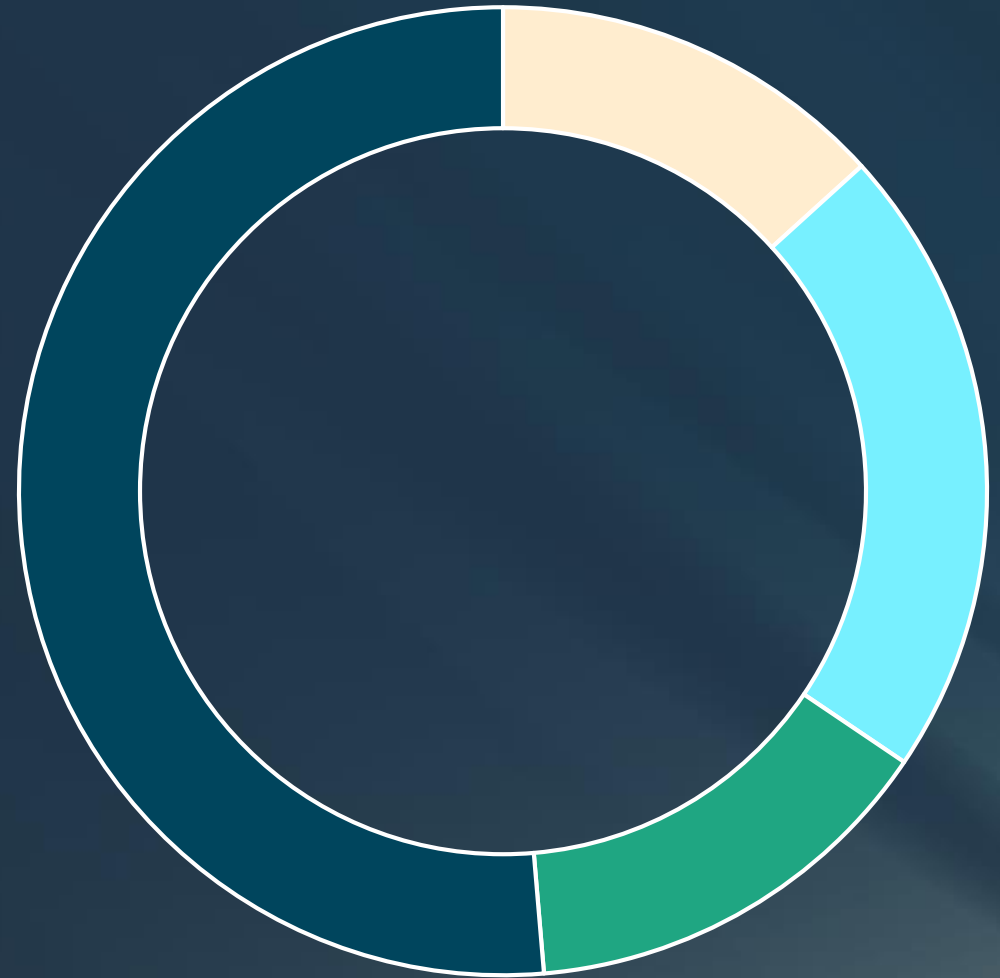
(2) How do we feel about our time spent?

?? %

have gotten into work disagreements because of misunderstandings due to virtual communication

?? %

get annoyed when coworkers ask them for documents or information they are unable to find

?? %

would consider leaving a job if there wasn't an efficient or easy way to access information

?? %

think their company's process for organizing shared documents and information is stuck in the dark ages and needs to modernize

(3) Search relevance

??%

of cases it took up to eight attempts to find the right information

??%

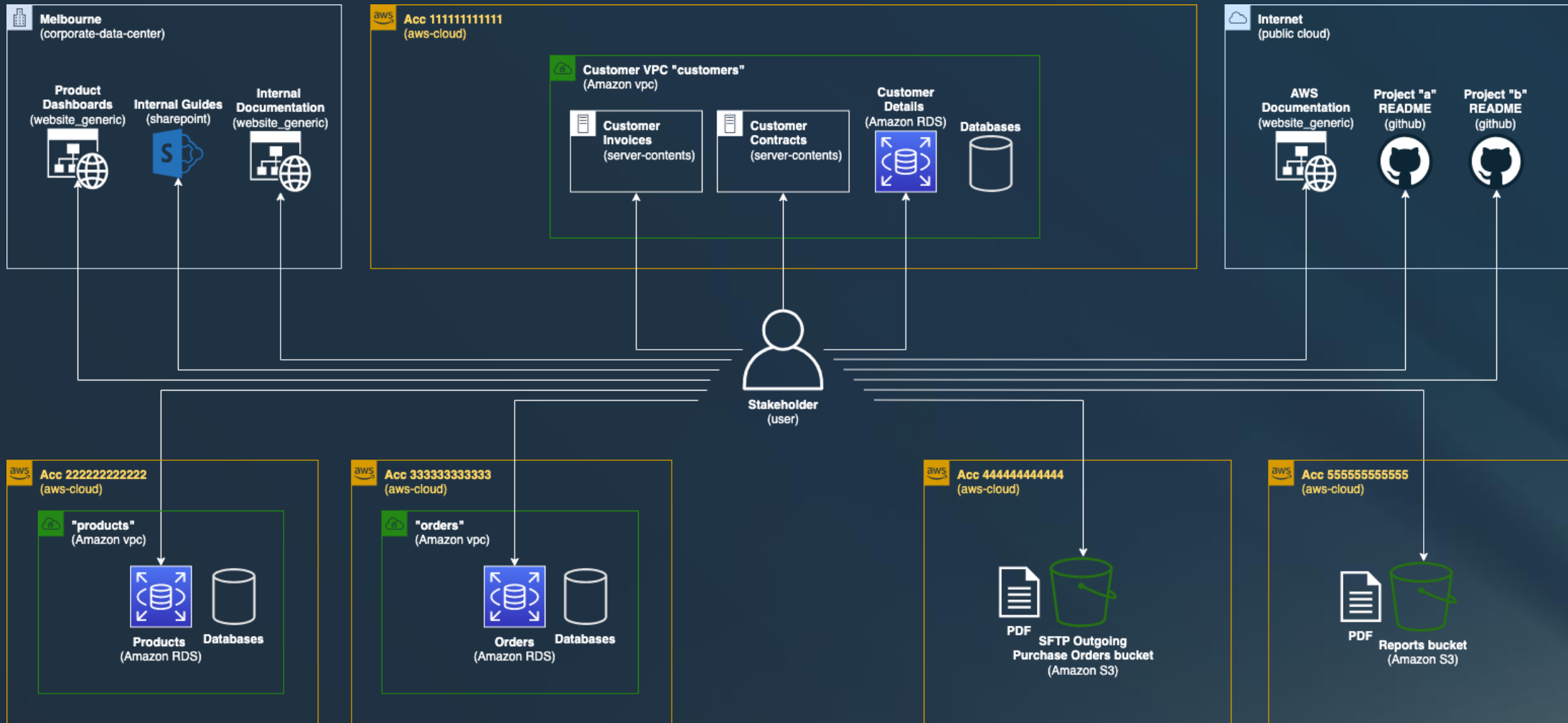of respondents reported first time successful searches
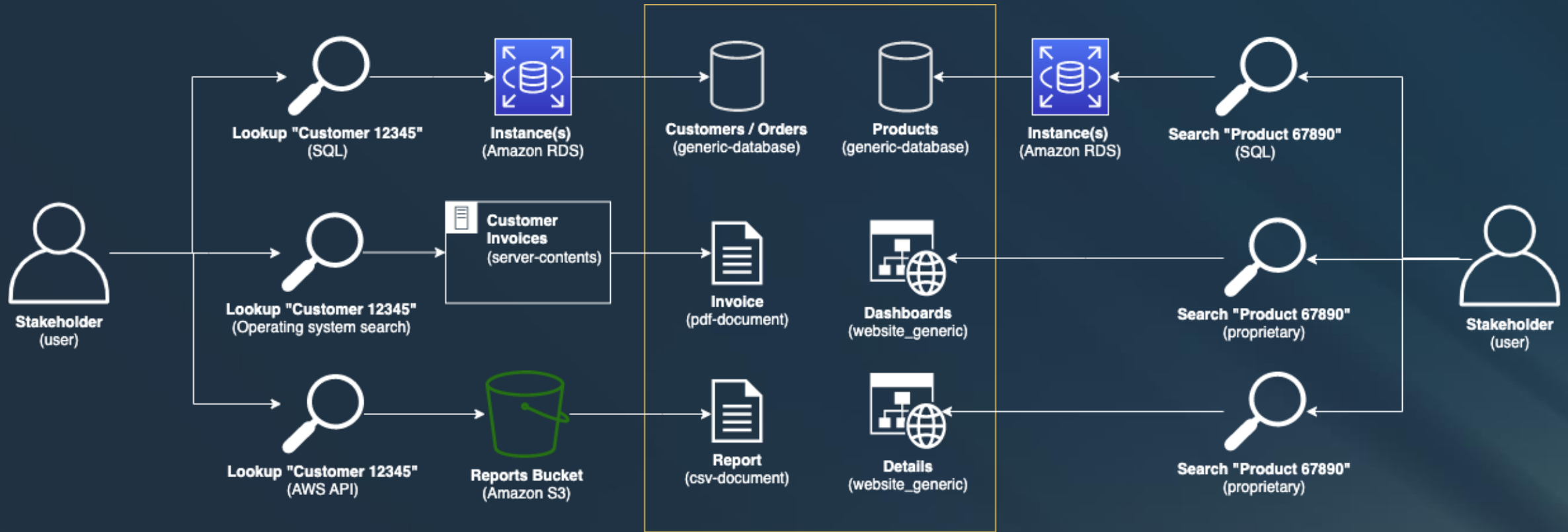
??%

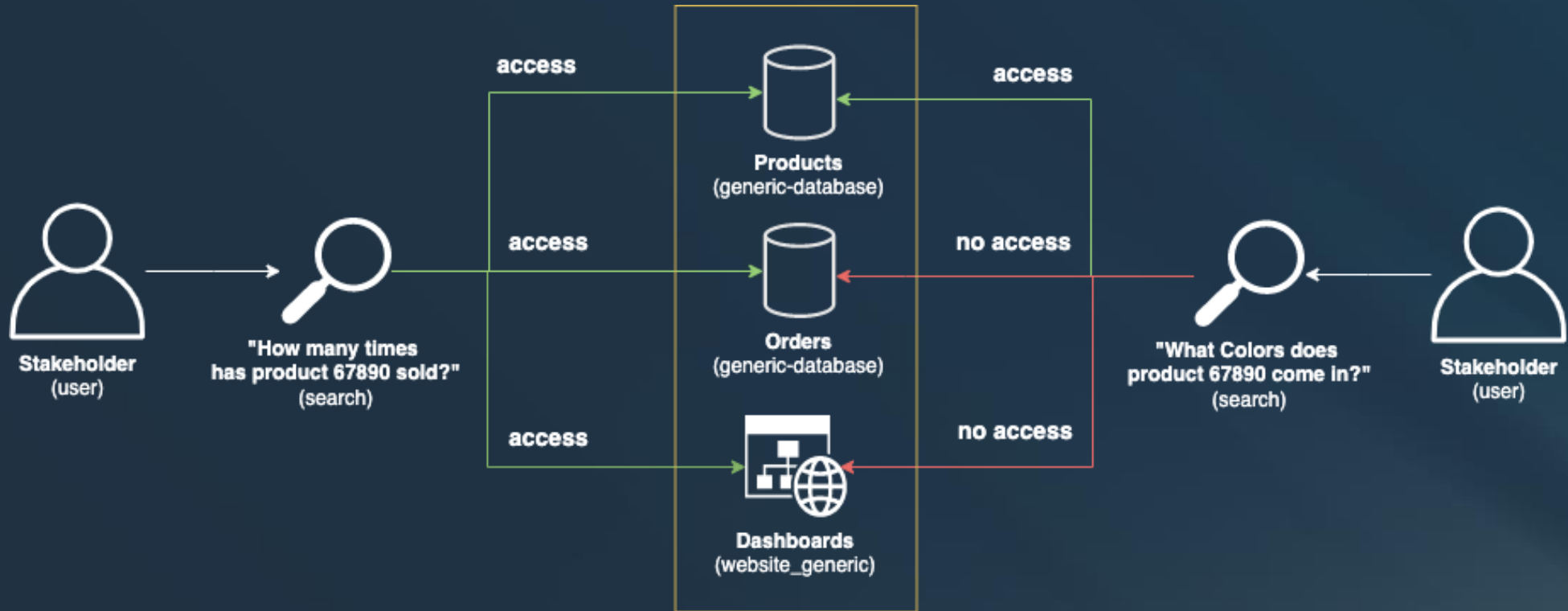of searches take at least three attempts

Challenges that we typically face

(1/6) - Distributed silos and varying data types

Melbourne (corporate-data-center)

- Product Dashboards (website_generic)
- Internal Guides (sharepoint)
- Internal Documentation (website_generic)

Acc 111111111111 (aws-cloud)

Customer VPC "customers" (Amazon vpc)
- Customer Invoices (server-contents)
- Customer Contracts (server-contents)
- Customer Details (Amazon RDS)
- Databases

Internet (public cloud)
- AWS Documentation (website_generic)
- Project "a" README (github)
- Project "b" README (github)

Stakeholder (user)

Acc 222222222222 (aws-cloud)
"products" (Amazon vpc)
- Products (Amazon RDS)
- Databases

Acc 333333333333 (aws-cloud)
"orders" (Amazon vpc)
- Orders (Amazon RDS)
- Databases

Acc 444444444444 (aws-cloud)
- PDF SFTP Outgoing Purchase Orders bucket (Amazon S3)

Acc 555555555555 (aws-cloud)
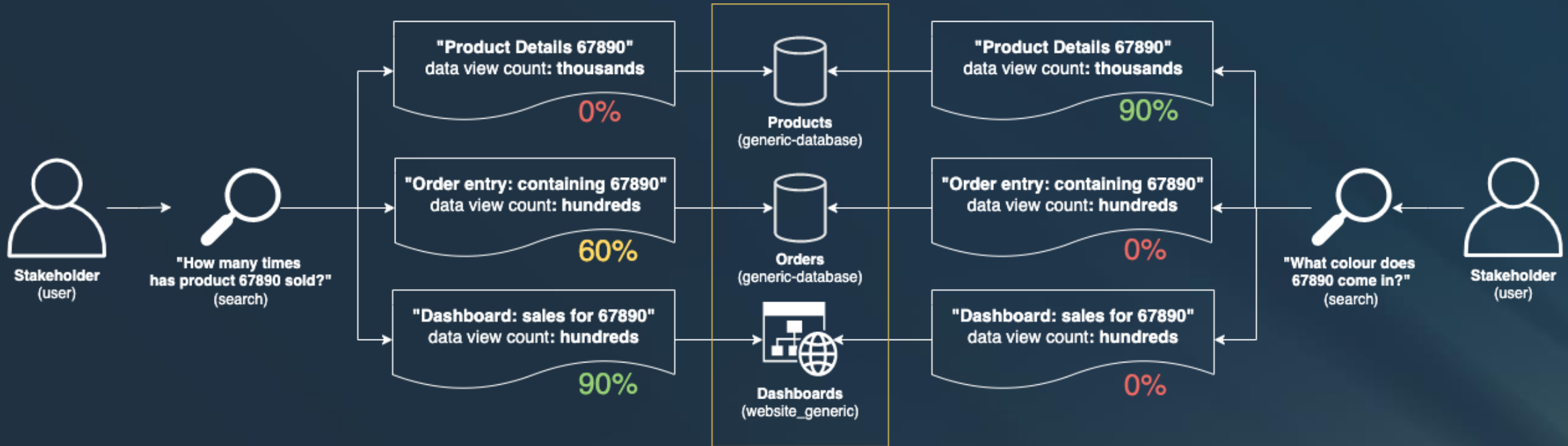- PDF Reports bucket (Amazon S3)

(2/6) – Varying interface / search capabilities

(3/6) – Data access control

(4/6) – Search relevancy

(5/6) – Search efficacy, and feedback loops

investigate search
behaviours

Products
(generic-database)

Orders
(generic-database)

Dashboards
(website_generic)

Stakeholder
(user)

"How many times
has product 67890 sold?"
(search)

onboard new
data

tweak
relavency

search new
data

Search Maintainer
(user)

PDF
(file)
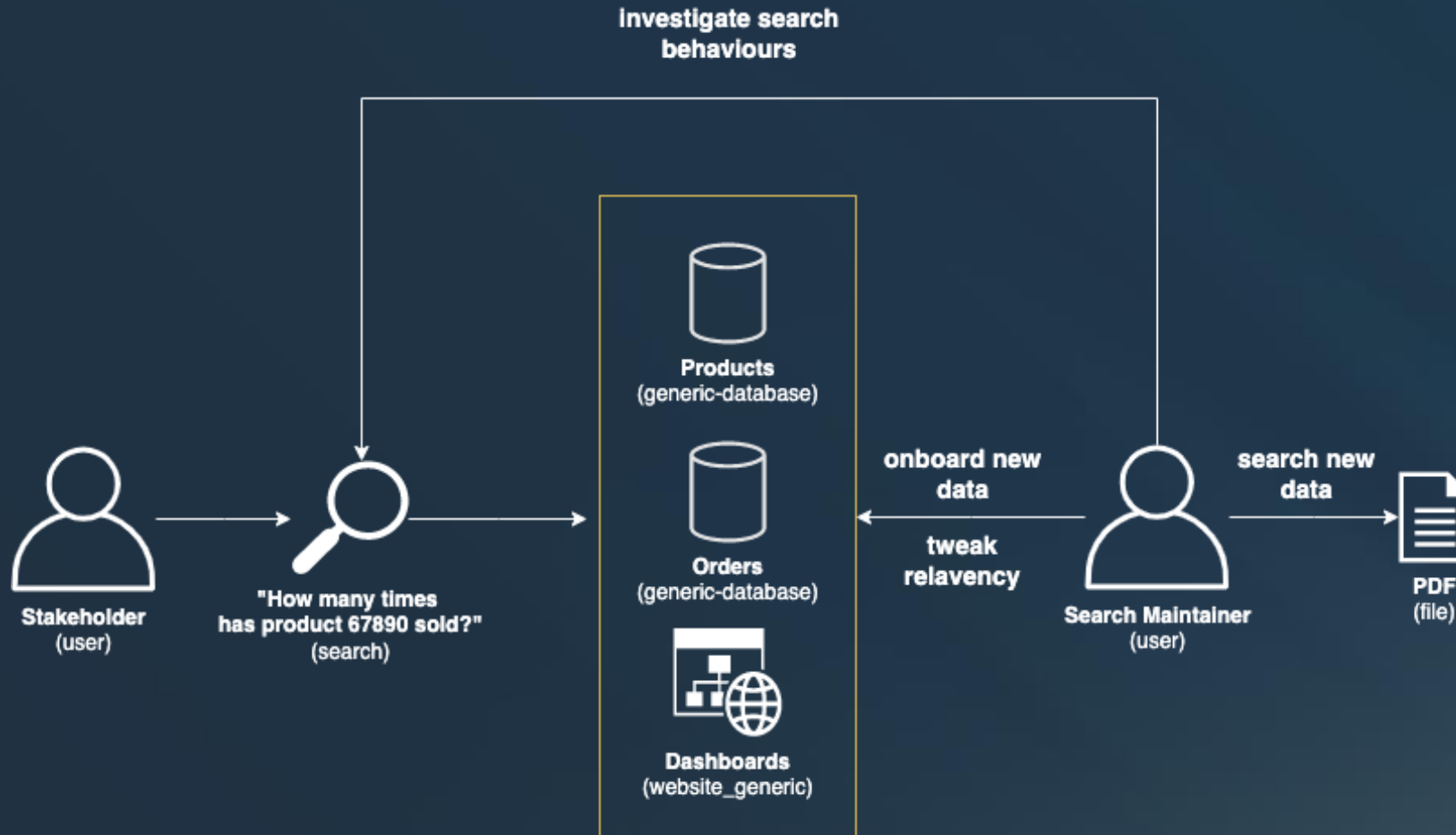
(6/6) – Manual processes
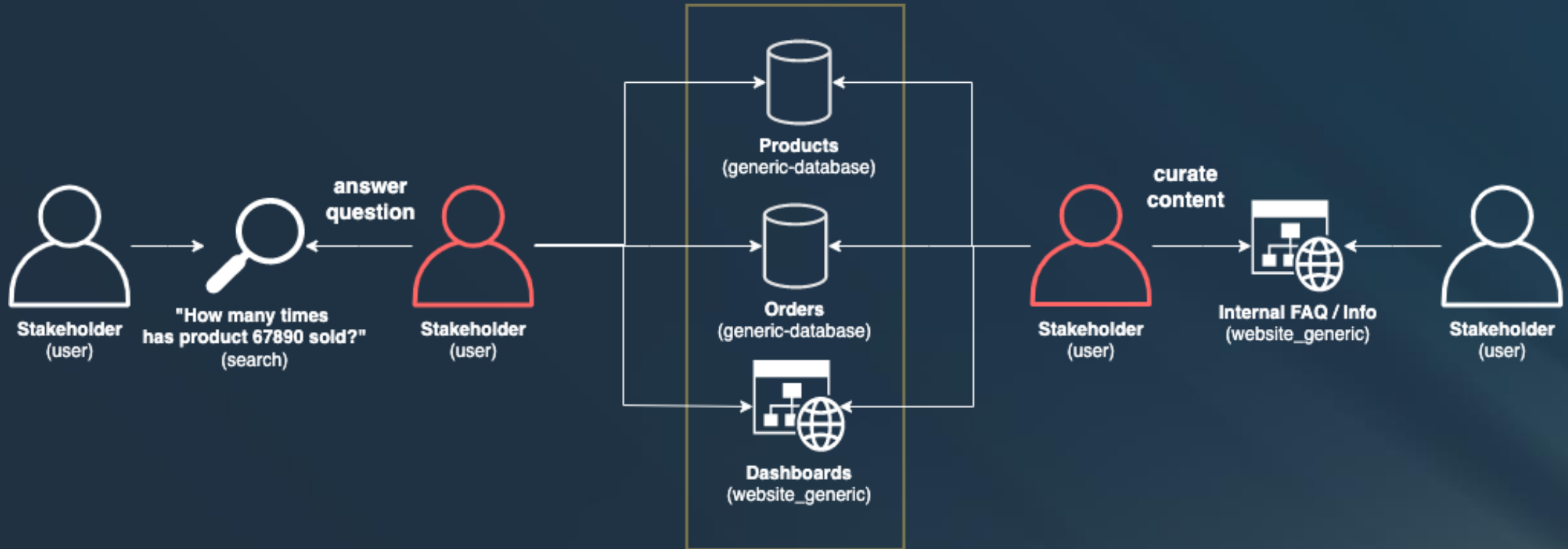
Where have we seen these problems solved before ?

Public search engines

History Lesson

- The first commonly accepted search engine came about in 1990

- Since then, consumers have progressively embraced unified searches across all aspects of the "day to day"

- As of October 2022, one of the largest globally accepted search engines hit approx. 97.5 billion monthly hits
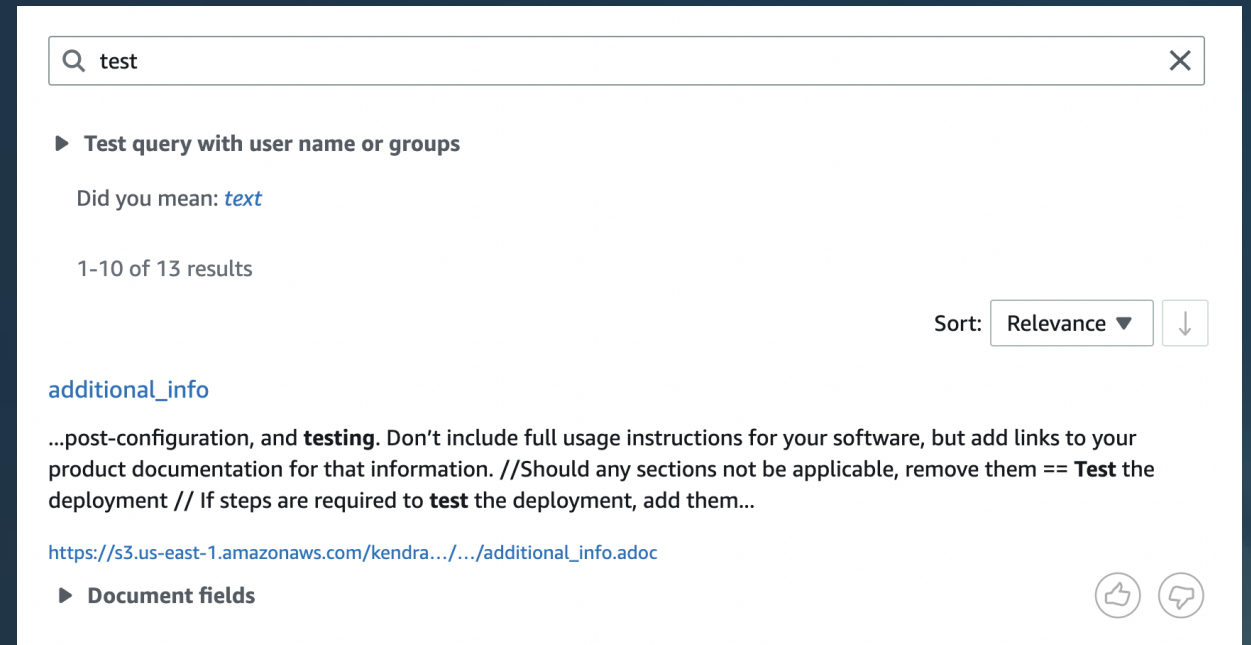
Solution to improve the tried and tested formula

Amazon Kendra

"A fully managed intelligent search service that makes it easy for customers to add search functionality to their applications."

- Unified interface / API to explore data

- Natively Integrates with AWS services

- Handles the management of all infrastructure required, so you can concentrate on the data instead of management
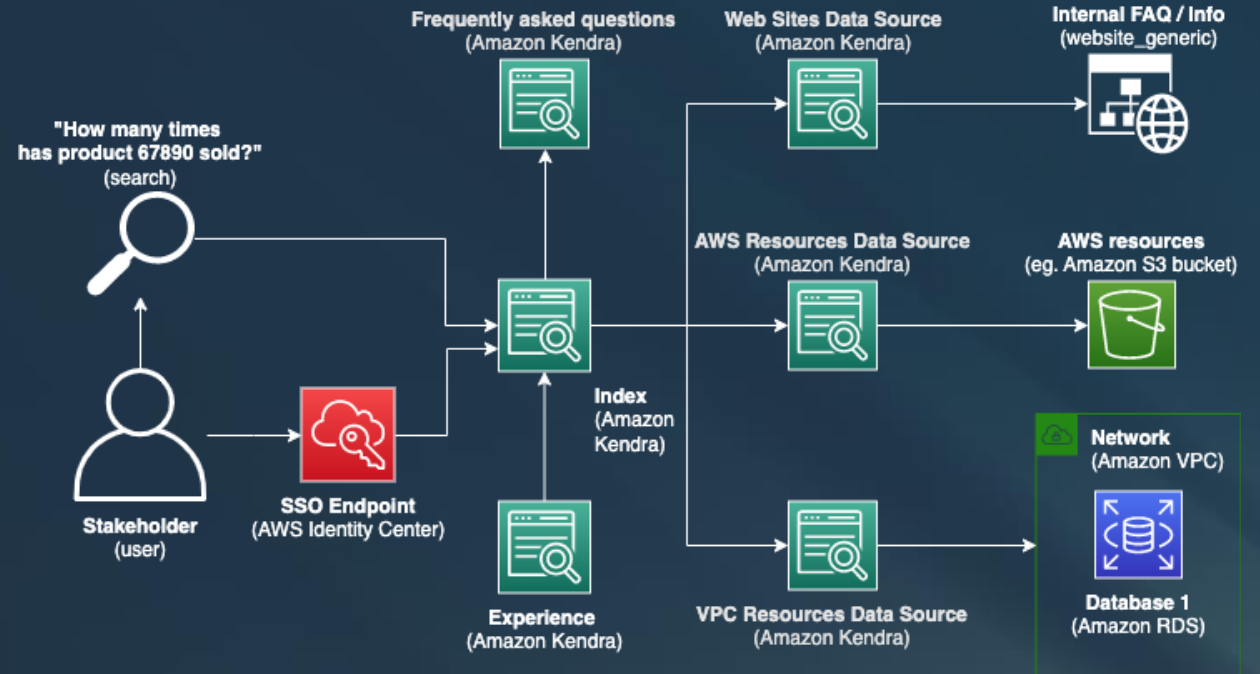
Data sources

"Amazon Kendra can index a variety of data sources, including documents, websites, and databases, and many more"

- Low to no code registration of data sources

- Removes the need to manually ingest content

- Allows for custom coded data sources

- Custom data enrichment can be performed natively via the connectors, or using AWS Lambda during ingestion

- Static FAQs can also be authored and added

Supported data sources are:

- Alfresco
- Amazon S3 buckets
- Amazon RDS for MySQL, Amazon RDS for PostgreSQL, Amazon Aurora MySQL, Amazon Aurora PostgreSQL databases
- Amazon Kendra Web Crawler
- Amazon WorkDocs
- Amazon FSx
- Atlassian Confluence
- Box
- Custom data sources
- Dropbox
- GitHub
- Google Workspace Drives
- Jira
- Microsoft OneDrive
- Microsoft SharePoint
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

For the up-to-date list of document types supported by Amazon Kendra, visit types of documents

(1) Relevancy – Natural Language Processing

"NLP is the practice of teaching machines to understand, interpret and process conversational inputs from humans"

- Amazon Kendra under the hood handles and abstracts the undifferentiated heavy lifting of common data pipeline activities such as

  - Tokenization, stop word removal, stemming/lemmatization, part of speech tagging

AI   NLP   Linguistics
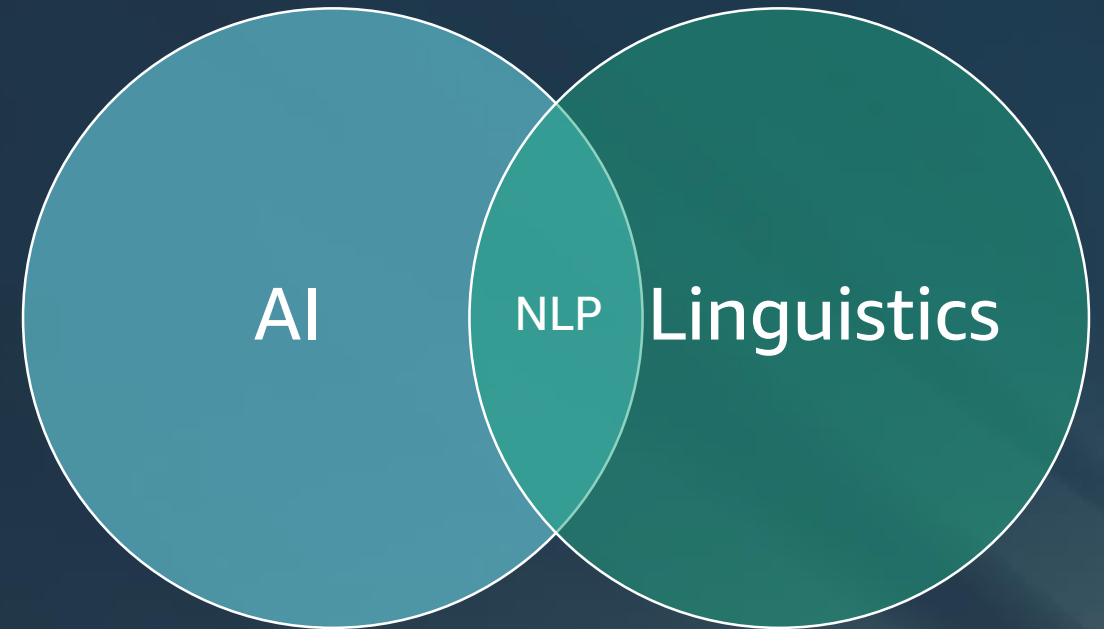
(2) Relevancy – Natural Language Understanding

"NLU is a technique within NLP which focuses on both understanding the meaning of words and sentences, as well as the broader context in which they are used"

- Amazon Kendra under the hood handles and abstracts the fundamental tasks associated with NLU around

  - Identifying "Intent"
  - Identifying "Entities" and extracting important information

NLP

NLU

(3) Relevancy – Domain optimization training

- IT
- Financial services
- Insurance
- Pharmaceuticals
- Industrial manufacturing
- Oil and gas
- Legal
- Media and Entertainment

- Travel and Hospitality
- Health
- News
- Telecommunications
- Mining
- Food and beverage
- Automotive

(4) Relevancy – Incremental learning

"Incremental learning tunes future search results quickly in a way that's data-driven, and cost effective without the need for continually re-train and deploy models"

- Amazon Kendra continually learns how users interact with search results and adapts its models to re-rank information, so the most commonly sought content is presented first.

(5) Relevancy – Tuning

"Relevance tuning allows developers to provide feedback on specific documents against queries, to improve the ranking of search results."

- Can either be done globally at the index, or per query for a "contextual" experience for the user

```
"DocumentRelevanceOverrideConfigurations" : [
    "Name": "department",
    "Type": "STRING_VALUE",
    "Relevance": {
        "Importance": 5,
        "ValueImportanceMap": {
            "HR": 2,
            "Legal": 8
        }
    }
]
```
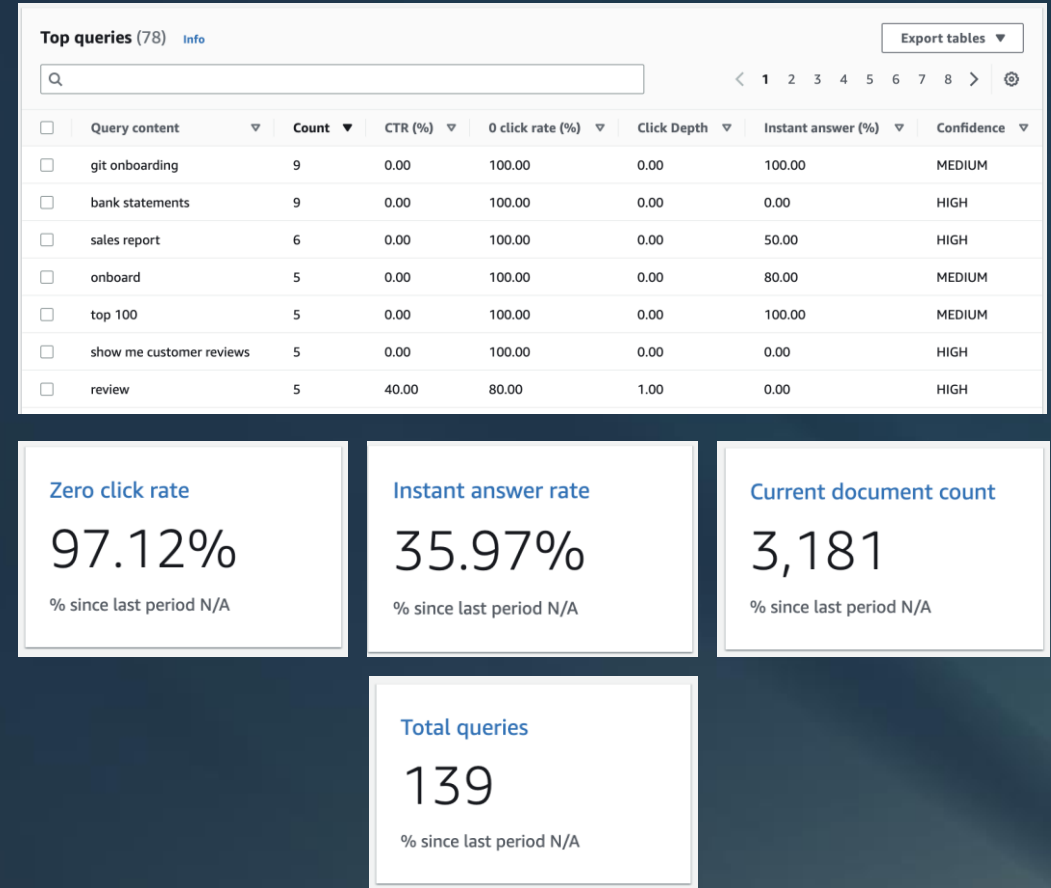
Mutates "importance" of documents with the property "department" based on if the value is ["HR", "Legal"]

Analytics and feedback loops

"Kendra also has analytics capabilities that allow you to track the performance of your searches, understand user behavior, and continually improve the search experience for your users"

- These can be viewed on a dashboard (on the right), or returned by API

Top queries (78) Info   Export tables ▼

🔍 [   ]   ‹ **1** 2 3 4 5 6 7 8 › ⚙

	Query content ▽	Count ▽	CTR (%) ▽	0 click rate (%) ▽	Click Depth ▽	Instant answer (%) ▽	Confidence ▽
☐	git onboarding	9	0.00	100.00	0.00	100.00	MEDIUM
☐	bank statements	9	0.00	100.00	0.00	0.00	HIGH
☐	sales report	6	0.00	100.00	0.00	50.00	HIGH
☐	onboard	5	0.00	100.00	0.00	80.00	MEDIUM
☐	top 100	5	0.00	100.00	0.00	100.00	MEDIUM
☐	show me customer reviews	5	0.00	100.00	0.00	0.00	HIGH
☐	review	5	40.00	80.00	1.00	0.00	HIGH

Zero click rate
97.12%
% since last period N/A

Instant answer rate
35.97%
% since last period N/A

Current document count
3,181
% since last period N/A

Total queries
139
% since last period N/A

Amazon Kendra Experience Builder

"Allows developers to quickly and easily build search experiences for their applications. It provides a visual drag-and-drop interface that makes it easy to design and customize their appearance and behaviour."

- Registers experiences with AWS Single Sign-On for discovery / authorization

- Can be built with relevancy tuning included
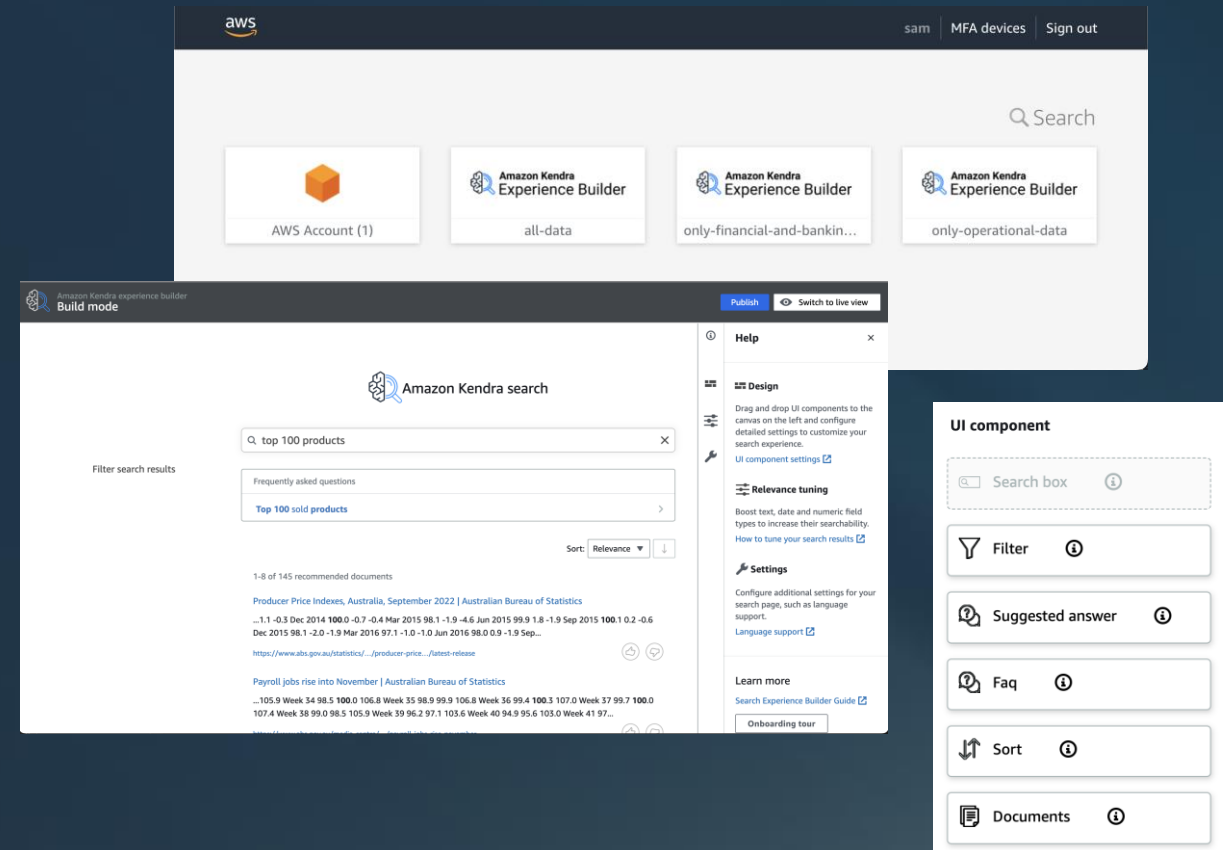
Customer adoption for "Unicorn Bank"

Unicorn Bank – Identified challenges

- Multiple data types.

- Variety of different technologies and applications in use.

- Data stored in multiple physical locations.

- Multiple searching solutions.

- Not knowing what is stored.

- Data silos.

Unicorn Bank - Existing data

Current data

- Database: Customer reviews.
- Object storage: Bank data.
- Internal wiki pages: Meeting minutes and operational runbooks.
- Curated FAQs from 3 internal teams.
- External web server: Financial data.

Possible future integrations

- Server Message Block (SMB) file storage solutions.
- Version control systems.
- Customer Relationship Management (CRM).

Unicorn Bank – Demo architecture
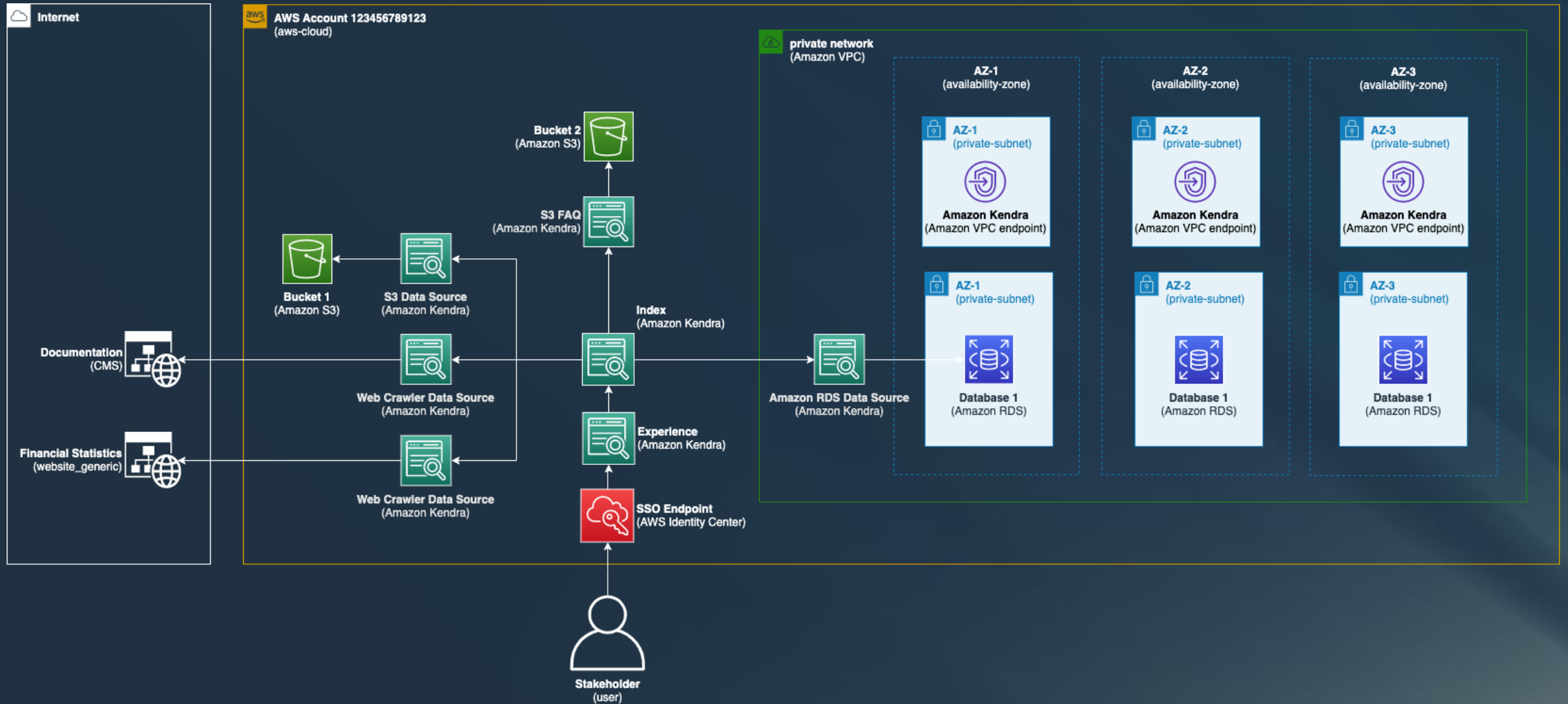
Demo

Recap

- How to build customizable search experience with Amazon Kendra

- Access control integrates with AWS Identity Center

- Low to no code connector of multiple data sources, and dynamic content enrichment

- Many ways to tweak result importance, both customer and AWS provided

- Start the journey to replace traditional manual processes with managed services.

Additional resources

- Amazon custom data sources:
  https://docs.aws.amazon.com/kendra/latest/dg/data-source-custom.html

- Amazon Kendra FSx connector:
  https://aws.amazon.com/blogs/machine-learning/securely-search-unstructured-data-on-windows-file-systems-with-amazon-kendra-connector-for-amazon-fsx-for-windows-file-server/

- Relevance tuning with Amazon Kendra:
  https://aws.amazon.com/blogs/machine-learning/relevance-tuning-with-amazon-kendra/

Visit the Data & AI/ML resource hub

Dive deeper into these resources, get inspired and learn how you can use AI and machine learning to accelerate your business outcomes.

- 6 steps to machine learning success e-book

- 7 leading machine learning use cases e-book

- Machine learning at scale e-book

- Achieving transformative business results with machine learning e-book

- Tackling our world's hardest problems with machine learning e-book

- Accelerating machine learning innovation through security e-book

- … and more!

https://bityl.co/FqdC

Visit resource hub

AWS Training and Certification

Access the AI & ML learning plan courses built by AWS experts on AWS Skill Builder

- Get started with digital self-paced, on-demand training and ramp-up guides to help you grow your technical skills

- Learn how to apply machine learning, artificial intelligence, and deep learning to unlock new insights and value in your role

- Take the steps today, towards validating your expertise with an AWS Certified Machine Learning – Specialty Certification

https://bit.ly/3FnxDH7

Learn your way explore.skillbuilder.aws »

Thank you for attending AWS Innovate – Data & AI/ML Edition

We hope you found it interesting! A kind reminder to **complete the survey.**
Let us know what you thought of today's event and how we can improve the event experience for you in the future.

aws-apj-marketing@amazon.com

twitter.com/AWSCloud

facebook.com/AmazonWebServices

youtube.com/user/AmazonWebServices

slideshare.net/AmazonWebServices

twitch.tv/aws

Thank you!