

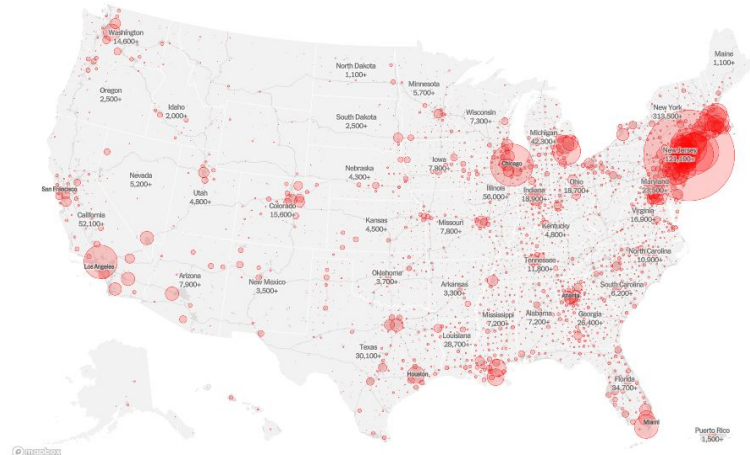
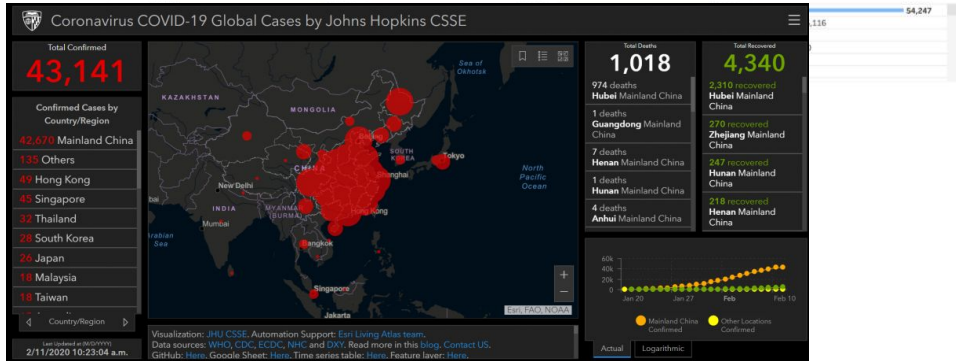
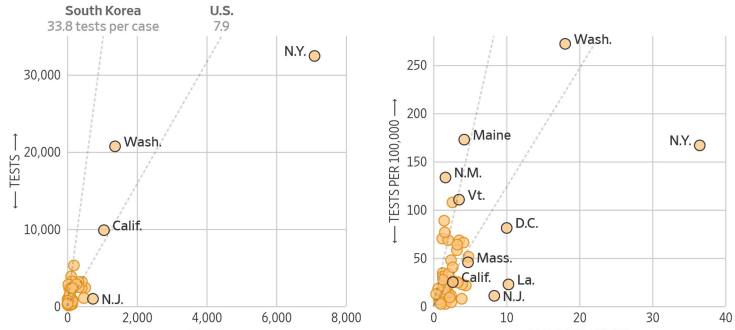
Exploring Data through Natural Language Conversations

Anand Ranganathan
Co-Founder & Chief AI Officer
anand@unscrambl.com

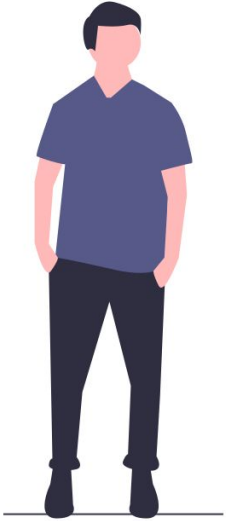
UNSCRAMBL

How do people consume data & analytics today?

- through charts & dashboards



What's wrong with dashboards



- ❌ Limited or no drill-downs
- ❌ Don't know how the data for the dashboards is produced
- ❌ Can't ask a slightly different question
- ❌ Representative of an opinion; easy to cherry-pick stats
- ❌ Can be misleading

THERE HAS TO BE



A BETTER WAY

What is Conversational Analytics?

Allow any user to ask text or voice questions of their data

and

*receive back a natural language + visual analysis
of statistically relevant and actionable insights for that user.*

**Note that in this talk, we focus on structured data stored in relational format (e.g. SQL databases, Excel sheets, etc)*

Qbo: Natural language conversations with data within collaboration platforms



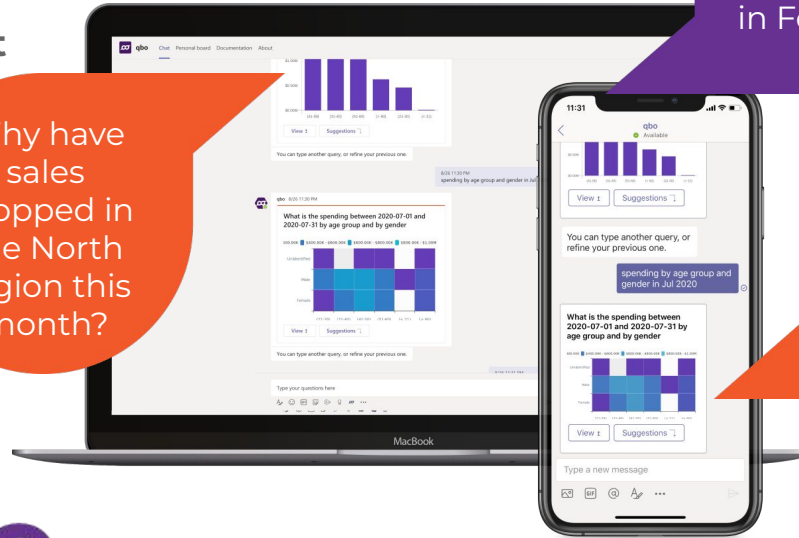
AI-Powered,
Data Analyst

Connect
Converse
Collaborate

Why have sales dropped in the North region this month?

Hey QBO, why were new acquisitions in Feb lower?

Hey QBO, how many policies are expiring in Oct 2022?



Qbo sits between users and disparate, siloed datasets

Support 20+ data connectors, and access via a web interface or Microsoft Teams



Unscrambl Qbo Demo : on Austin 311 Data

Anand Ranganathan
Co-Founder & Chief AI Officer
anand@unscrambl.com

UNSCRAMBL

UNSCRAMBL

A (very simplified) overview of NLU pipeline



Anand 7:29 PM

number of trips in winter 2017 by age and gender

↓ Entity Recognition & Construction

number of trips in winter 2017 by age and gender

↓ Identification of Query type and mapping to known concepts in DB

Type: Aggregation Query on Trips table with a group-by and a filter; **age** -> derived from birth year attribute; **gender** -> gender attribute; **in winter 2017** -> 2017-12-23 and 2018-03-19 (filter)

↓ Generate DB-specific SQL query

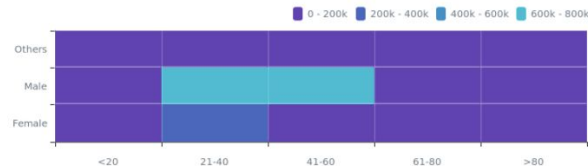
```
SELECT anon_1."age group", anon_1.gender, count(*) AS "Count"
FROM (SELECT "TRIP_ANALYSIS".end_station_id AS "end station id", "TRIP_ANALYSIS".program_id AS "program id", "TRIP_ANALYSIS".start_station_id AS "start station id",
"TRIP_ANALYSIS".bikeid AS bikeid, CASE WHEN (:birth_year_1 - "TRIP_ANALYSIS".birth_year < :param_1) THEN :param_2 ELSE CASE WHEN (:birth_year_2 -
"TRIP_ANALYSIS".birth_year < :param_3) THEN :param_4 ELSE ..., "TRIP_ANALYSIS".gender AS gender
FROM "TRIP_ANALYSIS"
WHERE ...
```

↓ Get results, decide on visualization and narratives, and present back to user



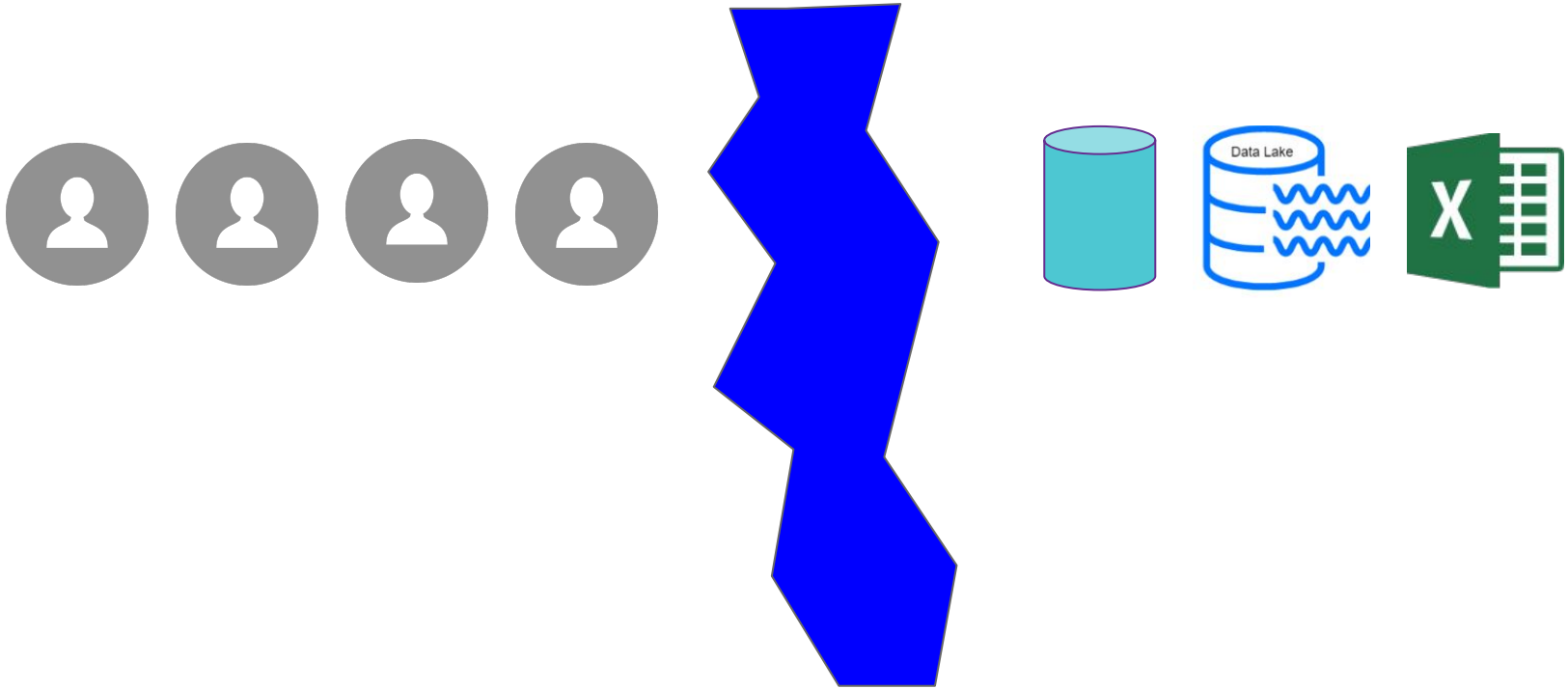
qbo 7:29 PM

What is the total number of the trips between 2017-12-23 and 2018-03-19 by age group and by gender

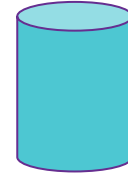
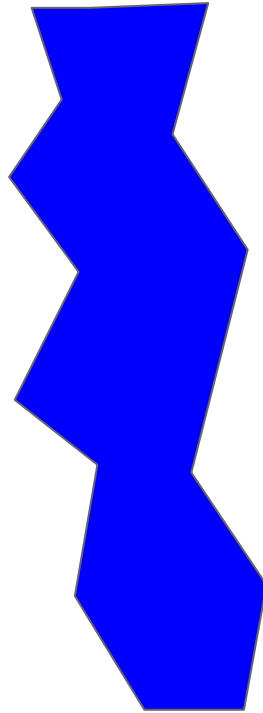


View

Key Challenge : Bridging the gap between users and data



Key Challenge : Bridging the gap between users and data

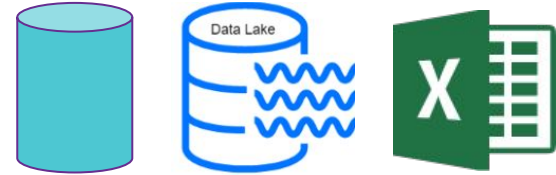
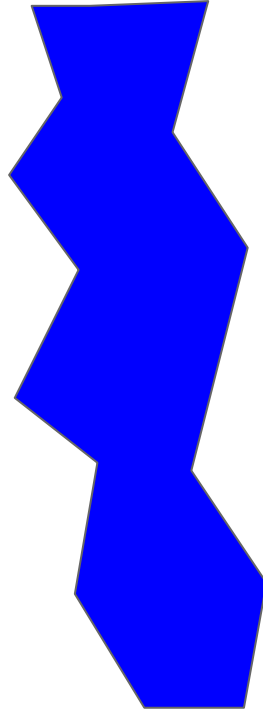


- . Users don't know what to ask
- . Users don't know how to ask
- . Users may pose questions in an ambiguous manner
- . Users may use terms not in the dataset

Key Challenge : Bridging the gap between users and data



- Users don't know what to ask
- Users don't know how to ask
- Users may pose questions in an ambiguous manner
- Users may use terms not in the dataset



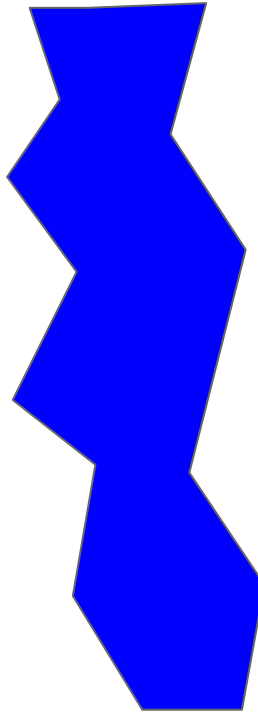
- Data may be modeled in a variety of ways
- Hidden semantics and assumptions behind different tables and columns
- Data may be incomplete, unclear
- Data may be spread across silos

Users ask questions the way they want -- & not be constrained by the physical data model



What is the total number of rides by age group

What is the ADPU in the last 6 months?



← SEMANTIC GAP -- →

MUST BE BRIDGED BY THE PERSONAL DATA ANALYST

PHYSICAL DATA MODEL

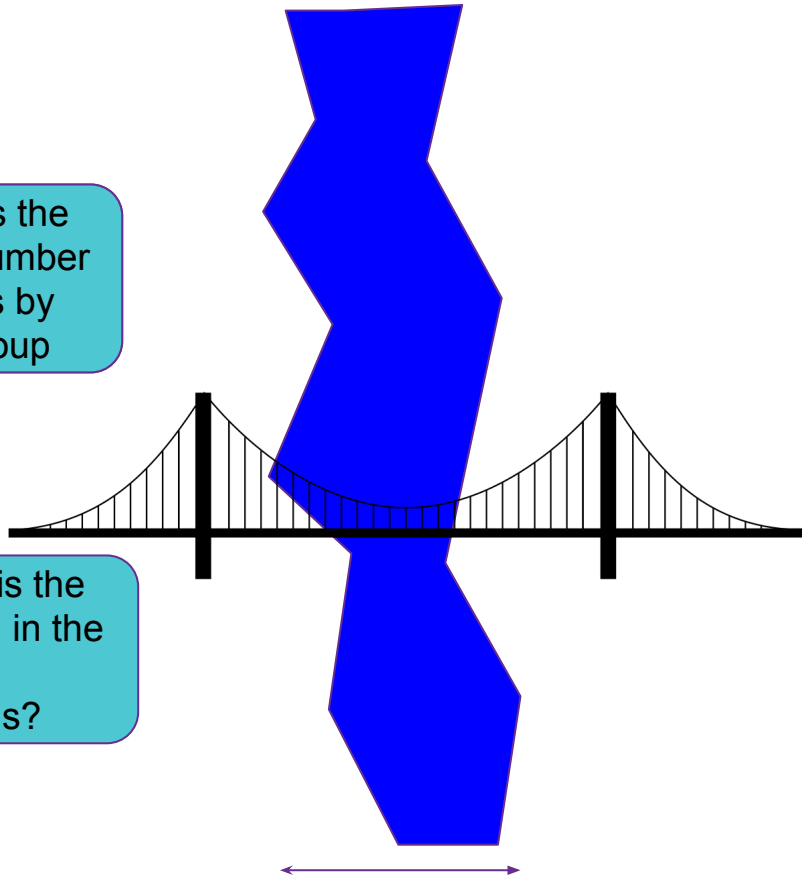
Trip
starttime
stoptime
tripduration
birth_year
bikeid
usertype

UNSCRAMBL



What is the total number of rides by age group

What is the ADPU in the last 6 months?



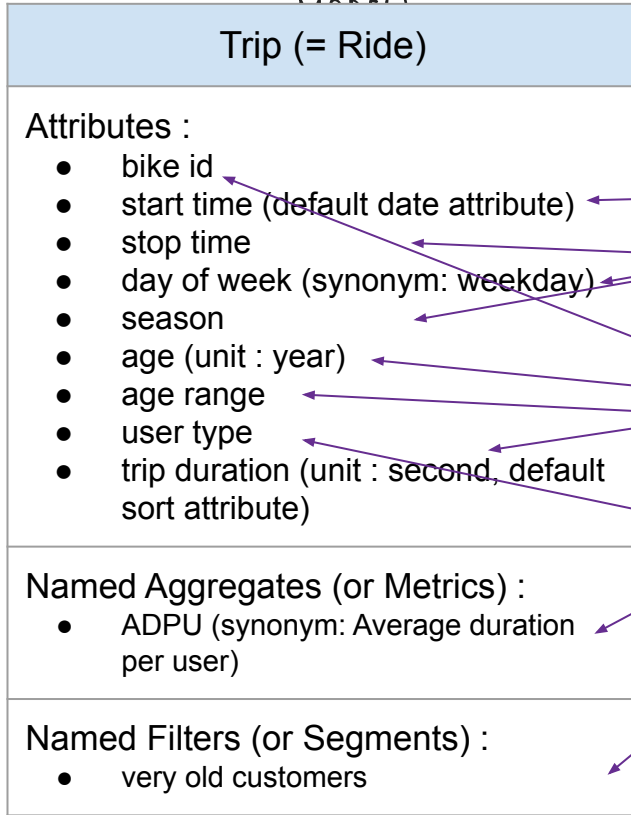
Semantic Gap

PHYSICAL DATA MODEL

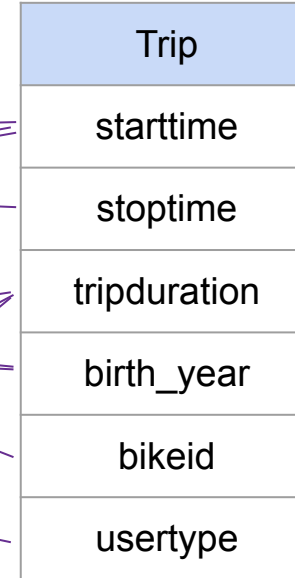
Trip
starttime
stoptime
tripduration
birth_year
bikeid
usertype

LOGICAL DATA MODEL

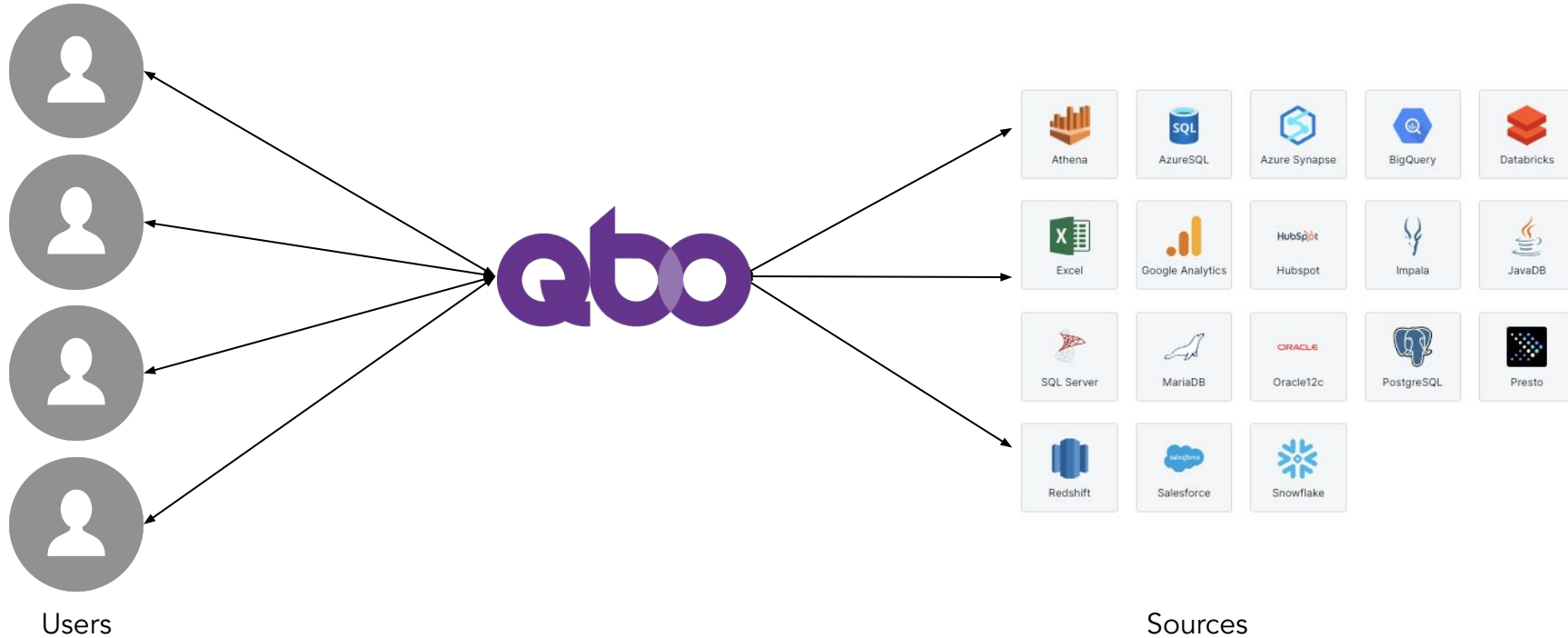
(THE BRIDGE BETWEEN THE USER AND THE PHYSICAL DATA)



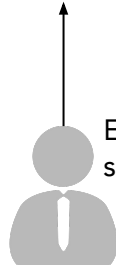
PHYSICAL DATA MODEL



Our approach to solving some of these challenges



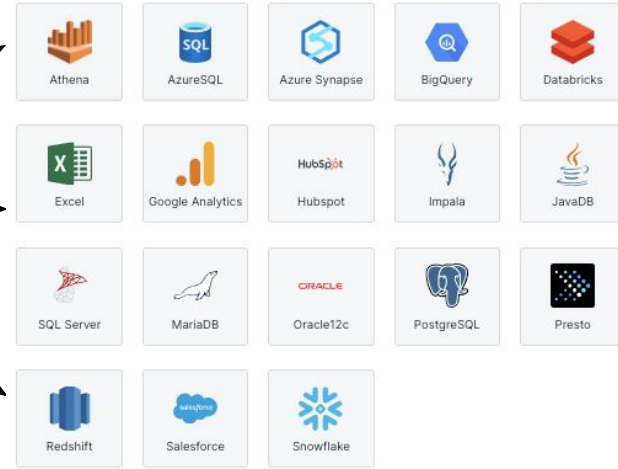
Step 1. Initial Data Discovery & Configuration



Admin/ Data Modeler

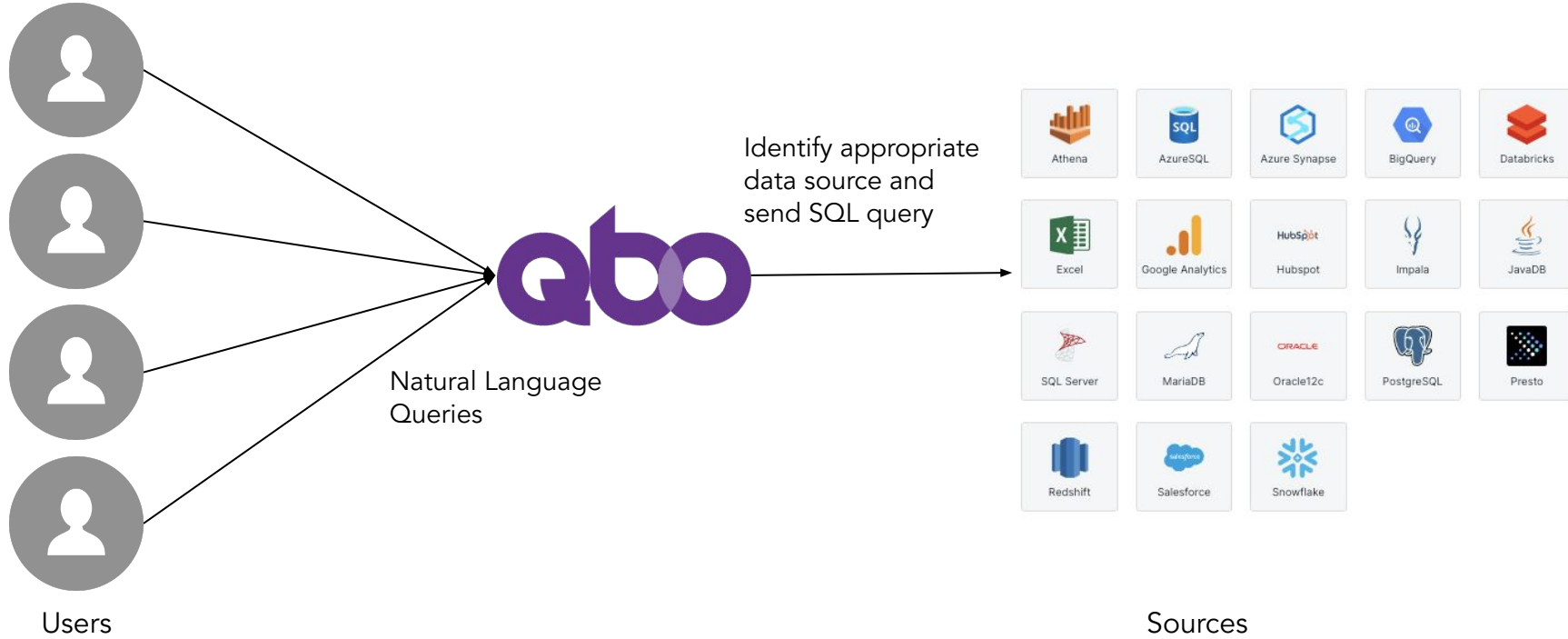
Enhance the logical / semantic model

Crawl and extract model of data sources

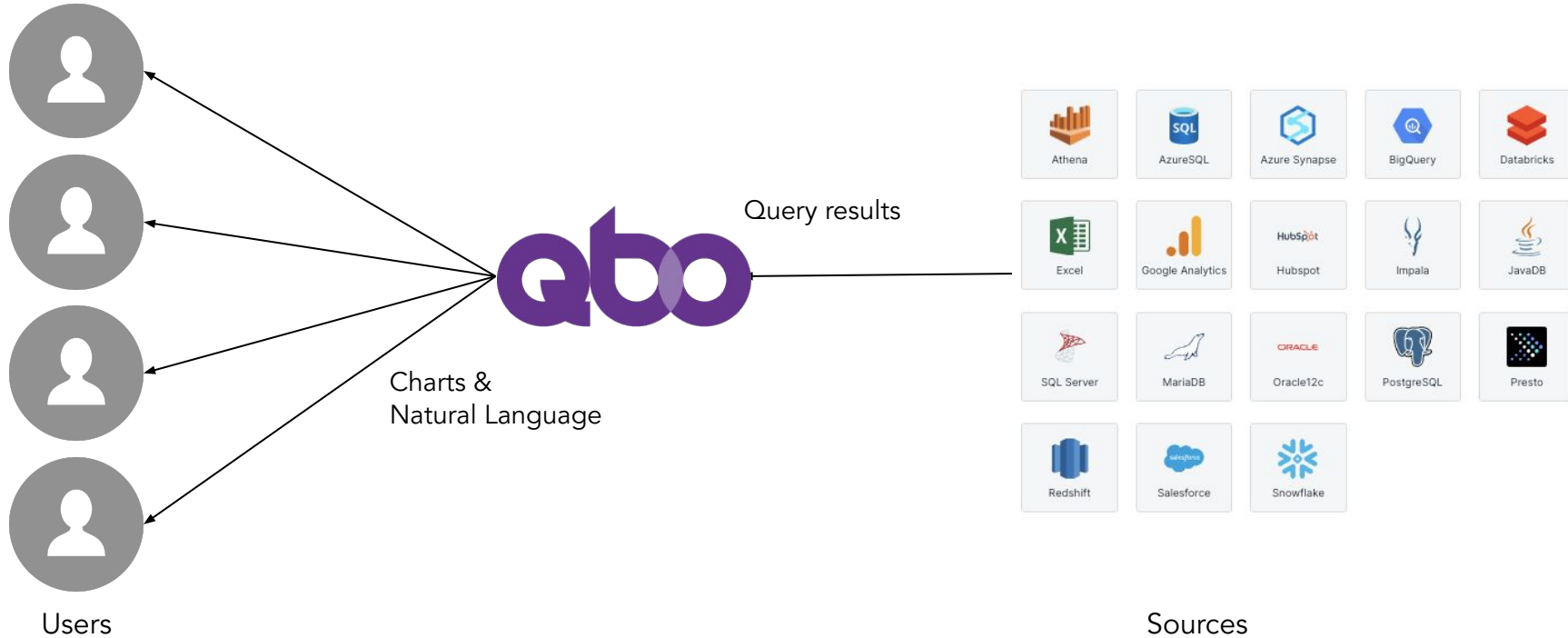


Sources

Step 2. Users can converse with qbo about their data



Step 2. Users can converse with qbo about their data



Imagine...



[#futureofwork](#)

[#futureofdata](#)



UNSCRAMBL