IBM **SQL Data Insights** & IBM SQL Data Insights **Pro**

Bringing AI to where enterprise Z data lives

Speakers:

Akiko Hoshikawa,
Distinguished Engineer - Db2 for z/OS Development

Richard Ruppel,
Brand Technical Specialist – Data and AI for IBM Z

Agenda

- Introduction of SQL Data Insights
- AI semantic queries
- Introduction of SQL Data Insights Pro
- Features in Pro
- Use Cases and Industry Starter Kit
- Q&A

The Problem

Data is rich, Insights are hard to unlock.

Enterprise data is growing rapidly —but insight is not keeping pace.

Data is fragmented across structured and unstructured sources

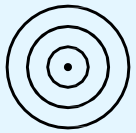
Traditional SQL retrieves exact matches—not meaning

AI adoption is slowed by:

- Data movement complexity
- Skill gaps
- Integration overhead

Result

Valuable insights remain hidden in your data



AI is Changing Data Interaction

AI will redefine how we explore enterprise data

From queries → **Discovery**

From syntax → **Intent**

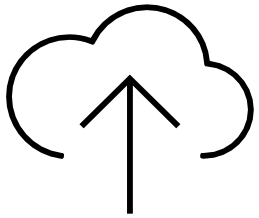
From data access → **Insights Generation**

Why Location Matters for AI Decisions

AI Off-Platform

Off-platform AI creates distance between data and decision, adding:

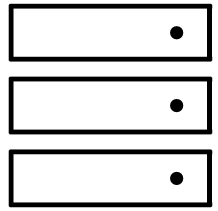
- X Latency penalty
- X Movement costs
- X Compliance risk exposure
- X Infrastructure complexity
- X Competitive disadvantage



AI on IBM Z

Run AI inference on the platforms where mission-critical data lives and decisions actually take place for co-location benefits:

- ✓ Real-time speed advantage
- ✓ Security advantage, no transfer risk
- ✓ Competitive moat



From SQL Queries to Meaning-Based Discovery

Example:

“Find hotels near a national park in Arizona, rating >4.5, under $300”

Requires:

- Exact filters
- Structured fields
- Complex query logic

Now:

“Recommend a hotel similar to Yellowstone Lodge, but in Arizona”

Requires:

- Understanding meaning
- Interpreting context
- Comparing experience, not just fields

Db2 13 with SQL Data Insights

Uncover and monetize hidden insights in your most important data

Automatically builds an unsupervised, Large Database Model against the Db2 data you tell it to, to answer business questions

Integrates AI within any Db2 application via common SQL – no data scientist required!

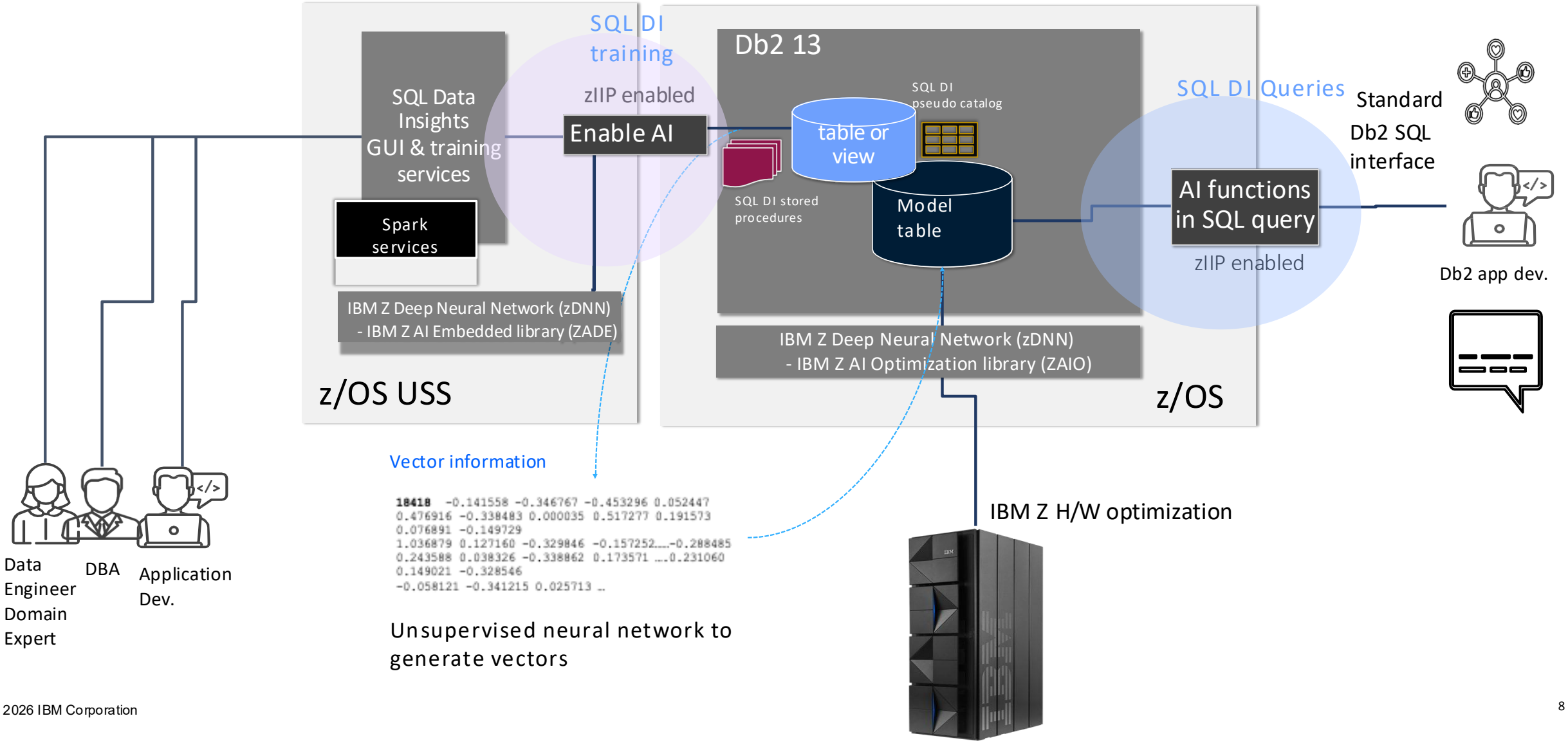
Leverages z stack AI functions (SIMD, zIIPs, Telum, …)

Ability to execute “fuzzy” queries

An example: Here is the account ID of someone who engaged in fraudulent activity.

Show me the 1000 accounts most similar to this one

SQL Data Insights – High Level Overview



Db2 database design

DSNAIDB1

Pseudo-catalog

The AIDB Pseudo-catalog Tables

- SYSAIOBJECTS
- SYSAIMODELS
- SYSAICOLUMNCENTERS
- …

DSNAIDB2

Container for Vector Tables

Tablespace <Db2 generated name>

Table DSNAIDB. <generated name>

Unique IX DSNAIDB. …

Tablespace <Db2 generated name>

Table DSNAIDB. <generated name>

Unique IX DSNAIDB. …

. . .

COL, Value, Vector, Occurrence,…

Two Databases

- One for “catalog”
- One for model tables

Pseudo-catalog

- Metadata tables for model tables
- Not for regular user access

Model tables

- Created by user via Admin UI through Db2 stored procedures
- Table space, table, indexes are given generated names
- Storage and buffer pool attributes inherited from the database
- Note : model TS pagesize is 16K

SQL Data Insights Training

DSNAIDB1

Pseudo-catalog

model metadata is inserted into pseudocatalog tables

SQLDI

training job

code, 1251, <1.8473, 0.3772, -0.3827, 0.2837, 1.286, … >

protein, c2, <0.2837, 1.2837, 1.3727, 0.8237, 1.3737, … >

fat, c3, <1.2822, 1.9846, 0.8736, -1.8373, 0.8736, …>

brand, Nabisco<-0.765, 0.983, -0.876, 0.96d3, 1.3836, …>

…

DSNAIDB2

Container for Vector Tables

Vector table is created in DSNAIDB2.

Vector data is populated by zLoad

Tablespace AIDBREM

Table DSNAIDB.AIDB_DEMO_FOOD

COL, Value, Vector, Occurrence,…

Example table: Food

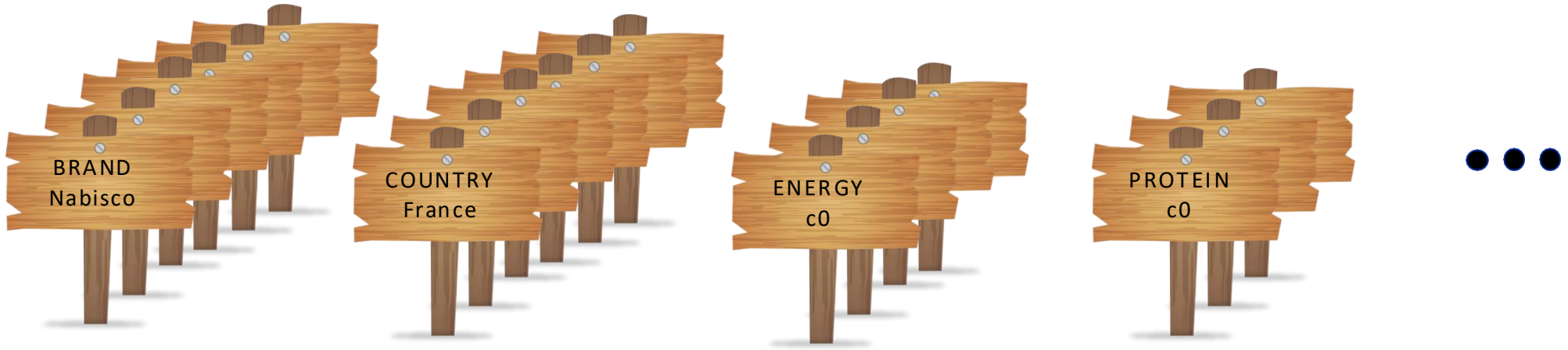
source: https://www.kaggle.com/datasets/openfoodfacts/world-food-facts

260K food products from 58 countries with nutritional information

CODE	PRODUCT_NAME	BRANDS	ENERGY_100G	FAT_100G	SUGARS_100G	PROTEIN_100G	SODIUM_100G	… *86 more numeric nutritional columns*
16087	Organic Salted Nut Mix	Grizzlies	2540.0	57.14	3.57	17.86	0.482	…
16124	Organic Muesli	Daddy's Muesli	1833.0	18.75	15.62	14.06	0.055	…
16872	Zen Party Mix	Sunridge	2230.0	36.67	3.33	16.67	0.633	…
18012	Cinnamon Nut Granola	Grizzlies	1824.0	18.18	21.82	14.55	0.009	…
18050	Organic Hazelnuts	Grizzlies	2632.0	60.71	3.57	14.29	0.004	…
18227	Organic Oat Groats	Pcc	1096.0	5.95	2.38	16.67	0.01	…
… *260171 distinct product codes*	… *187562 distinct product names*	… *48034 Distinct brands*	…	…	…	…	…	…

Concepts about SQL DI Training

Imagine: one signpost for every value in the table



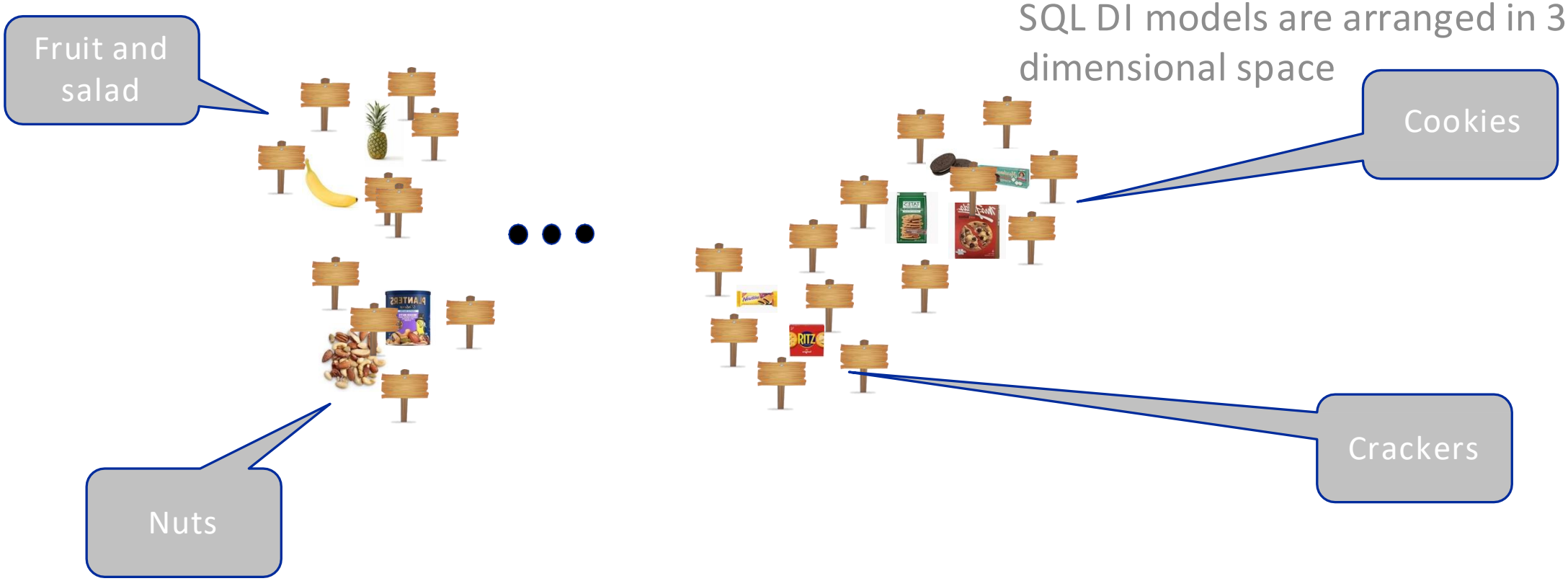
38

How to arrange every product to be near the signs that describe them?

Concepts about SQL DI Training

The final “model” is a full placement of every (255,474) sign and every (260,172) product.

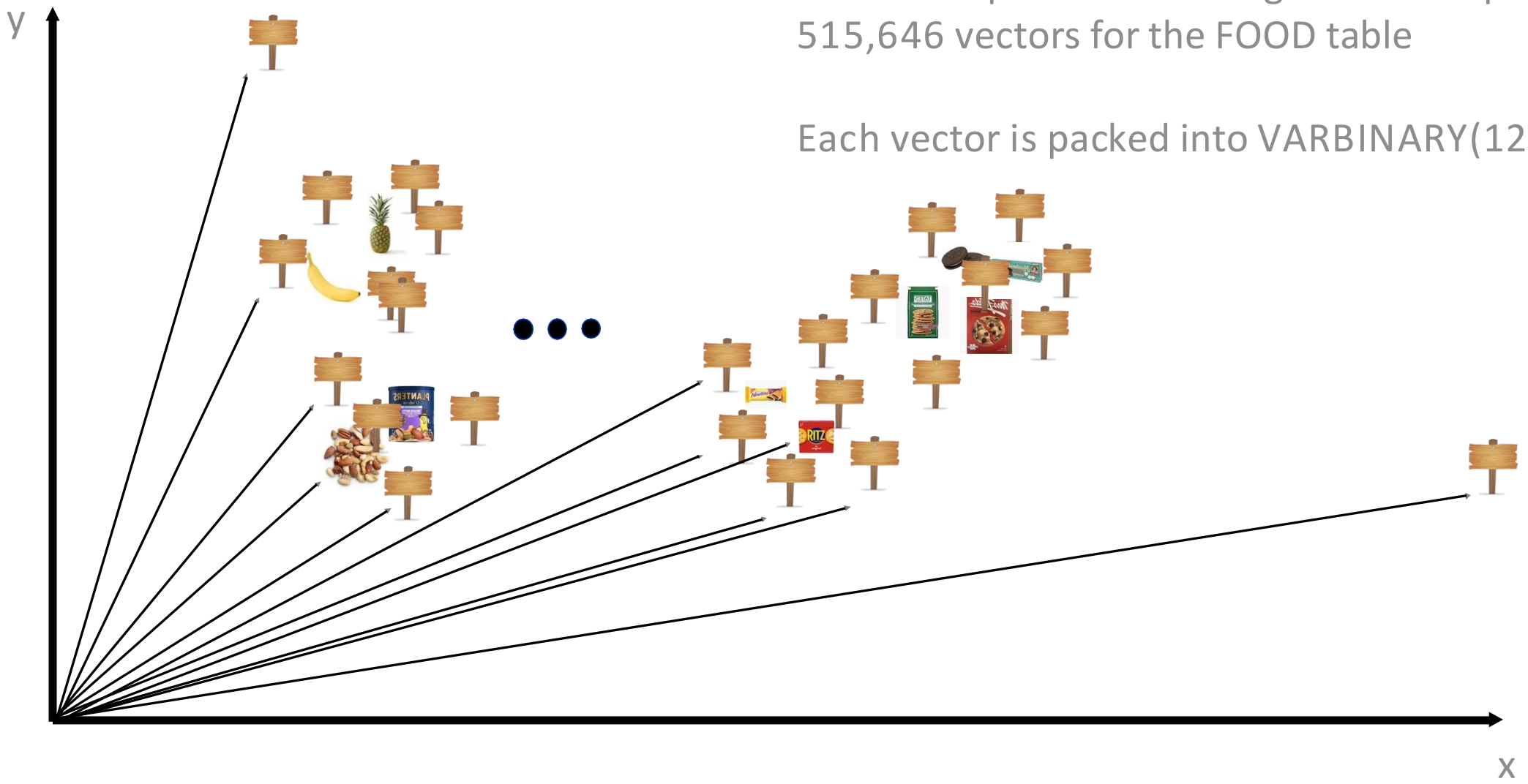
SQL DI models are arranged in 320-dimensional space



Concepts about SQL DI Training

A vector represents each sign and each product
515,646 vectors for the FOOD table

Each vector is packed into VARBINARY(1280) = 320*4



41

SQL DI Similarity Query

What foods are most similar to Oreo cookies (code '44000007492')?

1.0 = most similar
-1.0 = most dissimilar

```
SELECT F.CODE, F.PRODUCT_NAME, F.BRANDS, AI_SIMILARITY('44000007492', CODE) AS SIMI
FROM DEMO.FOOD F
WHERE PRODUCT_NAME NOT LIKE '%Oreo%'
ORDER BY SIMI DESC
FETCH FIRST 10 ROWS ONLY
```

CODE	PRODUCT_NAME	BRANDS	SIMI
4400002007	Mini Chips Ahoy!	Nabisco	0.886
44000088453	Nutter Butter 12 packs	Nabisco	0.884
418930	Soft baked originals milk chocolate chip	Mrs.Fields	0.864
14100074359	Milano Orange Flavored Chocolate	Pepperidge Farm	0.849
81029102115	Chocolate chip cookies	Tate's bake shop	0.848
24300041112	Marshmallow Pies	Little Debbie	0.845
14100079477	Milano Cookies-Mint Chocolate	Pepperidge Farm	0.838
703741000178	Colossal Chocolate Chip Cookie	Alternative Baking Company	0.831
44000012625	Newtons Fruit Thins	Newtons,Nabisco	0.831
14100079521	Chessmen	Pepperidge Farm	0.831

SQL DI Analogy Query

What foods are analogous to Oreo cookies in other countries?

```
WITH COUNTRIES (COUNTRY) AS
(SELECT 'France' FROM SYSIBM.SYSDUMMYU
 UNION ALL SELECT 'Belgium' FROM SYSIBM.SYSDUMMYU
 UNION ALL SELECT 'Germany' FROM SYSIBM.SYSDUMMYU
 UNION ALL SELECT 'Switzerland' FROM SYSIBM.SYSDUMMYU
 UNION ALL SELECT 'United Kingdom' FROM SYSIBM.SYSDUMMYU)
SELECT F.*
FROM COUNTRIES C,
TABLE (SELECT D.COUNTRIES_EN, D.PRODUCT_NAME, D.BRANDS,
        AI_ANALOGY('United States' USING MODEL COLUMN COUNTRIES_EN,
              '44000007492' USING MODEL COLUMN CODE,
              COUNTRIES_EN,
              CODE) AS ANALOGY_SCORE
       FROM DEMO.FOOD D
      WHERE D.COUNTRIES_EN = C.COUNTRY
      ORDER BY ANALOGY_SCORE DESC
      FETCH FIRST 1 ROW ONLY) F
WHERE C.COUNTRY = F.COUNTRIES_EN
```

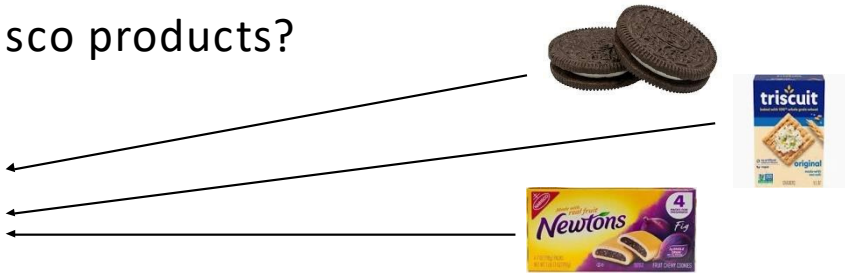
From a list of countries,
Find the top 1 most analogous food

COUNTRIES_EN	PRODUCT_NAME	BRANDS	ANALOGY_SCORE
France	kay Cookies Nougatine &Pépites de Chocolat	Rougier	0.984
Belgium	Cookies Big Rocks chocolat (x 8)	Carrefour	1.093
Germany	Creme Cookies	REWE Beste Wahl	1.098
Switzerland	Abricot biscuits	M-Classic,Migros, Mclassic	1.079
United Kingdom	shortbread fingers	Walkers	1.157

SQL DI Semantic Cluster Query

What products not branded Nabisco are in the vicinity of Nabisco products?

```
SELECT F.PRODUCT_NAME, F.BRANDS, AI_SEMANTIC_CLUSTER(CODE,
                                 '44000072742',
                                 '44000027957',
                                 '44000042257') AS SEM_SCORE
FROM DEMO.FOOD F
WHERE BRANDS NOT LIKE '%Nabisco%' AND BRANDS NOT LIKE 'Oreo'
ORDER BY SEM_SCORE DESC
FETCH FIRST 5 ROWS ONLY
```



AI_SEMANTIC_CLUSTER: Find a midpoint of up to 3 vectors, and compute distance from that point

PRODUCT_NAME	BRANDS	SEM_SCORE
Soft baked originals milk chocolate chip	Mrs.Fields	0.845
Chocolate chip cookies	Tate's bake shop	0.829
Chocolate Chip Cookie Dough	Toll House,Nestlé	0.828
Chocolate Chip	Famous Amos	0.824
Cinnamon Schoolbook Cookies	Trader Joe's	0.824

SQL DI Commonality Query

Is Oreo an outlier in our food nutritional model?

Needs FL504!

```
WITH COMMONALITY_SCORES AS
(SELECT CODE, PRODUCT_NAME, AI_COMMONALITY(CODE) AS COMM_SCORE FROM DEMO.FOOD)

SELECT 'Max', CS.PRODUCT_NAME, CS.COMM_SCORE
FROM COMMONALITY_SCORES CS
WHERE CS.COMM_SCORE = (SELECT MAX(S.COMM_SCORE) FROM COMMONALITY_SCORES S)

UNION ALL
SELECT 'OREO', CS.PRODUCT_NAME, CS.COMM_SCORE
FROM COMMONALITY_SCORES CS
WHERE CS.CODE = '44000007492'

UNION ALL
SELECT 'Min', CS.PRODUCT_NAME, CS.COMM_SCORE
FROM COMMONALITY_SCORES CS
WHERE CS.COMM_SCORE = (SELECT MIN(S.COMM_SCORE) FROM COMMONALITY_SCORES S)
```

AI_COMMONALITY computes a distance from the center of the model

	PRODUCT_NAME	COMM_SCORE
MAX	Deliciously Soft Fruit & Nut Confections	0.819
Oreo	Oreo	0.688
MIN	Nouilles Fines Wing Man	0.438

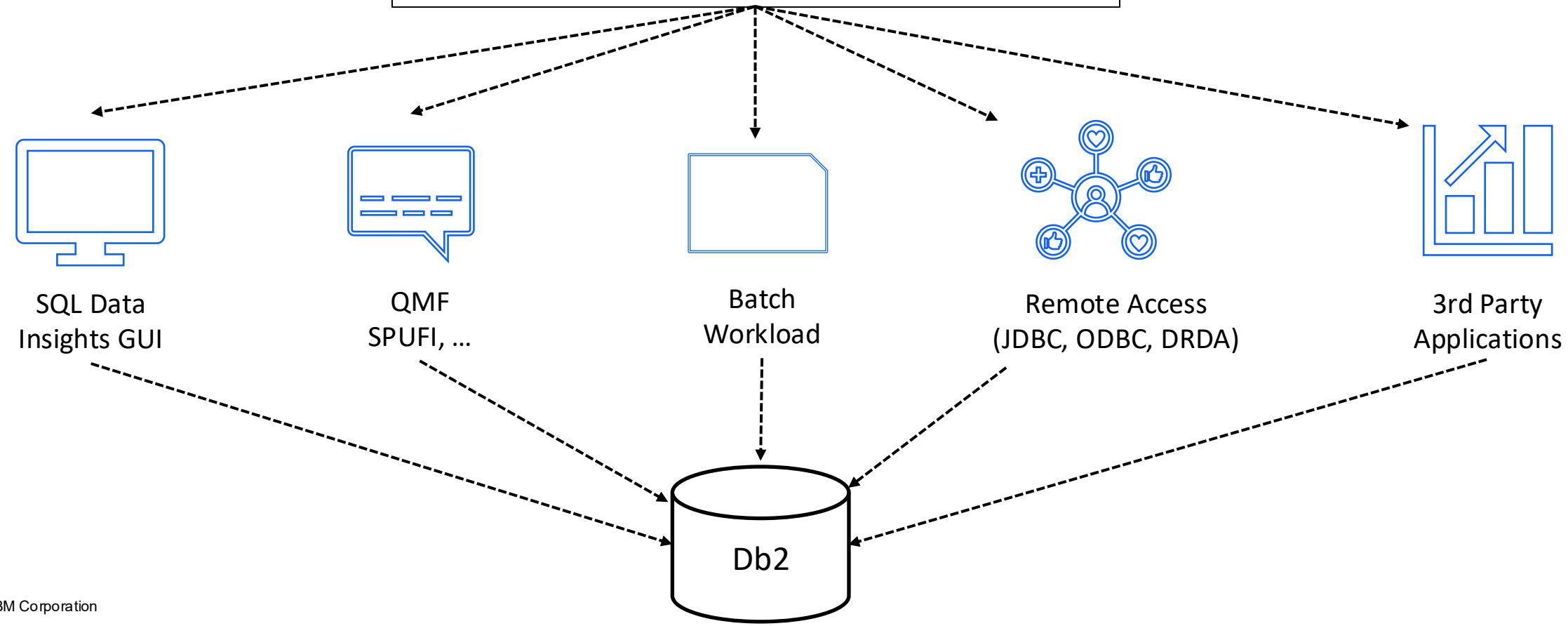
Closest to center of the model

Close to center of the model

Furthest from center of the model

Power any Db2 for z/OS application with AI enhanced SQL

```
SELECT CustomerID,
       AI_SIMILARITY('3668-QPYBK',
                     CustomerID) AS Report1
FROM USER.DATA_TABLE
ORDER BY 2 DESC
FETCH FIRST 5 ROWS ONLY;
```



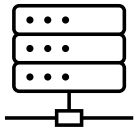
Introducing:

IBM SQL Data Insights Pro

AI-powered extension for IBM Db2 for z/OS that brings vector-based intelligence and semantic understanding directly to data on IBM Z - without moving data off-platform.

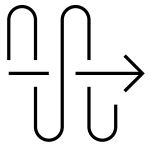
SQL Data Insights Pro

Key capabilities



Unified Data Embeddings

AI-powered analysis of structured and unstructured data (e.g., comments, logs, descriptions).

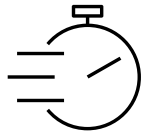


Optimized Training Infrastructure

Model training execution on IFLs for cost-efficient, high-performance workloads.

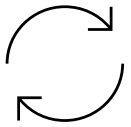
Builds on the foundation of SQL Data Insights

Extending its capabilities with additional advanced features



Accelerated Performance on IBM Z

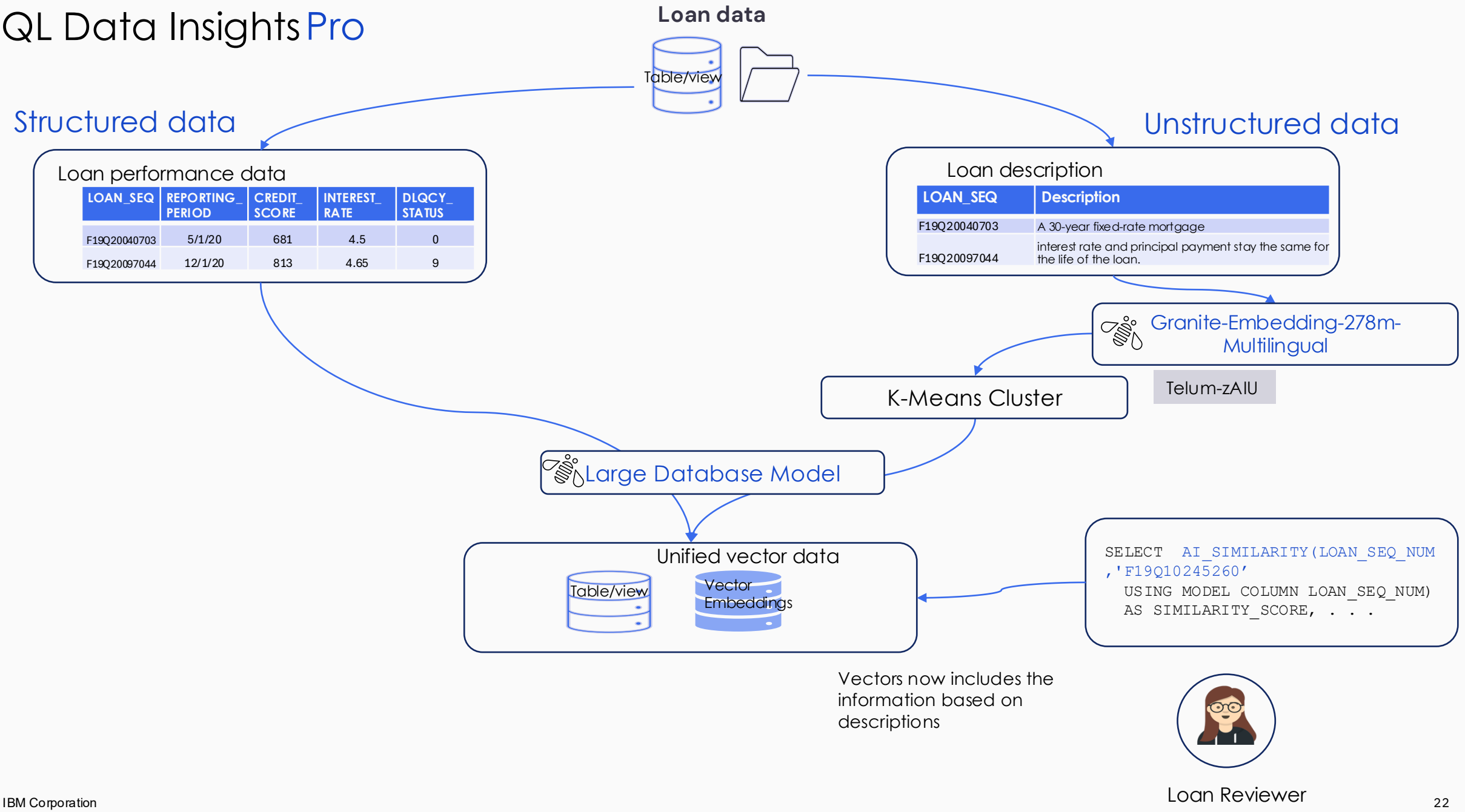
Integration with IBM Z hardware accelerators (AIU on Telum I & II) to optimize text embedding.



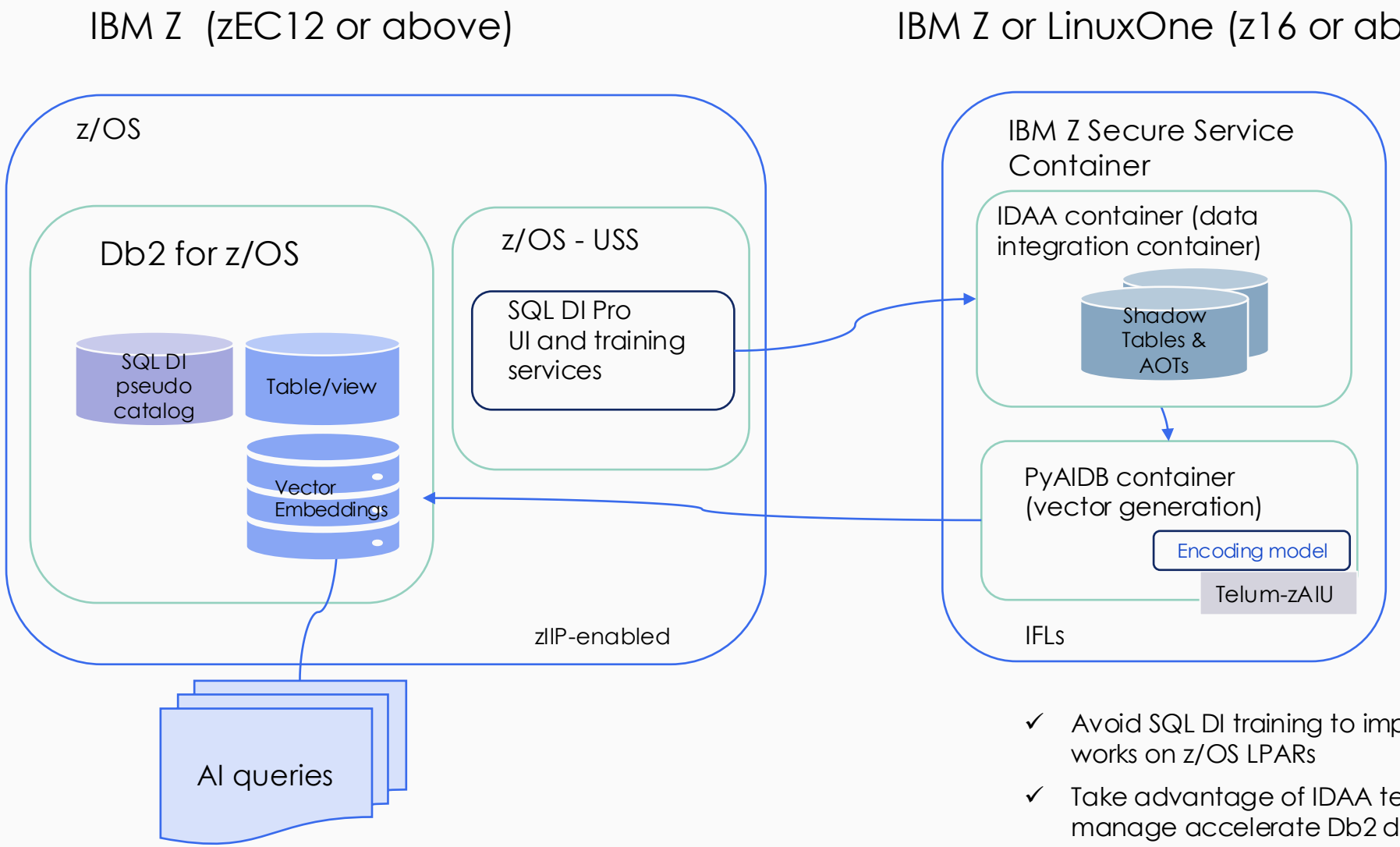
Incremental Retraining

Support for incremental model training using only changed data.

SQL Data Insights Pro



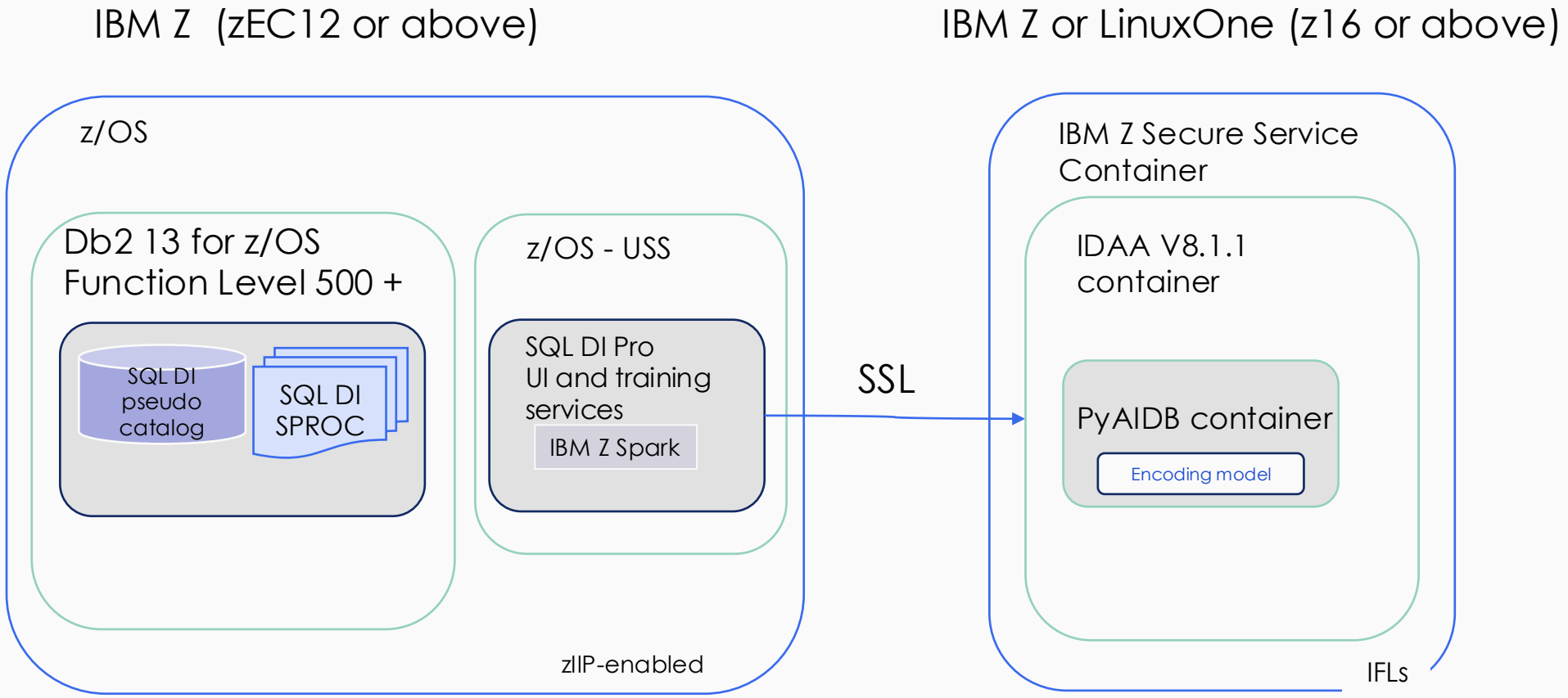
SQL Data Insights Pro – Architecture (high level)



Details
https://www.ibm.com/docs/en/sql-data-insights-pro/1.1.0?topic=overview-sql-di-pro

- ✓ Avoid SQL DI training to impact zIIP eligible works on z/OS LPARs
- ✓ Take advantage of IDAA technology to manage accelerate Db2 data
- ✓ Control resource (IFL and memory) between IDAA and PyAIDB containers

SQL Data Insights Pro – Installation & setup



Details
https://www.ibm.com/docs/en/sql-data-insights-pro/1.1.0?topic=installation-configuration-roadmap

- ✓ Db2 and UI services are same as SQL DI
- ✓ IDAA V8.1.1 comes with PyAIDB container
  - Requires license file to enable
- ✓ Training at zSpark is optional

IBM SQL Data Insights Pro

Change version

1.1.0

Show full table of contents

Filter on titles

Welcome to IBM SQL Data Insights Pro
Overview of SQL DI Pro
Planning for SQL DI Pro installation
System requirements
Installation and configuration roadmap
Installing and configuring SQL DI Pro
Enabling and running AI queries
Administering SQL DI Pro
SQL DI Pro REST APIs
SQL DI Pro shell commands
Db2 tables for SQL DI Pro
Db2 subsystem parameter for SQL DI Pro
DSNTIP81: Performance and optimization panel 2
MAX AI DATA CACHING field (MXAIDTCACH subsystem parameter)
Db2 built-in functions for SQL DI Pro
AI_ANALOGY
AI_COMMONALITY
AI_SIMILARITY
AI_SEMANTIC_CLUSTER
Db2 SQL statements for SQL DI Pro
CREATE FUNCTION (sourced)
CREATE FUNCTION (inlined SQL scalar)
CREATE FUNCTION (SQL table)
CREATE INDEX
CREATE MASK
CREATE PERMISSION
CREATE TABLE
CREATE VIEW

Download PDF

Roadmap for planning, installing, and configuring SQL Data Insights Pro

Last Updated: 2026-03-31

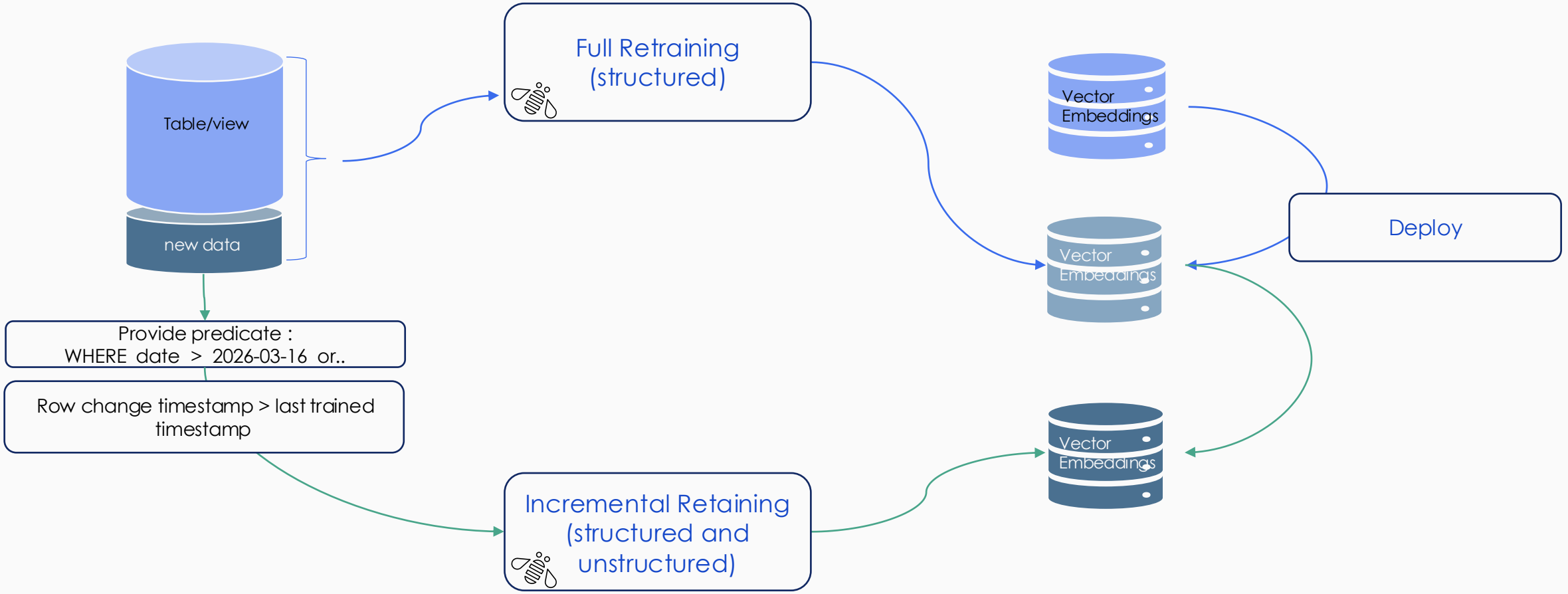
Getting IBM SQL Data Insights Pro (SQL DI Pro) up and running involves a sequence of planning, installation, and configuration tasks that might be best performed by experts in different roles and at different times. Create a high-level plan or roadmap to identify, track, and manage all the tasks and procedures.

Consider using the following roadmap for planning, installing, and configuring your SQL DI Pro:

Sequence	Task and instructions	Priority	Type	Role and skills
Step 1	Meeting system requirements for SQL DI Pro installation	Required	Planning, installation, configuration	z/OS® system administrator or programmer, Db2® system or database administrator, SQL DI Pro setup user
Step 2	Configuring system capacity and resources for SQL DI Pro	Required	Planning, configuration	z/OS system administrator, MVS system administrator
Step 3	Configuring network ports for SQL Data Insights Pro	Required	Planning, configuration	z/OS system programmer with Unix service system skills, network administrator
Step 4	Configuring setup user ID for SQL DI Pro	Required	Planning, configuration	z/OS system administrator or programmer with Unix service system skills, security administrator
Step 5	Configuring a keyring-based keystore (JCERACFKS) for SQL DI Pro	Required	Configuration	z/OS system administrator or programmer with Unix service system skills and z/OS keyring and keystore skills, network administrator, security administrator
Step 6	Configuring Db2 for SQL DI Pro	Required	Configuration	Db2 system or database administrator
Step 7	Enabling RACF PassTicket support for SQL Data Insights Pro	Optional *(Required if you use z/OS Spark for model training and PassTicket for authentication)*	Configuration	Db2 database, system, or security administrator
Step 8	Installing and configuring Db2 Analytics Accelerator for z/OS and dependencies	Required	Planning, installation, configuration	z/OS system, USS, security, network, and database administrators
Step 9	Configuring SSL connections for SQL Data Insights Pro	Required	Configuration	z/OS system, network, and security administrators
Step 10	Installing and configuring SQL DI Pro license	Optional *(Required if you plan to upgrade your existing Db2 SQL DI feature to the standalone SQL DI Pro)*	Installation, configuration	SQL DI Pro setup user, z/OS system programmer with Unix shell skills
Step 11	Installing and configuring the SQL DI server	Required	Installation, configuration	SQL DI Pro setup user, z/OS system programmer with Unix shell skills
Step 12	Verifying the installation and configuration of SQL DI Pro by using the web UI	Required	Installation, configuration	z/OS system administrator or programmer, Db2 system or database administrator, SQL DI Pro setup user
Step 13	Managing SQL DI Pro user permissions	Optional *(Required for post-installation configuration)*	Configuration, administration	Db2 administrator, SQL DI Pro administrator or setup user

The roadmap consists of the following elements:

SQL Data Insights Pro



SQL Data

Insights Pro

Next

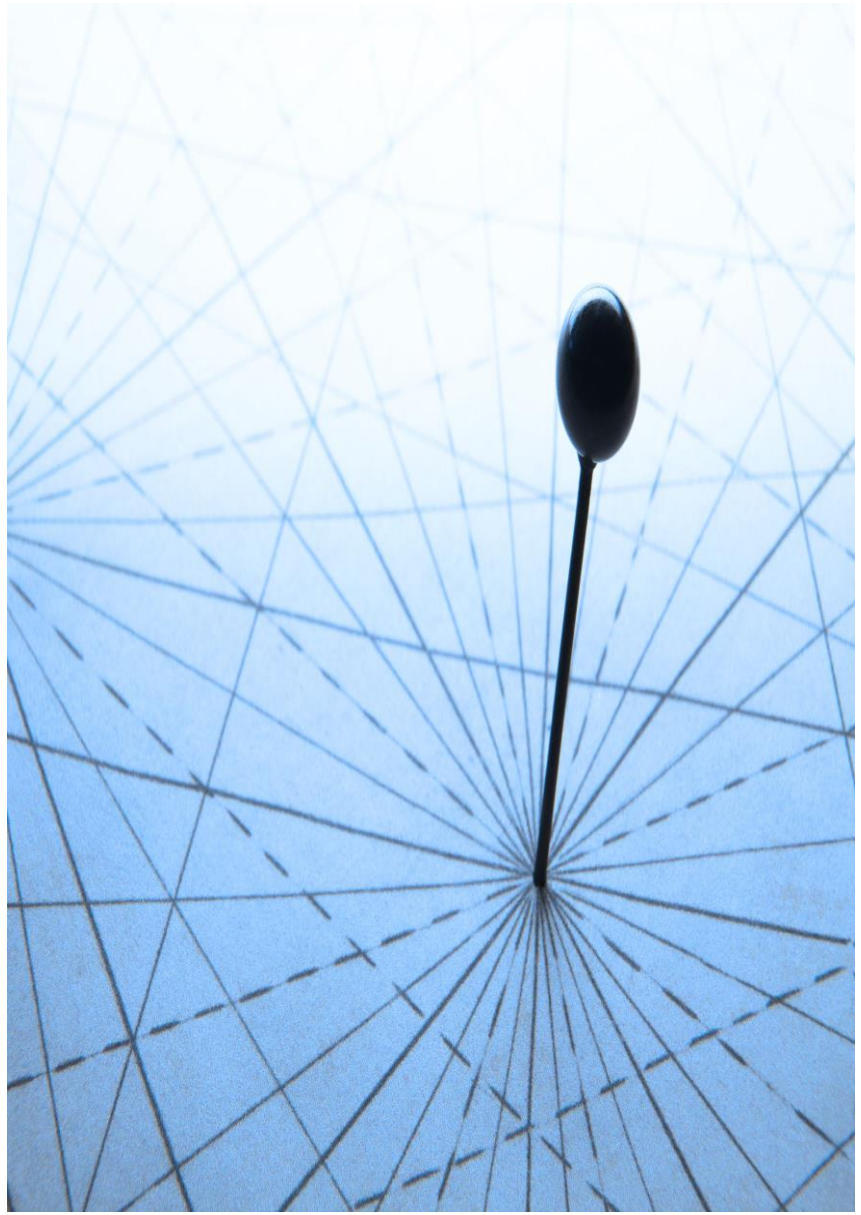
Further adaption of IBM Z AI acceleration

Model customization

Reasoning - detail explanation of ranking

Prediction

Combined with GenAI capability



SQL Data Insights Pro

Business Value

SQL Data Insights Pro adds AI models to Db2 that understand data meaning, enabling semantic queries—running securely on IBM Z without moving data.

More Relevant Insights

Get answers based on meaning — not just exact matches

Faster decisions

AI insights directly from Db2

No waiting on data pipelines

Low cost

No data movement

No duplication

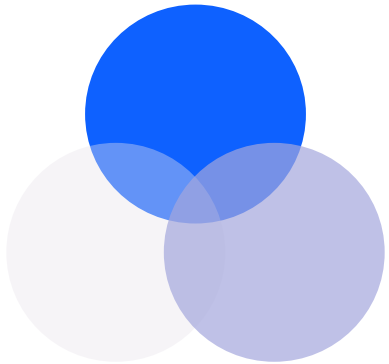
No additional infrastructure

Reduced Risks

Data stays secure and governed on IBM Z

Broader Access to Insights

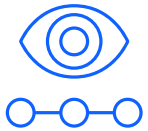
Use familiar SQL to explore AI-driven insights



Use Case

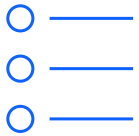
Credit card fraud detection

Industry: Banking



Problem

Fraud teams face massive alert volumes and high false positives(1). Traditional rules miss subtle behavior patterns, slowing investigations.



SDI Pro Capability

Semantic similarity search (finding things that are meaningfully alike, not just exactly the same words or values) across transactions, customer behavior, and merchant descriptions to instantly retrieve past fraud cases that resemble new activity.

Business Outcome

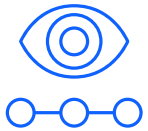
- Faster fraud triage and investigation
- Reduced false positives
- Analysts focus on truly high-risk cases

(1) False positive : The system says “this is bad”, but it’s actually normal

Use Case

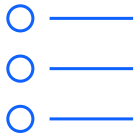
High risk claims detection

Industry: Insurance



Problem

Claims Adjusters rely heavily on experience to judge whether a claim is typical or risky, leading to inconsistency and long cycle times.



SDI Pro Capability

SQL DI Pro compares new claims against historical claims with similar structured attributes and unstructured adjuster notes, retrieving comparable outcomes instantly.

Business Outcome

- Faster, more consistent claims decisions
- Earlier anomaly and risk detection
- Reduced claim leakage(1) and cycle time(2)

(1) Claim leakage happens when the insurance company pays too much, pays incorrectly, misses fraud or mistakes

(2) Cycle time is how long it takes to process and settle a claim. Shorter = better

SQL Data Insights Pro

Additional Use Cases

Banking:

Apply semantic similarity and anomaly detection to identify transactions, accounts, or behaviors that resemble known risk patterns or deviate from learned norms.

- AML alert prioritization
- Post-transaction anomaly analysis
- Portfolio similarity (mutual fund analysis)
- Fraud ring similarity analysis

Insurance:

Use clustering and behavioral comparison to surface abnormal claim activity and detect patterns across structured claim data and adjuster notes.

- Upsell Opportunities Identification
- Claims anomaly clustering
- Underwriting risk segmentation
- Premium adequacy pattern analysis

Healthcare:

Identify patients with similar conditions or risks to improve care decisions and intervene earlier—using both medical records and clinical notes.

- Find patients similar to those who were readmitted
- Identify who may need follow-up care after discharge
- Detect unusual treatment or billing patterns
- Group patients with similar conditions for better care planning

Cross Industry:

Apply similarity and clustering to identify behavioral affinities that inform retention, targeting, and portfolio strategy.

- Customer churn identification
- Cross-sell / product affinity analysis
- Loyalty segmentation refinement
- Account behavior clustering

Industry Starter Kits

Accelerate adoption of SQL Data Insights Pro with ready-to-use industry use cases

What's included

- Curated sample datasets
- Prebuilt models
- Ready-to-run Jupyter notebooks for exploration
- Prebuilt semantic SQL examples leveraging AI functions

Availability

- Initial release focused on Financial Services (banking and insurance)
- Additional industries, including Healthcare, coming next
- Open for collaboration to expand use cases and industry coverage

What this enables:

Explore real use cases
Work with practical industry scenarios—not abstract concepts

Get started fast
See results in hours with ready-to-run data and notebooks

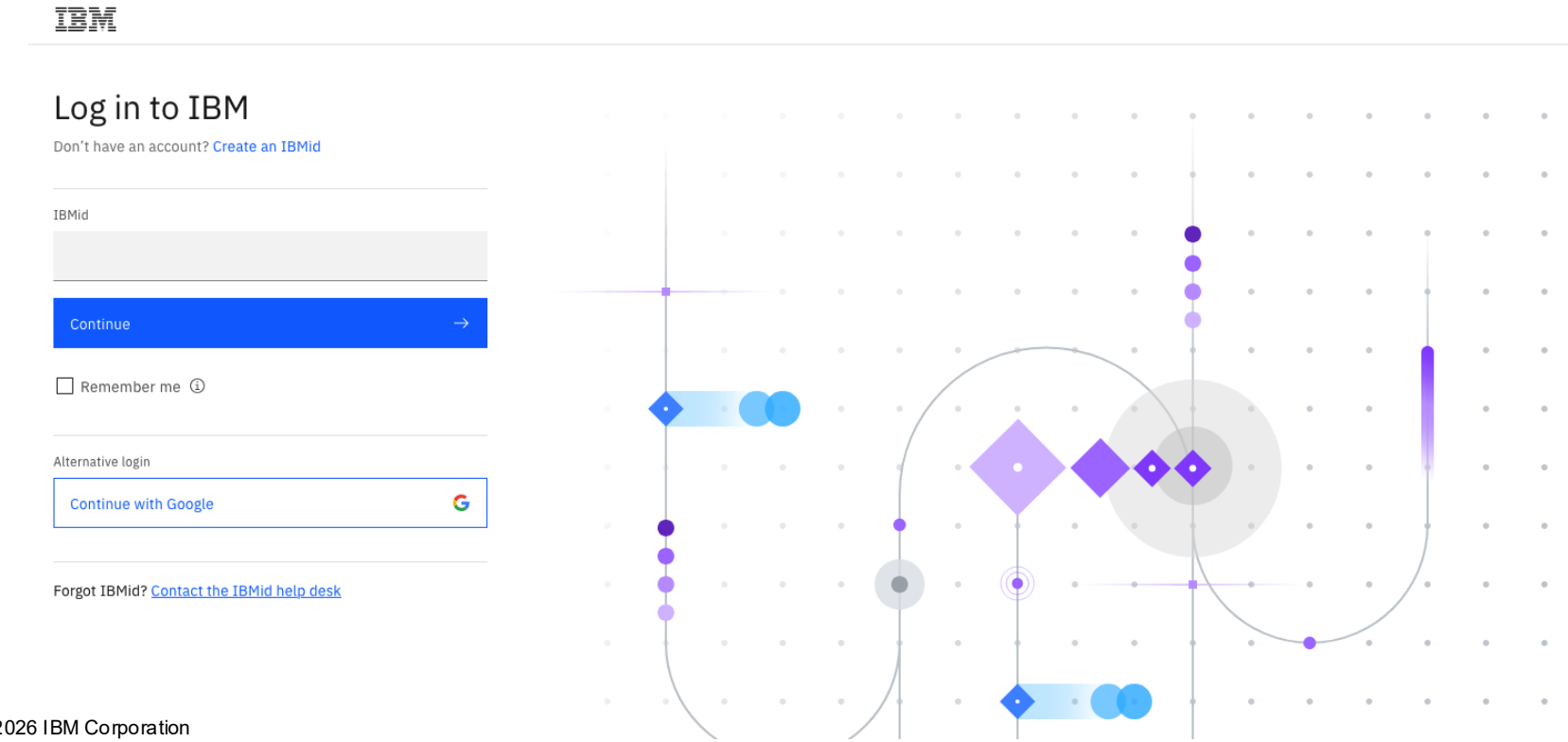
Accelerate adoption
Apply AI-powered SQL to your own data with minimal setup

How to download (free!)

Step 1 : Paste the following URL to your browser

https://ibm.biz/sqldipro_industrystarterkit

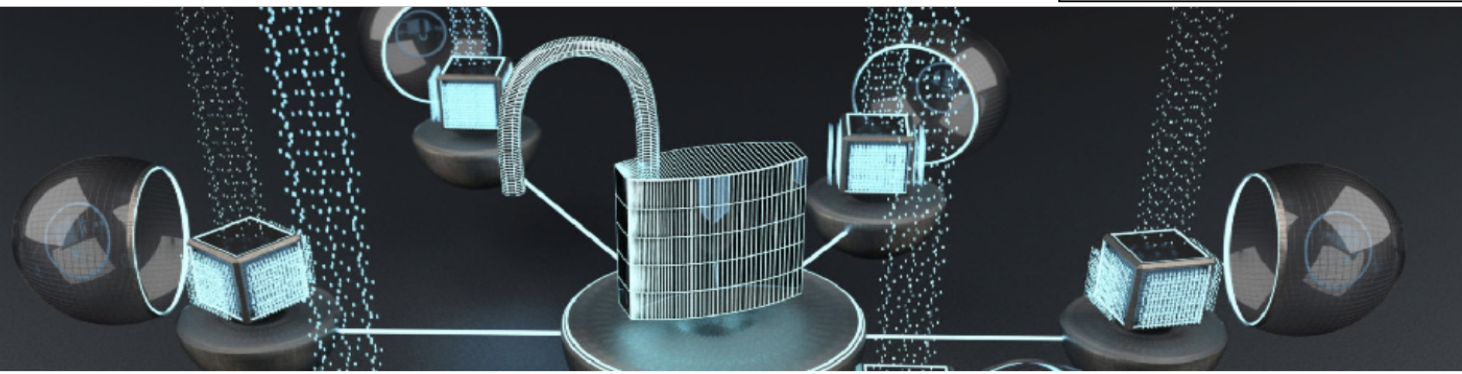
Step 2 : Log on with your IBM ID



IBM Db2 for z/OS AI Enterprise Hub

Assets and solutions to accelerate Db2 for z/OS transformation in the AI era

Marketing Registration Services



Downloads

By clicking `Download` you agree that you have had the opportunity to review the terms and conditions and that such terms and conditions govern this transaction

IBM Db2 for z/OS AI Enterprise Hub for Multiplatform
English
2026-03-20

Show 10 entries

Search:

Description	Filename	Size	Action
IBM SQL Data Insights Pro Industry Starter Kit - Credit Card Fraud	credit_card_transaction.zip	171.2 MB	Download
IBM SQL Data Insights Pro Industry Starter Kit License	License.txt	KB	Download
IBM SQL Data Insights Pro Industry Starter Kit - Anti-Money Laundering Transactions	aml_transaction.zip	502.7 MB	Download
IBM SQL Data Insights Pro Industry Starter Kit - Insurance Claims	insurance_claims.zip	609.4 MB	Download
IBM SQL Data Insights Pro Industry Starter Kit - Insurance Underwriting	insurance_underwriting.zip	510.3 MB	Download

- Credit Card Fraud
- Anti-Money Laundering
- Insurance Claims
- Insurance Underwriting

Need help?

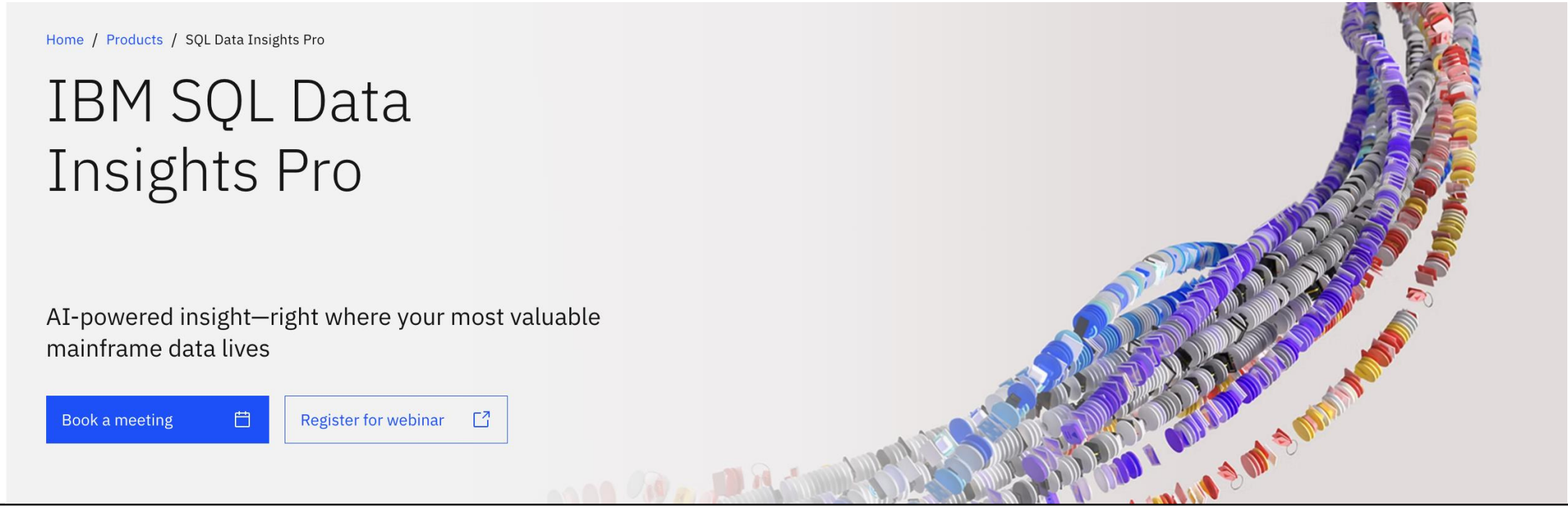
Sign up support (English only)

Sign up and Software Download FAQ

Client feedback

I did a small demo based on the industry pack for credit card transactions. The audience were some of the business and IT professionals working on credit card fraud detection and prevention. They were very interested and asked a lot of questions. The industry pack was a great help to open their minds.

Learn more about IBM SQL Data Insights Pro



01

Product Page
LINK

02

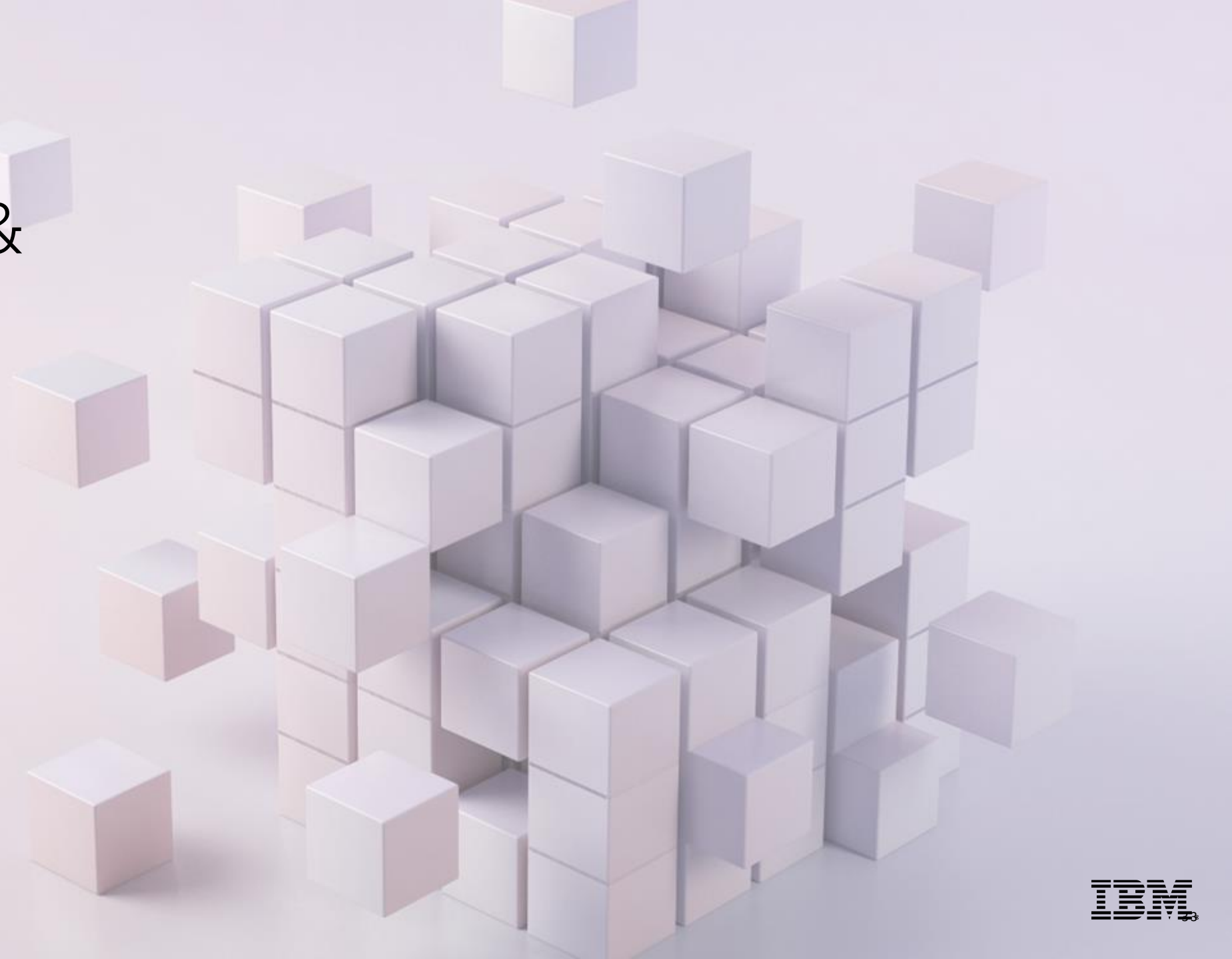
Documentation
LINK

03

Industry Starter Kit
LINK

Questions &

Answers



Notes and disclaimers

U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. This document is distributed “as is” without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity. IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts.

In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.”
Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used products and the results they may have.
References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer’s responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer’s business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law


IBM