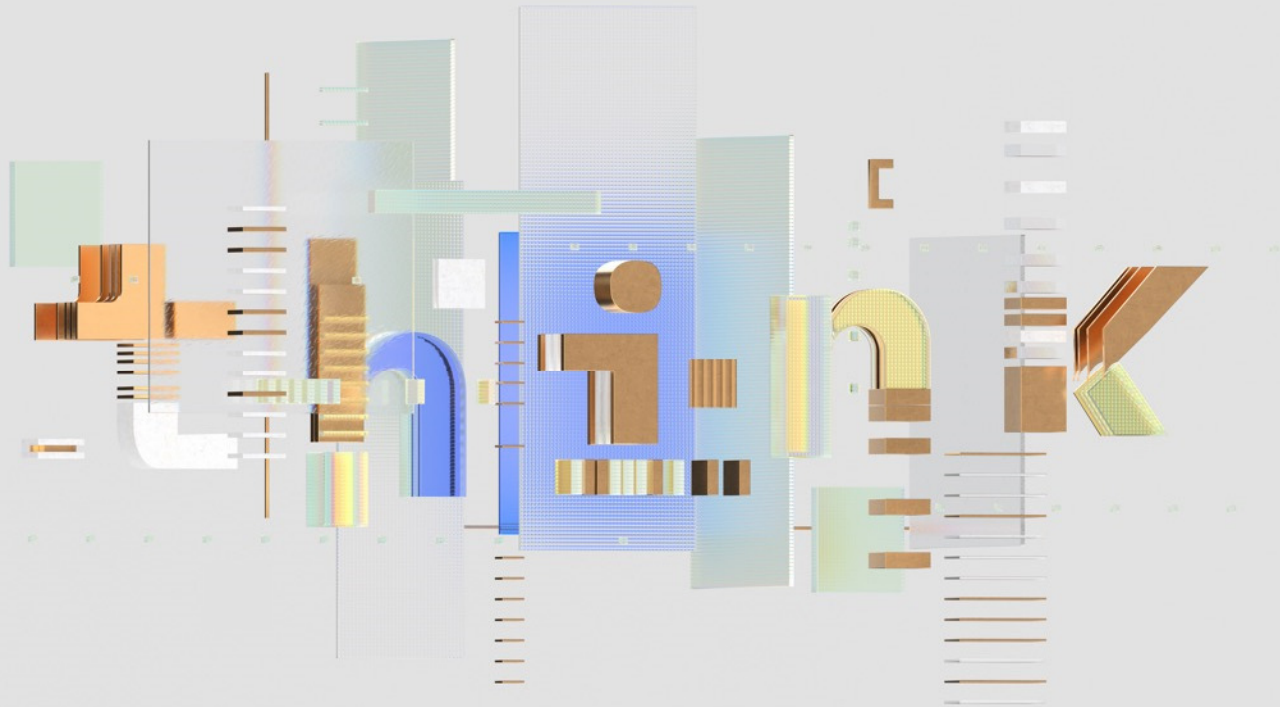


IBM Db2 – The AI database

think 2019



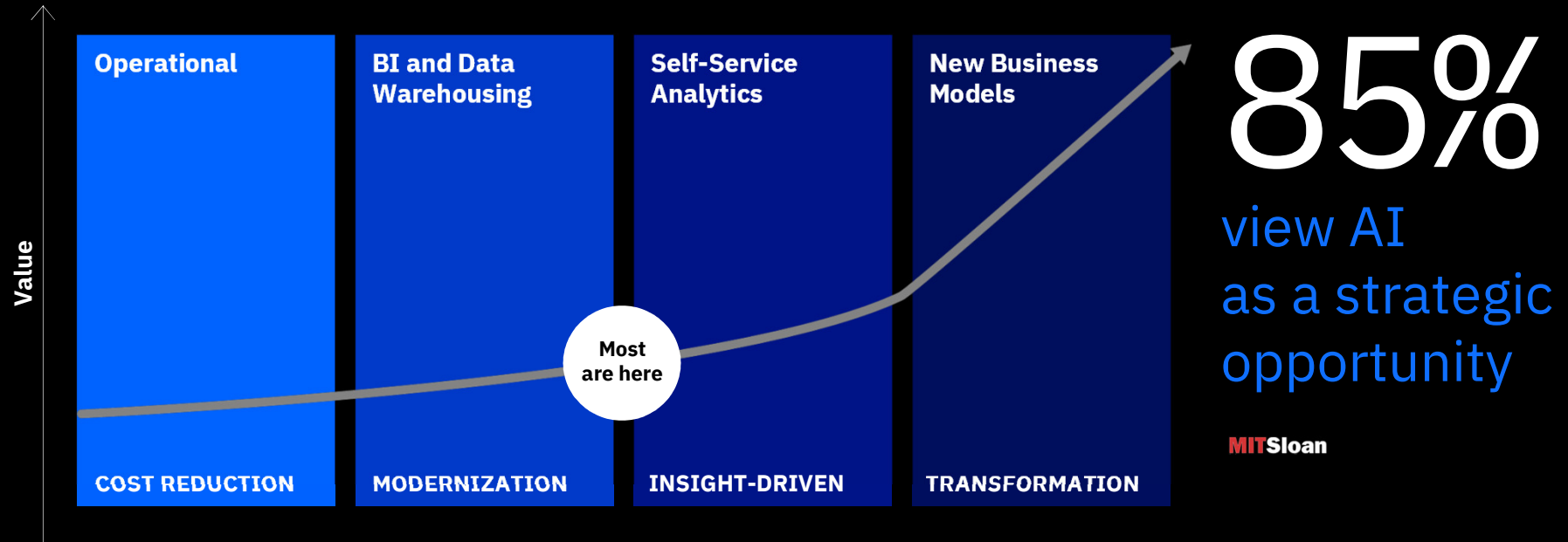
Pandit Prasad – Offering Manager, Db2 & Big Data

All business is becoming more data-driven



**“Data will be your
basis of competitive advantage”
- Ginni Rometty**

Our clients have declared the journey to AI a strategic priority



33%

Accessing and preparing data is a barrier to using machine learning

*451 Research Voice of the Enterprise: AI & Machine Learning, 2019

88%

AI and ML important for their data platform and analytics initiatives

*451 Research Voice of the Enterprise: AI & Machine Learning, 2019

IBM delivers the capabilities you need to build a ladder to AI



Infuse: Automate & scale across your processes

Trust: Achieve trust & transparency in outcomes

Analyze: Scale insights with ML everywhere

Organize: Create a trusted analytics foundation

Collect: Make data simple & accessible



IBM proudly presents:

IBM Db2

The AI database

- ***Transform*** your business with AI infused data management
- ***Deliver*** data to the business more quickly and with higher confidence
- ***Create*** AI apps faster with tools that make data and analytics accessible for AI development

The foundation of AI is data architecture. Be ready for everything that comes next,

with

the new generation of IBM

Db2

IBM Db2 - the AI database

Powered By AI



Confidence-based query results
leveraging ML-SQL



Faster data exploration
by using NLQ in Augmented Data Explorer



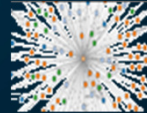
Up to 10x better query performance
powered by an ML-Optimizer



Build AI based applications
with Python, GO , JSON and Jupyter notebooks



single view on all data
delivered by Data Virtualization



Model Complex Relationships
by using Db2 Graph and SQL



Auto resource optimization
delivered by Adaptive Workload Management



Blockchain Ready
using Db2 Blockchain Connector

Db2 – The database powered by AI

Enables a spectrum of automated tasks for the administrator

Figure 5: The spectrum of AI use cases as they relate data management

Source: 451 Research



Automating repetitive tasks:

Automated data identification and tagging.

Augmenting humans to do everything in between:

Automation to make databases easier and more intuitive to work with.

Automating complex tasks:

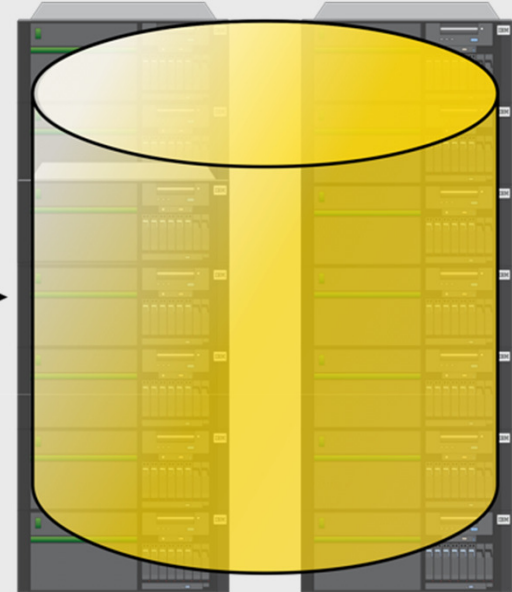
Integrated support for the development of data science and AI applications

Db2 scales ... easily and is optimized for maximum performance

Db2 is the only database that offers exclusive features for maximum performance



**Go from
gigabytes to petabytes ...
without compromising
performance**



- Industry leading cost optimizer
- ML Optimizer (New)
- Adaptive Workload Management (New)

Db2 - the only database that offers ML Optimizer in addition to Cost Optimizer

New

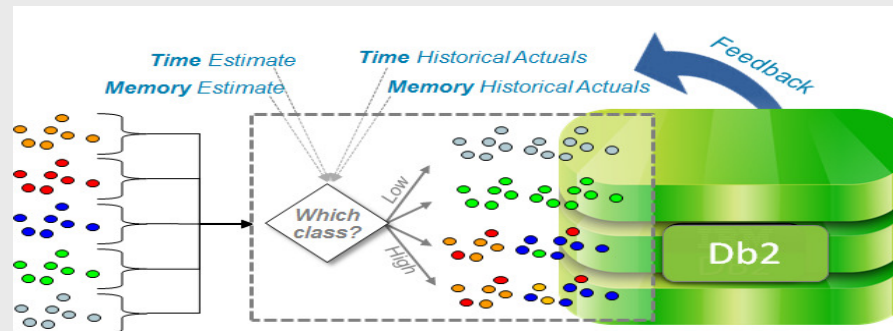
*Machine Learning
Optimizer that
improves from
experience to make
query optimization
simple, reliable and
stable*

- Number of rows flowing through the various operations impacts performance for most common issues and is calculated via cardinality estimates
- Basic tuning to improve cardinality estimation is high impact
- **Phase 1 – Cardinality Estimation**
 - Initial phase – Support equality & range local operators with no expressions
 - Future phases – Cardinality support for expressions, predicates & BLU, support for join enumerations,

Db2 delivers automatic workload management

Ensures system stability and responsiveness with zero tuning

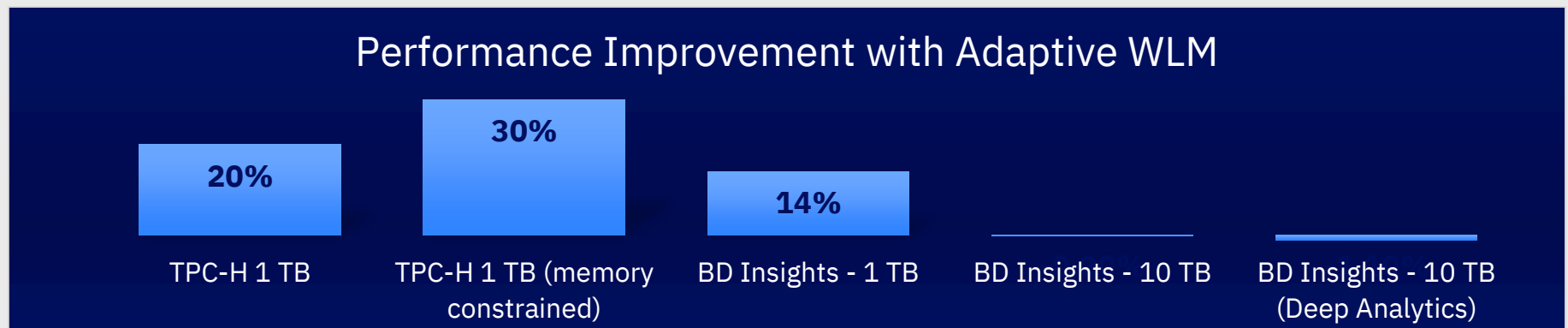
- **Adaptive Workload Management** is the “next generation” of Db2 workload management (WLM)
- Delivers automatic workload management within Db2 that ensures system stability and responsiveness with zero tuning
 - Make sure that the system is well-utilized but don't let it become overcommitted
 - Schedule jobs appropriately to ensure fairness and appropriate responsiveness



Some performance numbers for Adaptive WLM

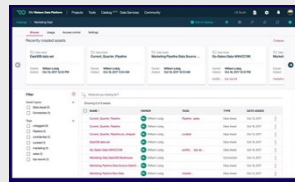
Analytical workloads performed at par or better with Adaptive WLM compared to current WLM using default concurrency thresholds across a set of internal workloads

- Since work is admitted into the system based on matching of CPU and memory needs versus current usage, processing of work is generally more efficient due to smoother, steadier consumption patterns
- System stability is maintained throughout

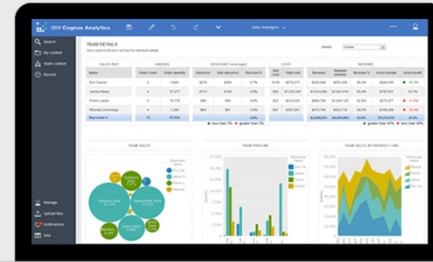


Db2 Data Virtualization is the next generation of federation

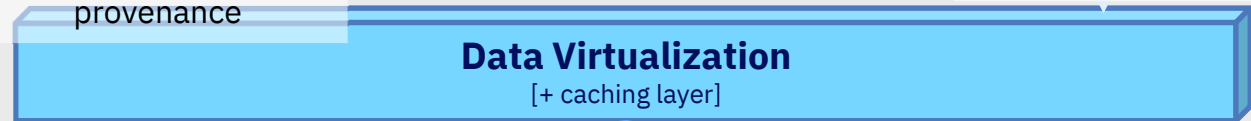
Data Virtualization will offer sophisticated capabilities to work with data distributed across multiple data stores



Unified data asset catalog, lineage and provenance



Unified access control and security policies



Data Warehouses & Marts

Relational Databases

No SQL

Spreadsheets & Text files

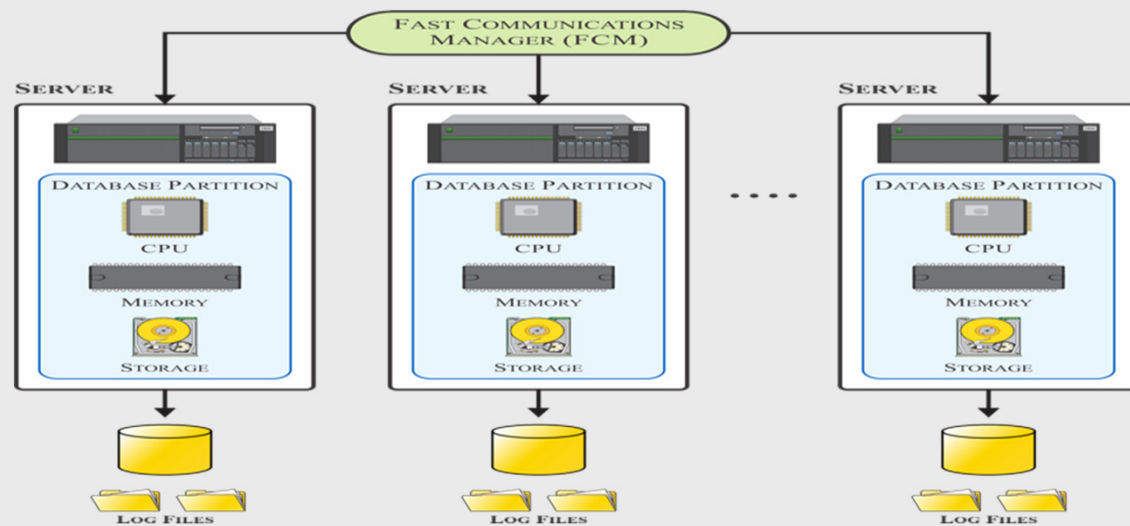
Big Data (Hadoop)

Locations: Private and public clouds, standalone systems, worldwide.

By the way we have auto-sharding too!

Sharding within Db2 (also known as DPF) lets you partition your database across multiple servers or within a large SMP server.

- *This allows for peta-byte scalability, since you can add new machines and spread your database across them.*
- *That means more CPUs, more memory, and more disks from each of the additional machines for your database!*
- *Ideal to manage large databases, whether you are doing data warehousing, data mining, or online analytical processing (OLAP)*



An AI Database automation unleashed

Machine learning has a key role to play in improving data base efficiency, performance and accuracy

Important but highly repetitive tasks can be accelerated through automation.

Improve query performance and accuracy



Increase system stability and reliability reducing manual tuning and monitoring of databases.

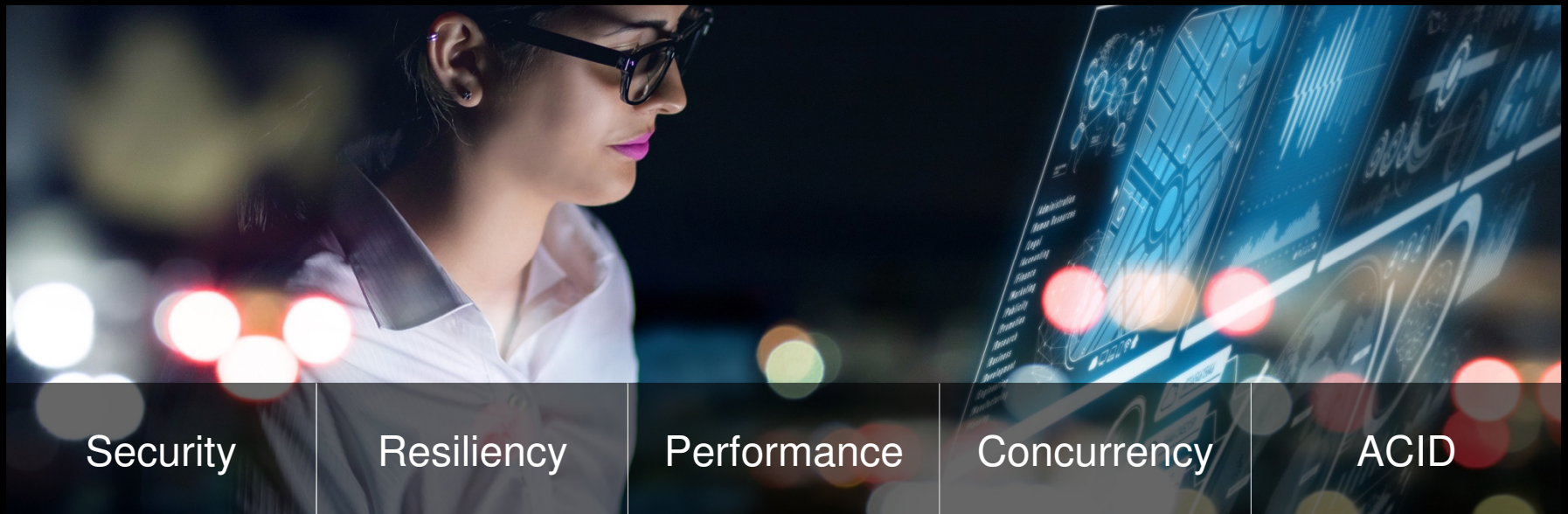


Decrease the time it takes to generate insights as well as make higher confidence-based business decisions.



Db2 – The database optimized for AI

Support modern AI applications & workloads ...



Security

Resiliency

Performance

Concurrency

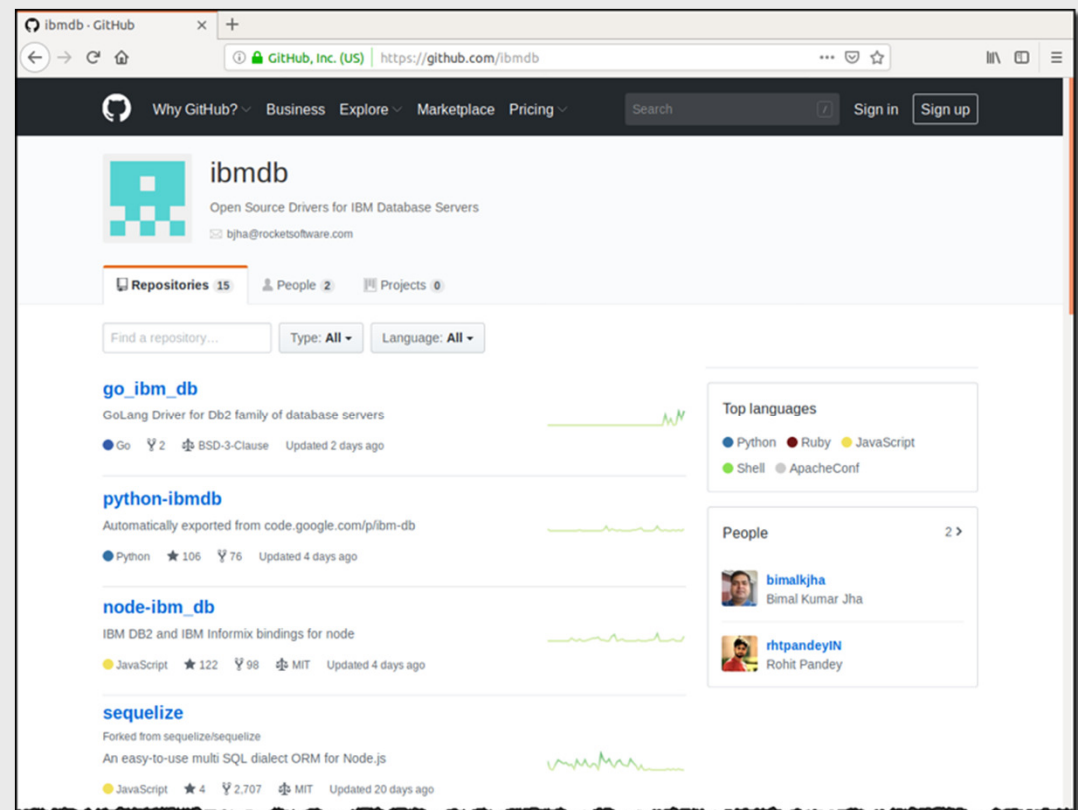
ACID

... while offering a single comprehensive view on all data

Developers and data scientists can now use Db2 for their AI applications

Eight open source drivers available from github -
<https://github.com/ibmdb>

- Go
- Python (Django, SQLAlchemy & Alembic)
- Jupyter Notebook
- Sequelize
- PHP
- Java (Spring)
- Node.js
- Ruby



Follow “Code Patterns” to build cognitive applications today

NEW!

CODE PATTERN | OCT 19, 2018

Train a cloud-based machine learning model using on-premise data

[Get the Code](#)

Analytics + Artificial Intelligence

CODE PATTERN | OCT 22, 2018

Continuous learning with Watson Machine Learning and IBM Db2 Warehouse on Cloud

[Get the Code](#)

Apache Spark + Artificial Intelligence

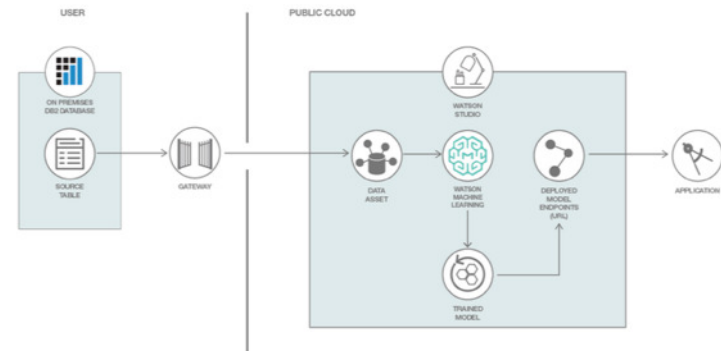
Train a cloud-based machine learning model from an on-premise database

build failing

Many companies and individuals struggle to use their on-premises data — the kind of data that lives on a local machine, within your data center, behind your firewall — for machine learning in the cloud. It can be challenging to find a quick, easy, and secure solution for connecting resources in a protected environment to resources in the cloud.

With Watson Studio & Machine Learning, Db2, and Secure Gateway, it is possible to establish a secure, persistent connection between your on-premises data and the cloud to train machine learning models leveraging cloud computing resources like Spark, elastic environments, and GPUs.

In this guide, we will create an on-premises Db2 database on our local computer, populate it, and then connect it to Watson Studio via Secure Gateway. Then, we'll read Buildings Violations data from this database and build a model to predict the likelihood that a particular building will FAIL an inspection based on historical data from the City of Chicago. After we build the model, we will deploy it as an API endpoint with Watson Machine Learning that only authorized users can access.



Db2 Code Patterns – Learning by Doing

(<https://developer.ibm.com/components/ibm-db2-database/>)

IBM Developer

Topics ▾

Community ▾

More open source at IBM ▾

NEW!

IBM Db2 Database

IBM's family of products including operational database, data warehouse, data la

CODE

Code Patterns

CONTENT

Announcements

Articles

Series

Tutorials

Series

Learning Path: Db2 for AI

This learning path demonstrates how data engineers and data scientist can predict the price of a house based on historical data. Using code patterns with sample code, you'll learn about...

Aug 09, 2019

Databases

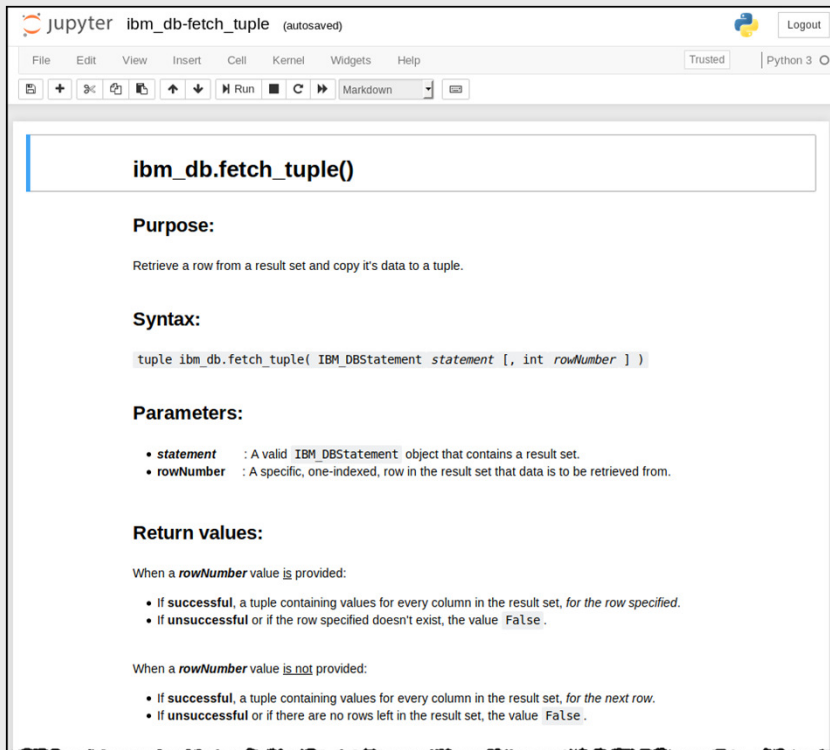
IBM Db2 Database



<https://developer.ibm.com/announcements/training-machine-learning-models-in-watson-studio/>

NEW!

... and an extensive set of Jupyter Notebooks to kick start your Db2 development



ibm_db.fetch_tuple()

Purpose:

Retrieve a row from a result set and copy its data to a tuple.

Syntax:

```
tuple ibm_db.fetch_tuple( IBM_DBStatement statement [, int rowNumber ] )
```

Parameters:

- **statement** : A valid `IBM_DBStatement` object that contains a result set.
- **rowNumber** : A specific, one-indexed, row in the result set that data is to be retrieved from.

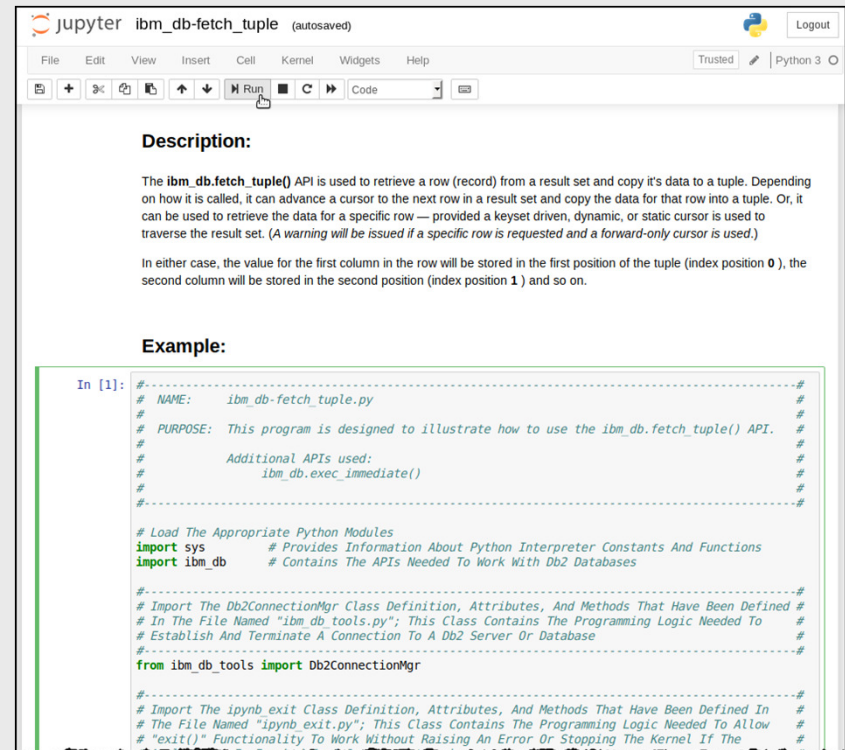
Return values:

When a **rowNumber** value is provided:

- If **successful**, a tuple containing values for every column in the result set, for the row specified.
- If **unsuccessful** or if the row specified doesn't exist, the value `False`.

When a **rowNumber** value is not provided:

- If **successful**, a tuple containing values for every column in the result set, for the next row.
- If **unsuccessful** or if there are no rows left in the result set, the value `False`.



Description:

The `ibm_db.fetch_tuple()` API is used to retrieve a row (record) from a result set and copy its data to a tuple. Depending on how it is called, it can advance a cursor to the next row in a result set and copy the data for that row into a tuple. Or, it can be used to retrieve the data for a specific row — provided a keyset driven, dynamic, or static cursor is used to traverse the result set. (A warning will be issued if a specific row is requested and a forward-only cursor is used.)

In either case, the value for the first column in the row will be stored in the first position of the tuple (index position 0), the second column will be stored in the second position (index position 1) and so on.

Example:

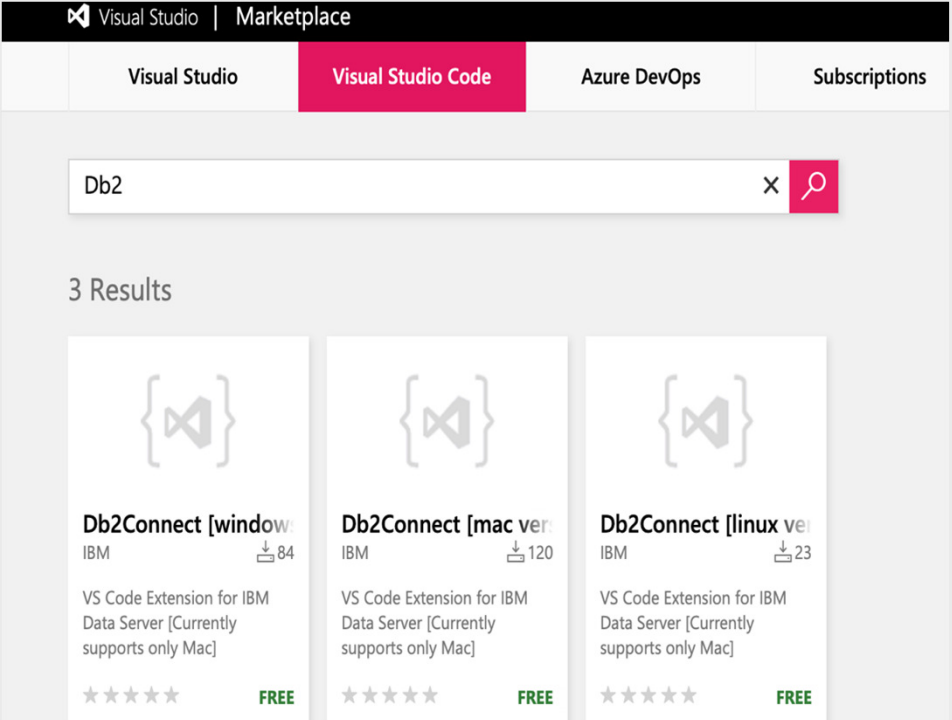
```
In [1]: .....#
# NAME:      ibm_db-fetch_tuple.py
#
# PURPOSE:   This program is designed to illustrate how to use the ibm_db.fetch_tuple() API.
#
#           Additional APIs used:
#           ibm_db.exec_immediate()
#
# .....#
# Load The Appropriate Python Modules
import sys      # Provides Information About Python Interpreter Constants And Functions
import ibm_db   # Contains The APIs Needed To Work With Db2 Databases
#
# .....#
# Import The Db2ConnectionMgr Class Definition, Attributes, And Methods That Have Been Defined
# In The File Named "ibm_db_tools.py"; This Class Contains The Programming Logic Needed To
# Establish And Terminate A Connection To A Db2 Server Or Database
# .....#
from ibm_db_tools import Db2ConnectionMgr
#
# .....#
# Import The ipynb_exit Class Definition, Attributes, And Methods That Have Been Defined In
# The File Named "ipynb_exit.py"; This Class Contains The Programming Logic Needed To Allow
# "exit()" Functionality To Work Without Raising An Error Or Stopping The Kernel If The
```

<https://github.com/IBM/db2-python>

We just announced VSCode Extension for Db2!

NEW!

- Based on Stack Overflow Survey 2019, Visual Studio code was voted as the favorite IDE amongst the developer community
- This extension provides a graphical user interface designed to enable development with Db2 from within Visual Studio so that developers can accelerate and simplify application development.
- The extension is named “db2connect”



The screenshot shows the Visual Studio Marketplace interface. At the top, there are navigation tabs for 'Visual Studio', 'Visual Studio Code' (which is highlighted in pink), 'Azure DevOps', and 'Subscriptions'. Below the tabs is a search bar containing the text 'Db2'. The search results section displays '3 Results'. Each result is a card for the 'Db2Connect' extension by IBM, available for Windows, Mac, and Linux. Each card shows the extension icon, the name 'Db2Connect [platform]', the publisher 'IBM', the number of downloads (84 for Windows, 120 for Mac, 23 for Linux), a description 'VS Code Extension for IBM Data Server [Currently supports only Mac]', a five-star rating, and the word 'FREE'.

LedgerSQL – Query your Blockchain data hassle-free

What

Empowering end-users to query Blockchain data leverage their existing skills (SQL) seamlessly just like any other traditional database.

Why

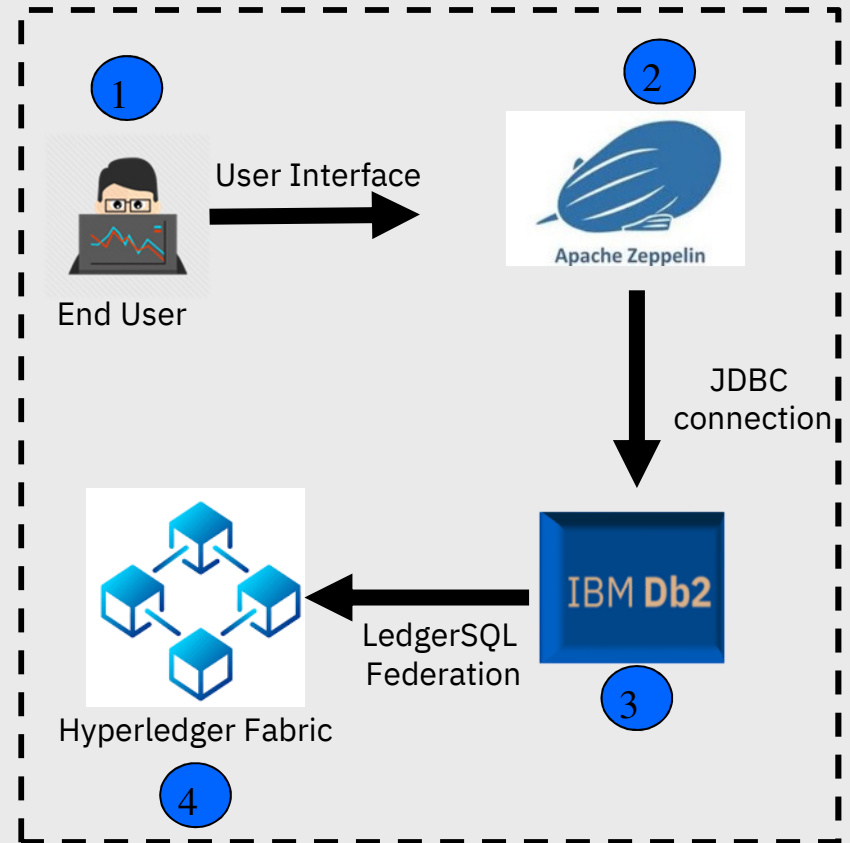
Make it easy for adopting blockchain, perform analytics/reporting on it and integrate with other data stores to give 360 view of information.

How

Build a Db2 Federation Wrapper to offer SQL querying capabilities on Hyperledger Fabric shared ledgers using the Fabric SDK & security model

Use Db2 SQL to query Blockchain

1. A Db2 federation wrapper that provides:
 - Connectivity to Blockchain Transactions and state databases via Hyperledger Rest API
 - Rewriting SQL queries into equivalent Hyperledger API calls to query data.
2. Using Zeppelin as end user tool to connect to Db2
3. Query Blockchain via Zeppelin using SQL.



Combine Blockchain data with other data stores for Analytics

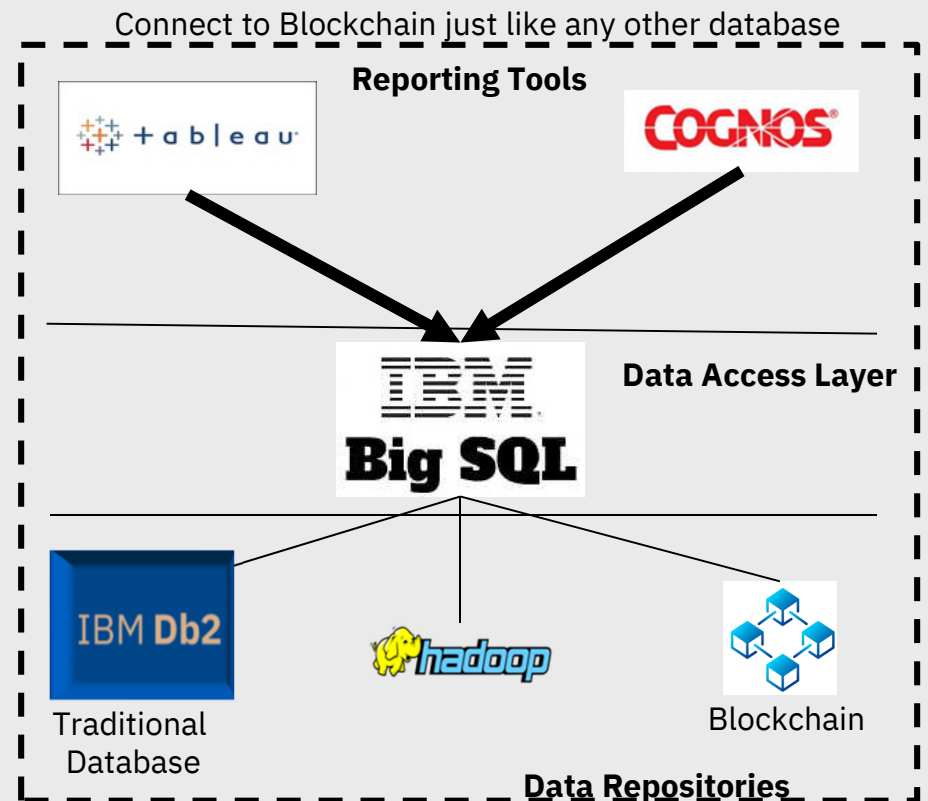
Easy and Secure: (Addressing the Skill Gap)

-In place querying, reporting & analytics of Blockchain data using the rich SQL language support provided without compromising the Blockchain security model.

Data Virtualization: (Fits into Analytics Infrastructure)

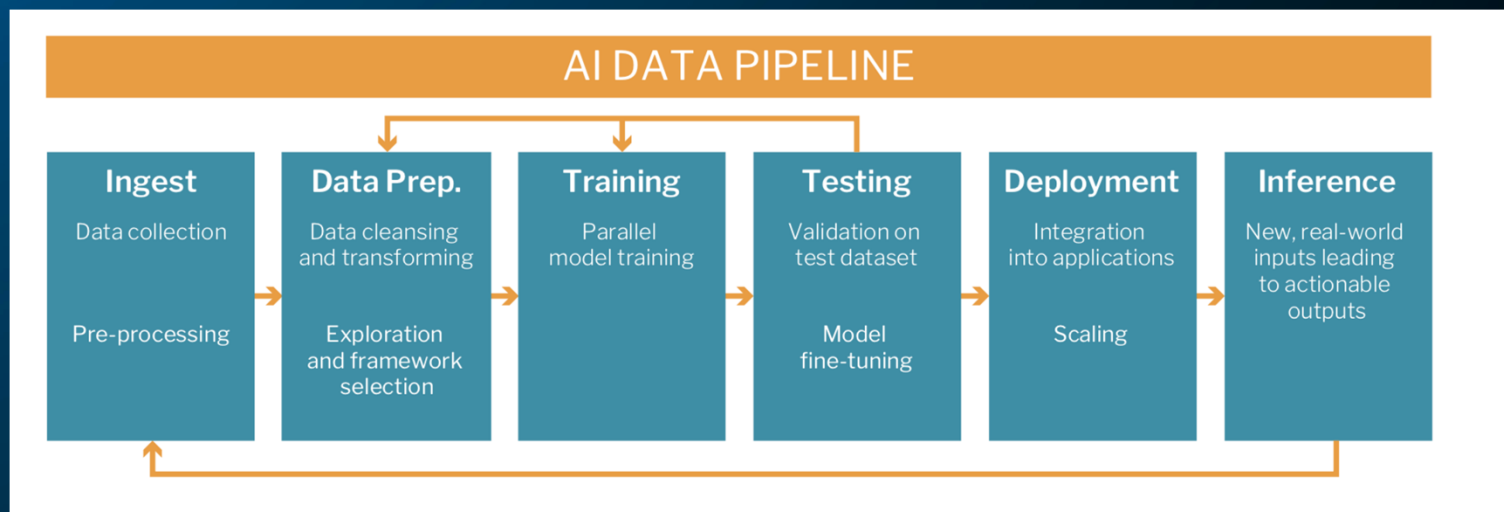
-Joining Blockchain data with other enterprise data set (Hadoop, Hive, Spark, external RDBMs) for complex use cases without additional work

-Immediate support for all DB2 based products through Federation



Data Exploration

“accessing and preparing data is one of the most significant barriers to ML adoption...database and data management products that accelerate data exploration can lay the foundation for more efficient development of AI applications” - 451 Research, Accelerating AI with Data Management, Accelerating Data Management with AI, 2019



Use Augmented Data Explorer tool for rich and deeper insights



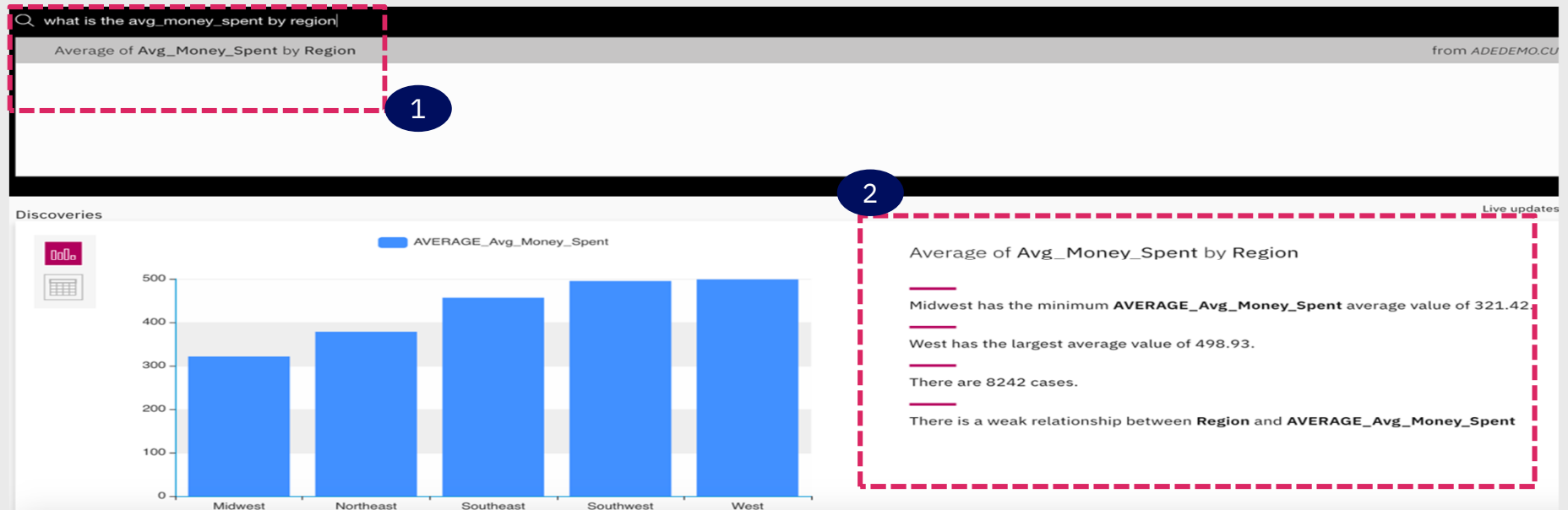
NEW!

*Get to the insights
faster – Use
Natural Language
Query (NLQ) for
data analysis*

Conversational analytics is the future...

- Data scientists often face the challenge of exploring data sets that are not familiar to them
 - Time consuming, Not sure what to look for
 - Need understanding of what is in the data set
- Need an intuitive data exploration tool
 - Alexa, Google Home, etc. prove trend
 - Natural extension of conversation to data
 - Provide multiple views of the data with minimal touch
- Free tool available to download - <https://www.ibm.com/us-en/marketplace/db2-augmented-data-explorer>
 - Currently in Beta
 - Seamlessly plugs and plays with Db2

Capabilities: Understanding NLQ to get accurate results



1. "avg sales by region" translates to: `SELECT AVG(SALES), REGION FROM xyz GROUP BY REGION`
2. Ability to automatically interpret the search result and create natural language based analysis that can be understood by user

Capabilities: Relevance

1. Understanding the user intent and providing type ahead queries
2. Providing suggestions for similar queries to the user
3. Providing source table of the data that the query will run against
4. Ability to use ML (Available) Model to augment insights

1 what is the total purchase by gender

- Sum of Number_of_Purchases by Gender
- Sum of Days_Since_Last_Purchase by Gender
- Sum of PURCHASE_TOTAL

2 Q Purchase Total

- PURCHASE_TOTAL
- Number_of_Purchases
- Days_Since_Last_Purchase

3 from ADEDEMO.TRANSACTIONS
from ADEDEMO.CUSTOMERS
from ADEDEMO.CUSTOMERS

4 PURCHASE_TOTAL, DATE

Discoveries

— PURCHASE_TOTAL — Forecast

2,500,000
2,000,000
1,500,000
1,000,000
500,000
0

01-01 2016 03-06 2016 07-13 2017 05-19 2018 03-25 2019 07-01 2019

PURCHASE_TOTAL and DATE

- The chart now includes the forecast we calculated automatically.
- The smallest PURCHASE_TOTAL value was 323493, which occurred for January 2019.
- The largest value was 2409507, which occurred for May 2018.
- PURCHASE_TOTAL began on January 2016 with a value of 525362, ending on January 2019 at 323493.
- PURCHASE_TOTAL had a declining trend, changing -38% from January 2016 to January 2019.


Try out the new Db2 Augmented Data Explorer

<https://www.ibm.com/us-en/marketplace/db2-augmented-data-explorer>


IBM Augmented Data Discovery ⚙️

🔍 Discover data and insights ×

Steps to get started


 *Set up your database*

Tell us where to find the data that you want to discover. [Database Setup](#)

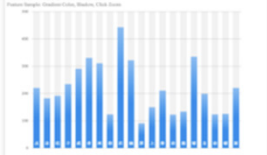
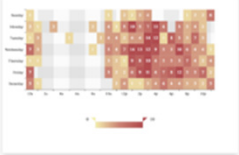

1010001101011
1010110101000
10100011
10100011  *Ask questions to your data*

Using the search bar, start asking questions containing column names or field values

- “ Average Sales by Region ”
- “ Average Profit for Northwest region where Sales > 10,000 ”
- “ Max Profit for all region and department ”

 *Discover and Refine*

Discover through Visualization and Natural language insights



Refine your search through the interactive pills in the search bar

How to be the best database for ML applications?

Access to more Data
=
**Better Machine Learning Application*



Let's try to bring Machine Learning closer to the data!

Testing and Deployment

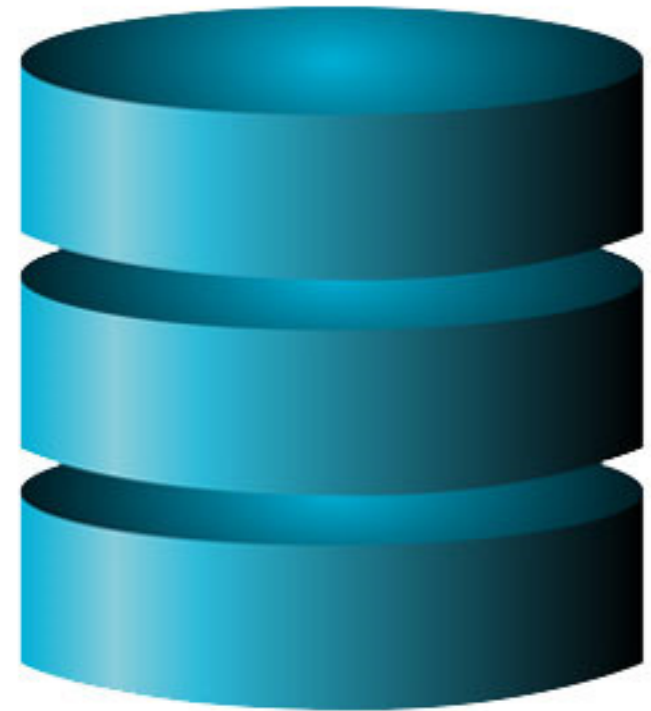
Model Development

Data Quality and Transformation

Data Exploration

Data Access

Data Storage



IBM Db2 - the AI database

“IBM is adding AI into Db2 through data science methods that will inject Machine Learning into it.”

-Forbes

“Db2 11.5 will feature better hooks for the languages and development tools that data scientists work with”

-Datanami

“Db2 version 11.5 has added drivers for leading AI languages to help ease AI app dev with Db2.”

-DBTA

One Db2 download, three Db2 licenses

	Community License	Standard License	Advanced License
Price	Free, unsupported	License + S&S (perpetual)	License + S&S (perpetual) subscription
License	Development / Production	Development / Production	Development / Production
Metric	Not applicable	Virtual Processor Core (VPC)	Virtual Processor Core (VPC)
Target	Developer community	Departmental	Enterprise
Caps	4 cores, 16 GB RAM	16 cores, 128 GB RAM	Unrestricted
Features	All features	All features	All features

Write once, run anywhere, with a Common SQL Engine

Anchored by a **Common SQL Engine** enabling true, highly scalable hybrid data warehousing solutions with portable analytics

Unified application and user experience across the Db2 family

Managed public Cloud DBaaS

Db2 on Cloud
Db2 Warehouse
on Cloud

Software defined warehouse On-Premises or in Cloud

Db2 Warehouse

Dedicated analytics appliance

Integrated
Analytics System

Custom deployable database

Db2
Database

Open source Hadoop with Hortonworks

Db2
Big SQL

- **Application compatibility**

Write once, run anywhere

- **Operational compatibility**

Reuse operational and housekeeping procedures

- **Licensing**

Single entitlement for flexible consumption enabling business agility and cost-optimization

- **Integration**

Data virtualization capabilities for query federation and data movement

- **Standardized analytics**

Common programming model for in-DB analytics

- **Ecosystem**

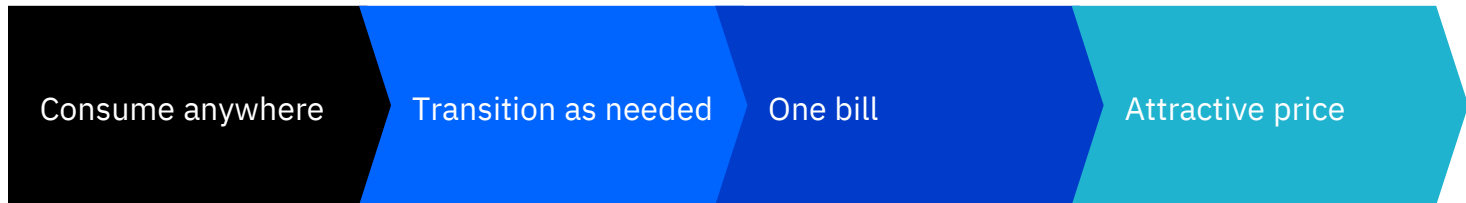
One ISV product certification for all platforms

Deploy and scale data, when and where you want

IBM Hybrid Data Management Platform

One Db2 family

One HYBRID entitlement



Available for purchase via two licensing models:

HDMP On-Premises

*(licensed **perpetually**)*

- ❑ Supports OLTP, warehousing, SQL on Hadoop, and Fast Data on private cloud and on-premises infrastructure
- ❑ Intended for customers who need to maintain ownership of their software licenses in perpetuity

HDMP Subscription

*(licensed **monthly**)*

- ❑ Supports OLTP, warehousing, SQL on Hadoop, and Fast Data on public cloud, private cloud, and on-premises infrastructure
- ❑ Intended for customers who want to procure software and cloud services as an operational expense

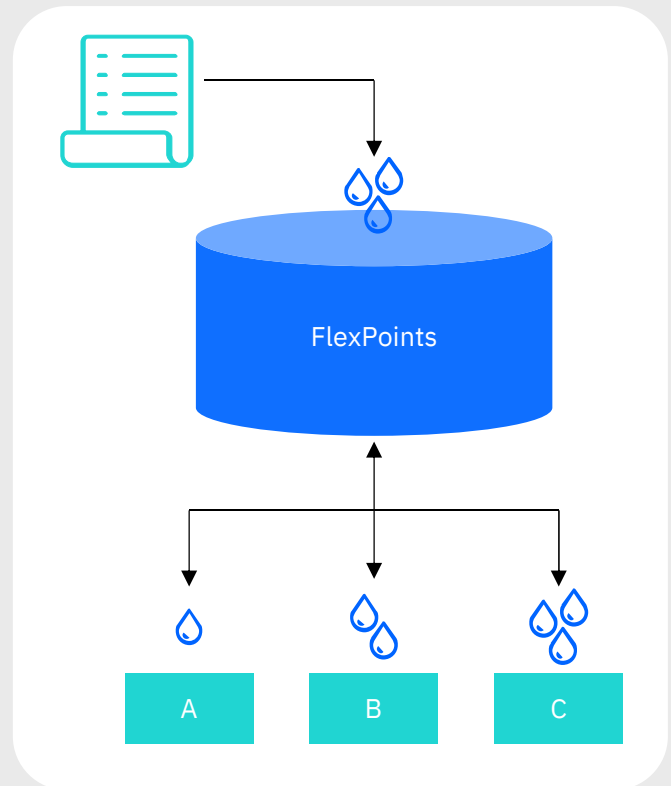
Hybrid Data Management Platform

Now supporting flexible, consumption-based licensing models

For both HDMP On-Premises and HDMP Subscription:

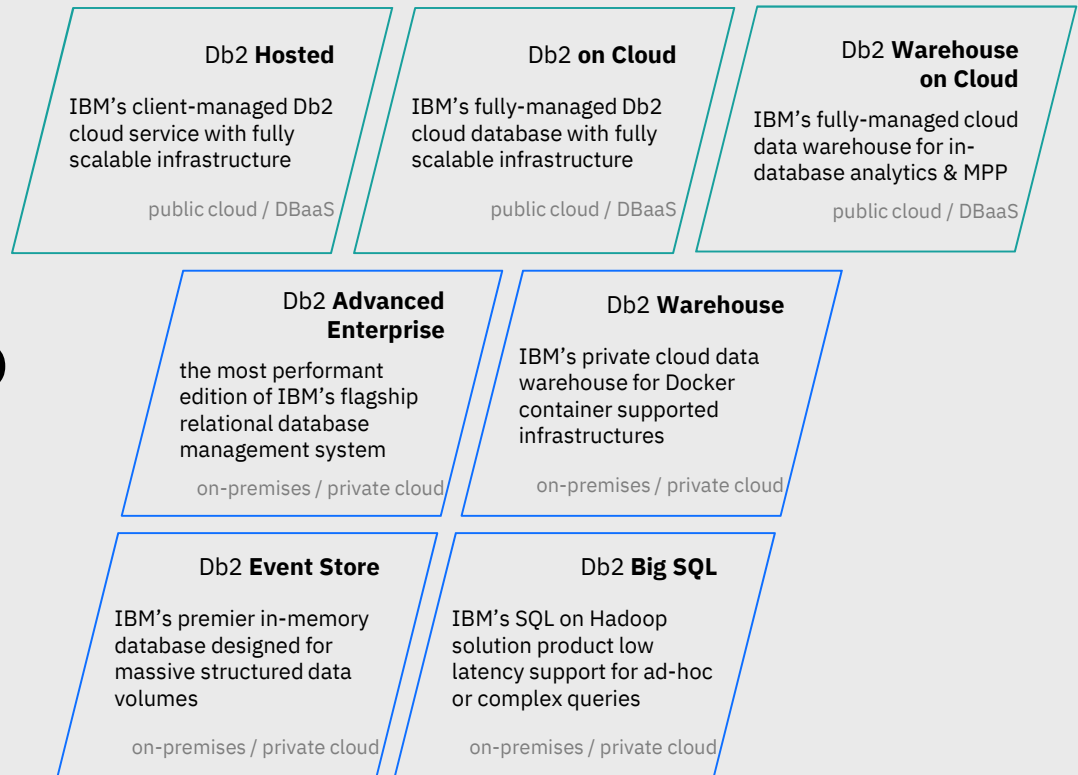
- ❓ **Step 1:** Customer purchases FlexPoints towards a license
- ❓ **Step 2:** Customer deploys those FlexPoints towards any of the bundled products and services included in the respective bundle
- ❓ Customer can mix/match their deployments and change them at any time
- ❓ No intervention is needed from IBM in order for customers to move between the bundled products and services!

[Estimate the costs - https://hdmp-calculator.mybluemix.net/](https://hdmp-calculator.mybluemix.net/)



What's included in the platform?

Customers purchase a monthly subscription and gain entitlement to seven of IBM's best-in-class data management products, spanning across on-premises and the cloud.

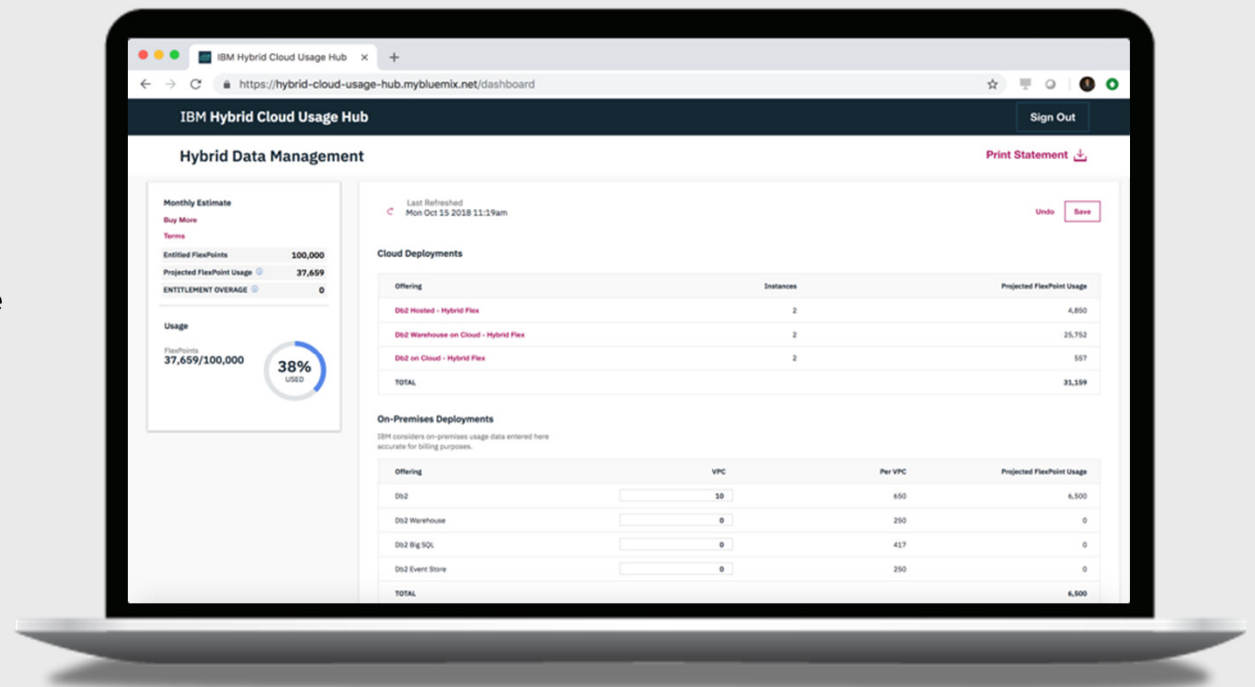


The Hybrid Data Management Platform provides your business with **peace of mind**. No matter how your business needs evolve in the future, deployments can be changed at any time by simply reallocating your **FlexPoints**.

Hybrid Cloud Usage Hub

HDMP Subscription customers gain access to the **Hybrid Cloud Usage Hub**, a web-application used to manage and deploy your FlexPoints across all of the bundled products.

- ❓ Seamlessly integrate across on-premises and cloud using one dashboard
- ❓ Gain a 360 degree view of your resources and how they're being consumed, so that you can optimize your infrastructure costs
- ❓ Easily and quickly expand your hybrid footprint, thereby reducing cost of development to production and saving valuable time



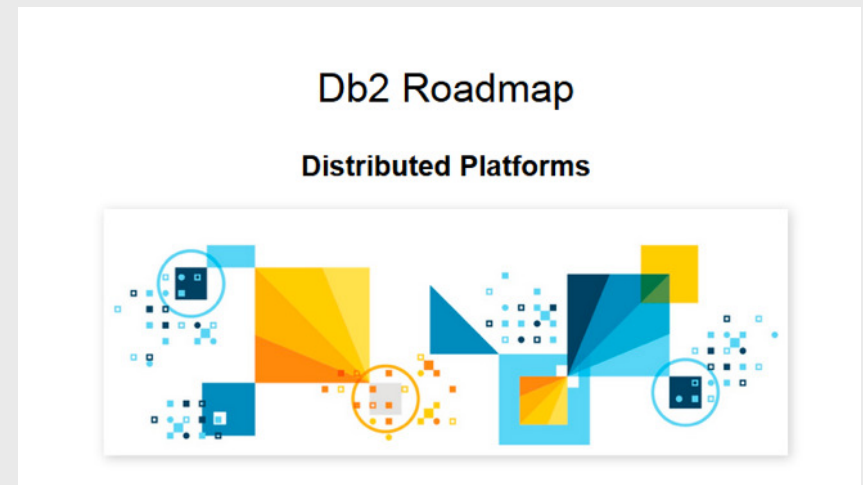
Follow our plans using Aha

We revisit development priorities frequently (e.g. every quarter) in response to customer and market demand/feedback

- As a result: some items move up, some down, some in, and some out.

We have committed to keeping our core roadmaps visible to the public eye using Aha

- <http://ibm.biz/AnalyticsRoadmaps>



Pandit Prasad

Offering Manager – Db2, Big Data

pprasad@us.ibm.com

310-648-1934

IBM Developer

Code patterns,
Tutorials, Videos &
Articles

<https://developer.ibm.com>

**Digital
Technical
Engagement
for Db2**

<https://ibm-dte.mybluemix.net/>

Read more about Db2 and application
development:

Thank you!

IBM Analytics/ Db2 for Developers/2018 / © 2018 IBM Corporation

