IBM

Cloud Pak for Data

The data and AI platform invented to achieve modern innovations faster.

Data and AI is primed for reinvention

In today's data and AI landscape, we have addressed compute capacities, maturity of point tools, availability of data sources, and automated AI functionality. What used to be limiting factors are now adding unnecessary complexity and stifling the ability to achieve modern outcomes with speed and agility. Instead of how individual tools advance, business today requires emphasis on how capabilities evolve together.

Data and AI space is primed to reinvent itself with a holistic platform approach. Opportunities for cost reduction, tools consolidation and streamlining processes are immense.

Interoperability removes redundant and inefficient spending by maximizing current and future investments

Speed to capability delivery empowers teams removing lag-time when something is needed today

Unification of experience removes hurdles, handoffs, and duplication of work -- emphasizing process efficiency and innovation velocity

Business outcomes and priorities have not changed. It is challenges in delivering those outcomes faster and more accurately which forces reinvention. With our Ladder to AI strategy, a connected data development lifecycle approach, IBM has reinvented how organizations will achieve faster and more targeted value from their data. Cloud Pak for Data is the data and AI platform that delivers all of the capabilities you need to attack any data challenge.

AI augmentation will create

$2.9 trillion

of business value in 2021*

*Gartner, INC. Aug 2019

AI Ladder
A holistic data strategy

Five critical steps you need on your Journey to AI.

Modernize
Make your data ready for an AI and hybrid multi-cloud world.

Collect
Make data simple and accessible.

Organize
Create a business-ready analytics foundation.

Analyze
Build and run AI with trust and transparency.

Infuse
Operationalize AI throughout the business.

What an end-to-end data lifecycle should look like

Collect

Increase business and analytics user's ability to consume all data in a self-service and governed way, decreasing time to insight while reducing cost and risk.

- Utilize all compute
- Visibility across the enterprise
- Leverage 100% of the data you need
- Connect IT and data governance

Organize

Decrease time to value in your enterprise through a DataOps approach to ensure that analytics and AI are built on high-quality assets, resulting in increased accuracy, faster time to deployment, and reduced risk throughout the organization.

- Trace back to data sources
- Make data governance agile
- Safeguard data with policy controls
- Know what data is available

Analyze

Increase deployment speed, accuracy, and trust of AI models through an integrated toolset that allows users of all skill levels to build and deploy AI.

- Collaboration is powerful
- A lifecycle fed directly into AI Apps
- Reliably sourced models

Infuse

Truly operationalize AI. Enable humans and machines to cooperate and complete ever more complex tasks at speed and scale.

- Automate business processes
- Employ AI-assisted intelligence
- Automate planning, budgeting, and forecasting analytics

Find out exactly where you're at on your Journey to AI.

ibm.com/data/modernize

Cost efficiency, production proficiency

Statistics from the 2020 Forrester Total Economic Impact study

ibm.biz/cloudpakfordata-tei

65-85% decrease

Containers And Container Management Efficiencies

Application management and efficiency will save 65% to 85% in infrastructure and associated management

25-65% reduction

Data Virtualization And Governance

Automation and faster access with governance improved productivity and reduces ETL requests by 25% to 65%

~$3M in savings

Data Science, Machine Learning and Artificial Intelligence Benefits

50% improvement in Data Scientist Productivity (recorded savings of $1.2 million to $3.4 million)
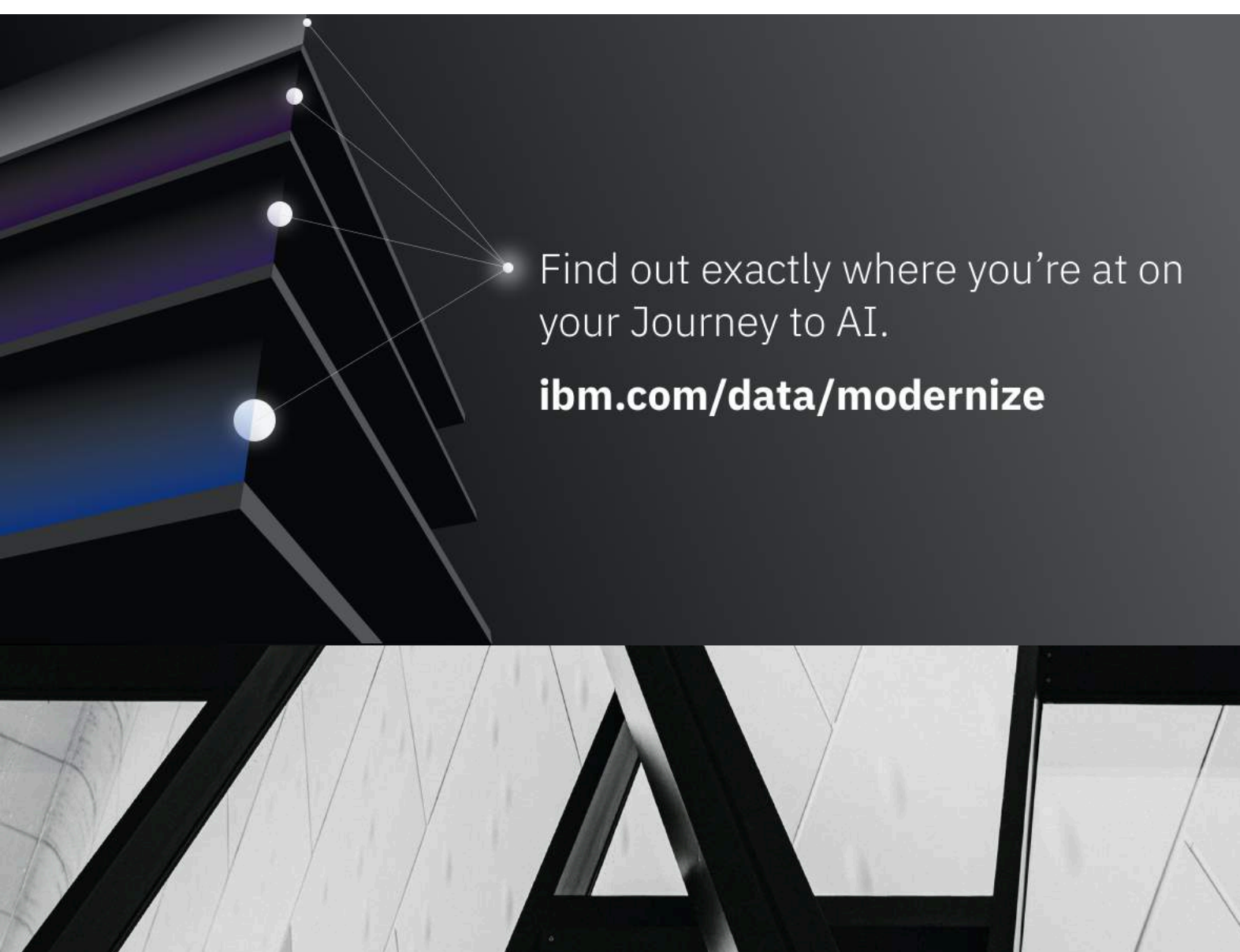
$7.3M

in benefits gain
(TEI study)

430%

faster than federation
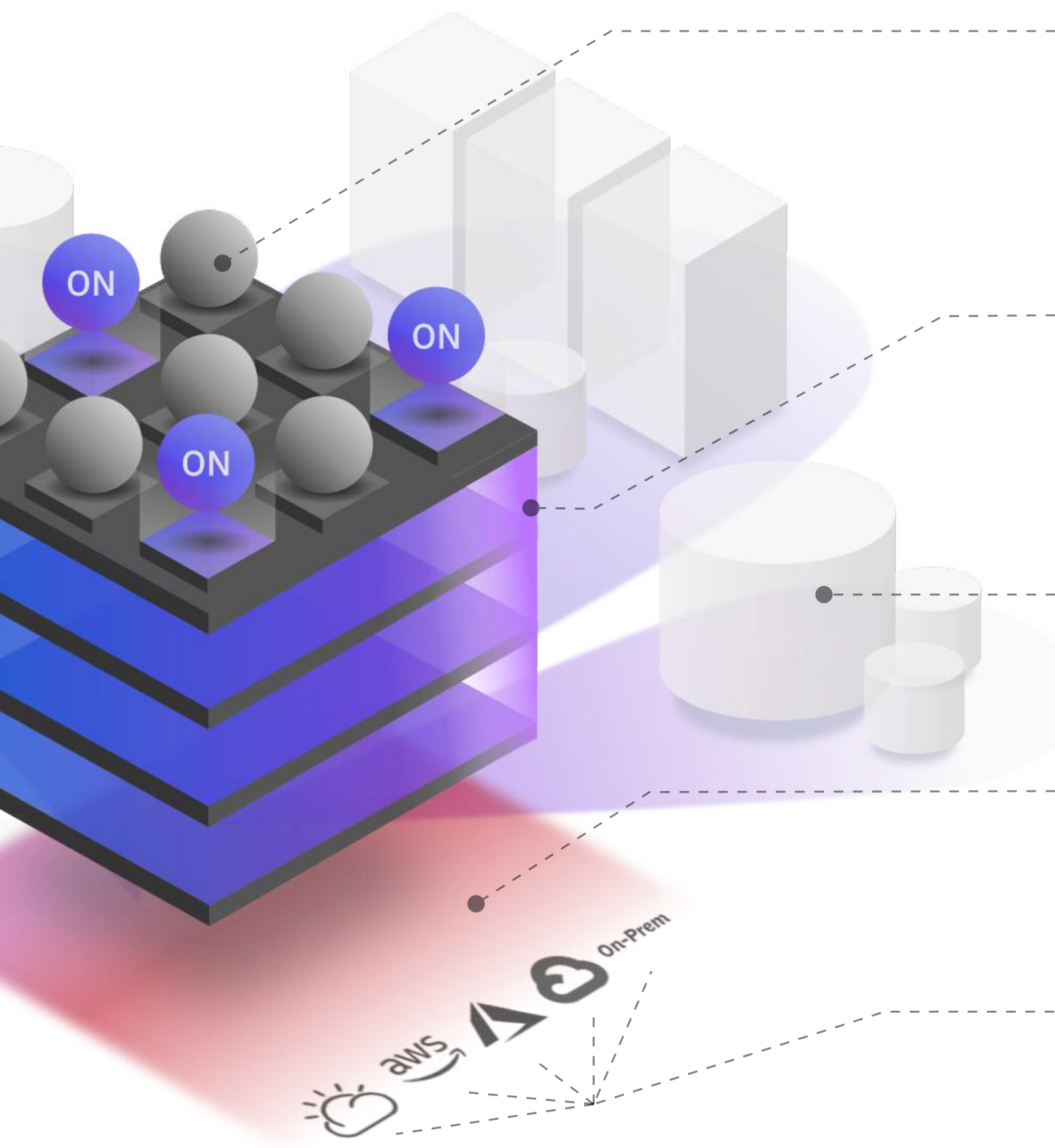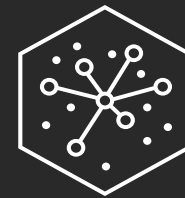(IBM testing)

65%

reduction in ETL jobs
(TEI study)

Cloud Pak for Data

A whole greater than the sum of its parts

Cloud Pak for Data automates the Ladder to AI. It drives a modern enterprise strategy by providing the connective fabric throughout the end-to-end data lifecycle for every tool and persona in the process. By incorporating 50+ IBM services, connecting to an even greater partner ecosystem of capabilities and leveraging an any-cloud environment model, it delivers a seamless user experience where the **whole is greater than the sum of its parts.**

Capabilities available, when needed
Services Ecosystem
With a click, search and access a growing ecosystem of data and AI services, IBM industry accelerators, and partner offerings to expand upon Cloud Pak for Data's end-to-end capabilities.

Unified User Experience
Collaborative Data Lifecycle
Speed time-to-value with a single user experience that integrates data management, data governance and analysis for greater efficiency and improved use of resources.

Existing infrastructure Utilization
Data Virtualization
Quickly and easily query across multiple data sources without moving your data.

Kubernetes implementation
Red Hat OpenShift
Leverage the leading hybrid cloud, enterprise container platform for an innovative and fast deployment strategy.

Hybrid Multi-Cloud world
Any Cloud
Avoid lock-in and leverage all cloud infrastructures with our multi-cloud approach.

01 Foundational Services

A full data lifecycle of services integrated in one platform

Collect

Db2 Eventstore

Store and analyze over 250 billion events per day with a memory-optimized database for real-time analytics, event-driven apps and AI with machine learning.

Data Virtualization

Makes queries across multiple data sources fast and easy without moving your data.

Db2 Warehouse

A client-managed, highly flexible operational data warehouse for private clouds and containerized deployments.

BigSQL

An enterprise-grade, hybrid ANSI-compliant SQL-on-Hadoop engine, delivering massively parallel processing (MPP) and advanced data query.

Organize

Master Data Connect

Comprehensive method to define and manage an organization's critical data. It provides a single, trusted view of data across the enterprise, agile self-service access, analytical graph-based exploration, governance and a user-friendly dashboard.

Watson Knowledge Catalog
Includes Data Refinery
Catalog, understand, govern, analyze and deliver business-ready data to your data citizens.

Information Governance Catalog
Included in WKC
IBM InfoSphere Information Governance Catalog (IGC) is a web-based tool that allows you to explore, understand and analyze information.

Open Source Management

Open source software management; OSS.

Instascan

Uses native index to give you visibility into your cloud data. Perform Risk Assessment and Compliance Check with speed and confidence by taking advantage of industry-recognized sampling algorithms.

InfoSphere Information Analyzer
Included in WKC
Provides data profiling and analysis to accurately evaluate the content and structure of your data for consistency and quality.

Regulatory Accelerator
Included in WKC
Streamline the process of complying with regulations by harnessing the power of machine learning.

Analyze

Watson OpenScale

Infuse your AI with trust and transparency. Understand how your AI models make decisions to detect and mitigate bias.

Watson Studio

Simplify and scale data science to predict and optimize your business outcomes.

Watson Machine Learning
Includes Auto AI
Deploy machine-learning models into production at scale.

Guardium **(integration component)**

Integration with external security information and event management (SIEM) systems

Analytics Engine for Apache Spark

A powerful unified analytics engine for large-scale distributed data processing and machine learning.

IBM Streams

Leverage continuously available data from all sources to discover opportunities faster.

Cognos Dashboard

An API-based solution that lets developers easily add end-to-end data visualization capabilities to their applications.

Administration Simplified

Empower

Grant administrative and editorial privileges to the right people and put analytics capabilities at their fingertips.

Visualize

From quickly identifying data governance rules assigned to datasets to tracking compute consumption of your various capabilities; a single vantage point for your entire data estate.

Collaborate

Accelerate time to team insights. Centralize workflows in a single interoperable platform where multiple roles and teams can operate as one.

Model Development

Build

At this phase in the AI lifecycle, it is critical to ensure companies use the right algorithms to build their models for making predictions.

Run

Put custom models into production, in an application or business process. Once a model is built and running, the question becomes: how can it be scaled with trust and transparency? This is where management plays a crucial role.

Manage

By having the management in place, organizations can track who changed the model, when the model was deployed, and the lineage on the model. By tracking all these items, organizations can ensure their models are not biased, and that they're explainable and transparent.

Extended Services

A growing ecosystem of additional Capabilities to utilize

Cognos Analytics

Cognos Analytics is an AI-fueled business intelligence platform that supports the entire analytics cycle, from discovery to operationalization. Visualize, analyze and share actionable insights about your data with anyone in your organization.

Db2

IBM Db2® is a family of hybrid data management products offering a complete suite of AI-empowered capabilities designed to help you manage both structured and unstructured data on premises as well as in private and public cloud environments. Db2 is built on an intelligent common SQL engine designed for scalability and flexibility.

Decision Optimization

IBM Decision Optimization represents a family of optimization software that delivers prescriptive analytics capabilities to enable organizations to make better decisions and deliver improved business outcomes.

Financial Crimes Insights (FCI)

Simplify the process of detecting and mitigating financial crimes with AI and regulatory expertise.

Financial Services Workbench

Design, implement, and deliver microservices that allow you to build AI banking apps faster.

Information Server

IBM InfoSphere® Information Server is a market-leading data integration platform which includes a family of products that enable you to understand, cleanse, monitor, transform and deliver data, and to collaborate to bridge the gap between business and IT. InfoSphere Information Server provides massively parallel processing (MPP) capabilities to deliver a highly scalable and flexible integration platform that handles all data volumes, big and small.

InfoSphere DataStage

IBM® InfoSphere® DataStage® provides a graphical framework for developing the jobs that move data from source systems to target systems. The transformed data can be delivered to data warehouses, data marts, and operational data stores, real-time web services and messaging systems, and other enterprise applications.
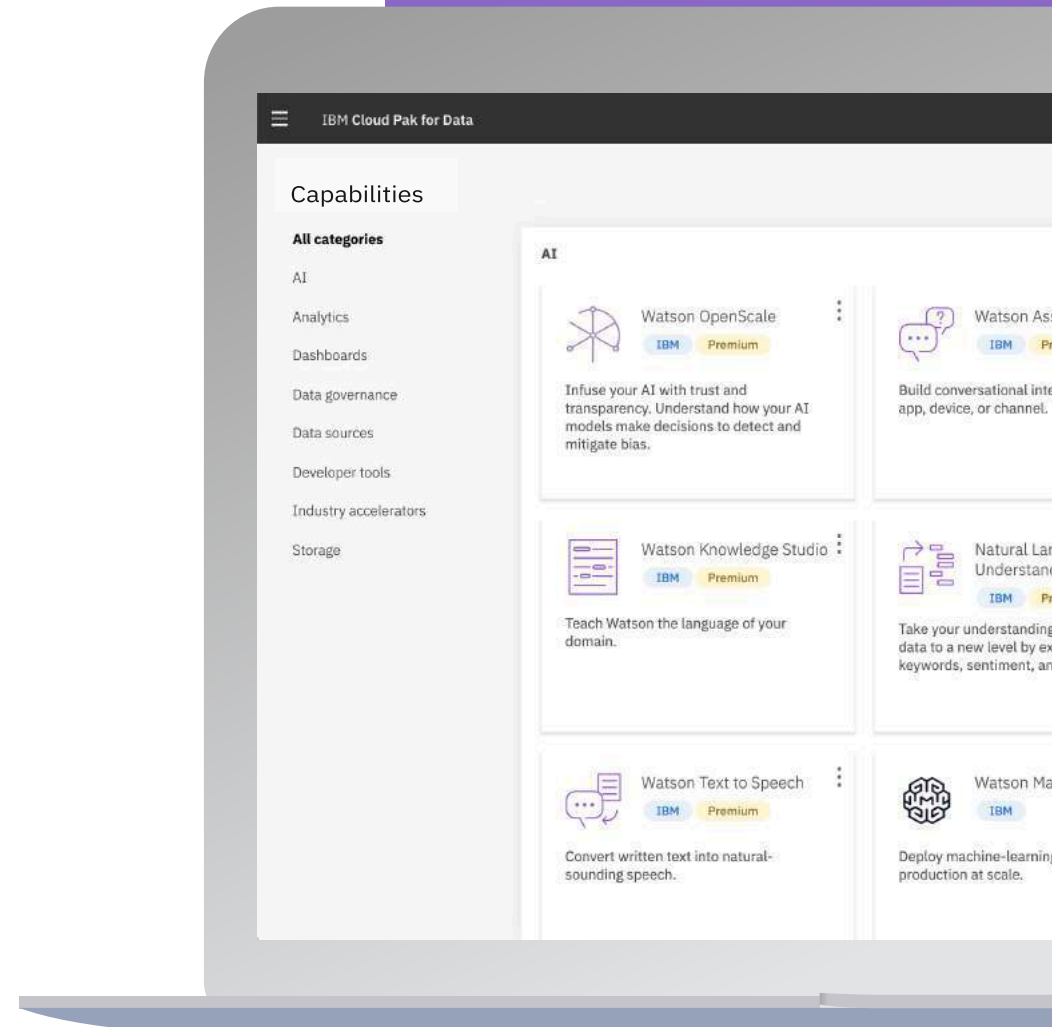
Master Data Management (MDM)

Master data management (MDM) is a comprehensive method to define and manage an organization's critical data. It provides a single, trusted view of data across the enterprise, agile self-service access, analytical graph-based exploration, governance and a user-friendly dashboard.

Planning Analytics

IBM Planning Analytics is an integrated planning solution that uses AI to automate planning, budgeting, and forecasting and drive more intelligent workflows.
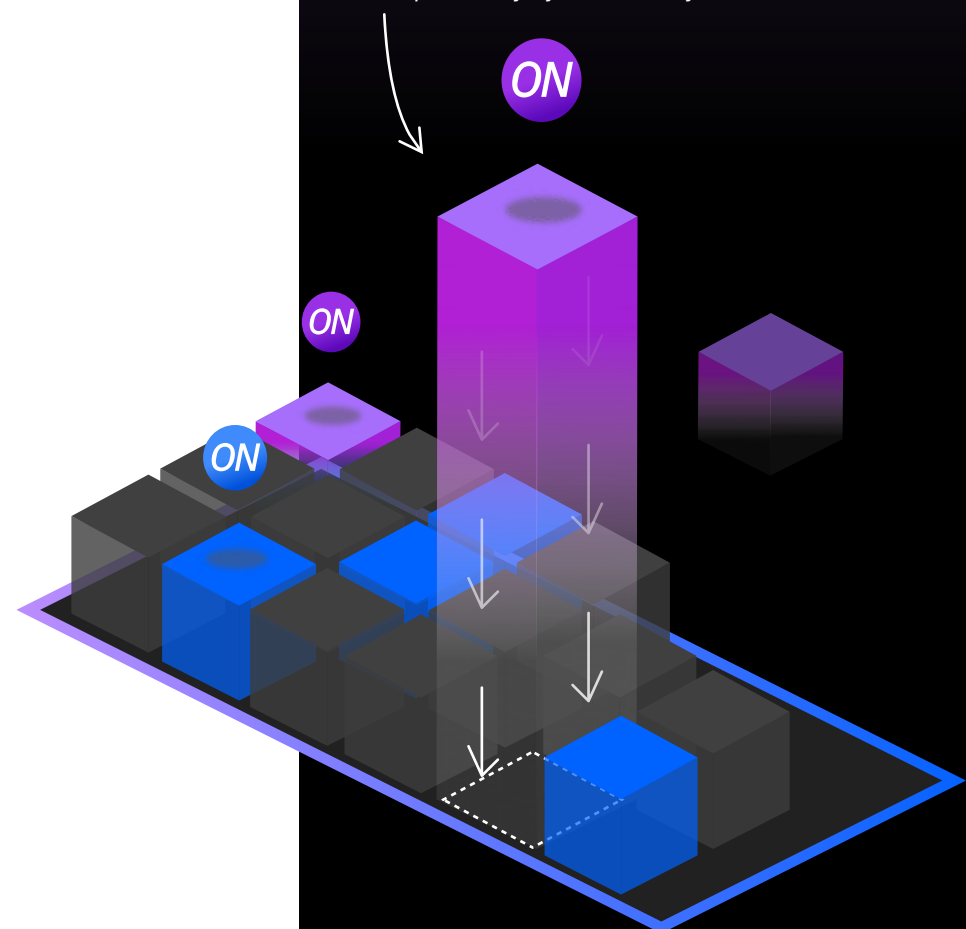
SPSS Modeler

Drive ROI and accelerate time to value with an intuitive, drag-and-drop data science tool

A whole greater than the sum of its parts.

Connect more capabilities with the interoperability of Cloud Pak for Data

ON
ON
ON

Virtual Data Pipeline (Actifio)

Turns data silos into data pipelines, to supercharge Data Driven Enterprises.

Watson Studio Premium

Watson Studio Premium helps accelerate time to value with your AI investments. It includes IBM Decision Optimization, SPSS Modeler and Hadoop Execution Engine. This powerful combination helps predict and optimize business outcomes in a single unified platform. You can empower citizen data scientists to use a drag and drop interface for data preparation and model creation. Also, by bringing analytics to where your data resides, it obviates the need to move data and recapitalize on your Hadoop investments.
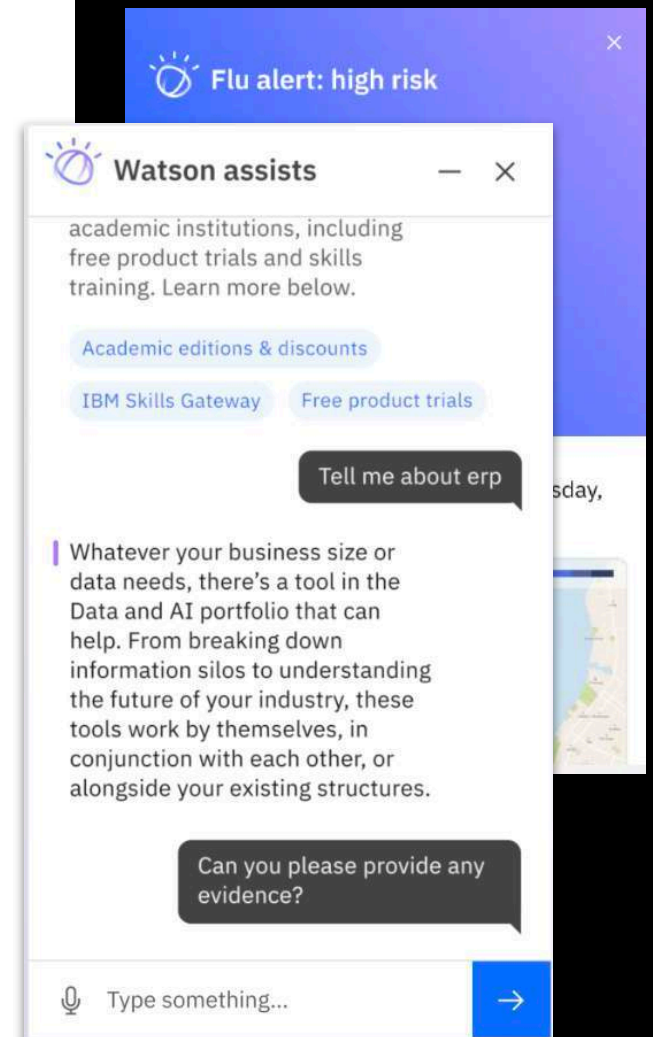
Watson API Kit

Watson™ API Kit provides the capability to combine and choose several Watson APIs within one offering, with the goal of building AI applications or enhancing other applications.

Watson Assistant

Use IBM Watson™ Assistant to build your own branded assistant into any device, application, or channel. Your assistant connects to the customer engagement resources you already use to deliver an engaging, unified problem-solving experience to your customers.

Watson Discovery

IBM Watson™ Discovery is an award-winning AI search technology that eliminates data silos and retrieves information buried inside enterprise data. Watson Discovery applies the latest breakthroughs in machine learning, including natural language processing capabilities, and is easily trained to your domain. With the flexibility of the IBM Cloud, users can operate anywhere

03   Two Editions

Standard Edition

Standard Edition is a great way for an enterprise new Cloud Pak for Data to get started on some use cases. All functionality is available; however, the consumption restriction is up to a maximum of 64 VPCs.
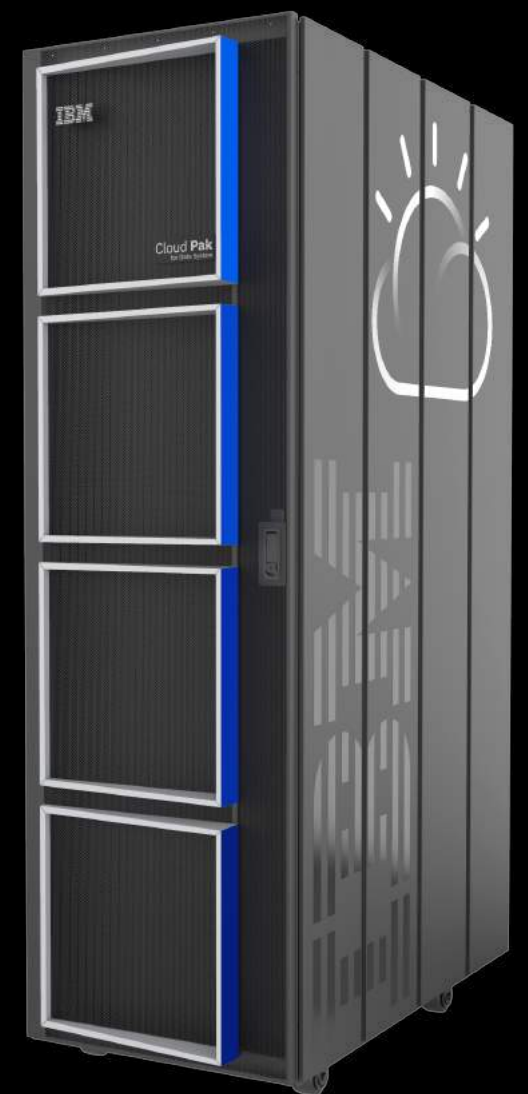
Enterprise Edition

Enterprise Edition will grow from beginning use cases all the way throughout your enterprise. Utilize capabilities requiring higher compute capacity, leverage the platform behind your firewall with the System, and take your enterprise's data strategy to the next level.

04   Cloud Pak for Data System

Cloud Pak for Data System

A hyper-converged system combining storage, compute, networking and software. With Cloud Pak for Data pre-installed, this modular system gives you the ability to stand-up the end-to-end data platform behind your firewall.

Cloud Pak for Data System