

# The data and AI platform buyer's guide

Practical questions every business should ask when choosing an enterprise-grade data and AI platform



**IBM**

## Contents

1. Consolidating and deploying data resources
2. Accessing, governing and securing data
3. Optimizing business processes with AI
4. Choosing a provider

## Introduction

More and more businesses are attempting to consolidate all their data and AI initiatives onto a single platform. However, not all platforms are created equal. Lately the term has been used very loosely by many providers who claim to deliver the whole spectrum of data and AI capabilities—but often with mixed results. So, what defines a true enterprise-grade platform for data and AI?

Leaders evaluating the myriad of options available need to ensure that optimizing for today doesn't prevent them from flexibly meeting their needs tomorrow. With the challenge and IT complexity of disparate solutions, it's imperative that a true platform provides interoperability, scalable provisioning services, and a shared unified experience to complete any data and AI task from data management and data governance to analytics and AI.

In this guide you'll find the top questions you need to consider when looking for the right data and AI platform for your organization. This is an important decision; the right platform can help you improve business outcomes, drive innovation and digital transformation, and achieve successful ROI.



## 1. Consolidating and deploying data resources

### Can you easily provision and scale all the data and AI services you need on a modular architecture—and can you deploy them on any cloud or on-premises environment?

The platform you invest in needs to support services that span the entire analytics lifecycle, including data management, DataOps, governance, analytics and AI. Look for a platform that runs on cloud-native architecture, such as Red Hat® OpenShift® Container Platform. With a cloud-native, modular architecture you have the open source capabilities required to run on any private or public cloud, including IBM Cloud®, AWS, Microsoft Azure or Google Cloud Platform. This eliminates vendor-lock in and enables you to select and define the key data and AI services you want in your platform experience.

With cloud-native architecture, your services inherit a number of benefits, including auto provisioning and scaling, seamless upgrades, built-in high availability, common logging, metering, monitoring, and more.

### Does the platform have capabilities to bring together all data regardless of type, structure or source?

A true data and AI platform should eliminate data silos and allow you to work with data without having to move it, no matter that data's type, structure or source. When choosing your data and AI platform, look for one that has the capability to make queries across multiple data sources without having to copy and replicate data. This query capability helps to reduce costs and can simplify your analytics, making it more up-to-date and accurate because you're accessing the latest data at its source.

In particular, a platform that can bring together all of your data should include integrated solutions for databases, data warehouses and data lakes. Its databases should employ high-performance and scalable transactional processing with query optimization. Its data warehouses should be able to perform analytics across on-premises environments. And its data lakes should be able to help you store and query structured and unstructured data no matter the data volume.



## 2. Accessing, governing and securing data

### Are all of the data and AI services interoperable, and do they enable self-service across your organization?

Once you narrow down a platform that is open and has all the data and AI services you need, it's important to evaluate if these solutions are easily interoperable, and if they're available in a self-service fashion for every role in your data team. You don't want a platform that requires extra time and resources to integrate services together. Instead, you want services that are intrinsically connected and work together easily. The right data and AI platform must enable interoperability of all analytics services, while also providing a collaborative self-service dashboard view to help your teams unite and get the job done.

### Does the platform have organizational capabilities to help you govern data, including a shared metadata catalog across all services?

Look for a solution that has cataloging capabilities to help you easily find and share your data and other assets. In order for your teams to receive a holistic and complete view of data, information from databases, data warehouses, data lakes or files, as well as tools for data modeling, business intelligence and applications all needs to be imported into a metadata catalog such as [IBM Watson® Knowledge Catalog](#). This is the most effective way to help organize your data and resources for multiple uses including data science projects, and to implement universal governance policies to help ensure data is secure.

### Are you able to automate governance policies to secure data of all sources and structures so you can comply with regulations as well as reduce manual work?

When looking for a data and AI platform best suited for your organization, the ability to govern your data should be a major consideration. That's because in order to work with data safely, you have to ensure it is governed appropriately. You should have the tools to quickly and efficiently respond to changing regulations with embedded and automated governance services.

The key here is automated data governance capabilities. Look for a platform that has data catalogs to help your organization find, curate, analyze, prepare and share data while protecting against misuse. You want to have active policy management and a dynamic masking of sensitive data to help ensure compliance, maintain audit-readiness and preserve client trust. As mentioned earlier in this guide, you also want to ensure these capabilities are self-service for your teams, so you can improve efficiency.

### 3. Optimizing business processes with AI

#### Does the platform have tools to help your team members, regardless of their data science skill level, automate the entire AI lifecycle?

No matter where your organization is in its journey to AI, you must ensure you have the skills required to build and train models. When researching the ideal data platform, look for one that can help more people participate in model building, and help those who participate be more productive. With increased engagement from your team, you can more easily build, run and manage AI models at scale.

A platform that provides automated AI lifecycle management and optimization technology can improve productivity and expand your AI-capable talent pool, as well as drive better business planning and forecasting.

Regardless of the skills your data scientists and individual team members may have, services like [IBM Watson Studio](#) can help unite your teams to work comfortably with AI, ultimately improving efficiency and accelerating time to value. Through a single environment view, your teams can collaborate to optimize model performance and simplify the AI lifecycle management with data preparation, model development, engineering and hyper-parameter optimization.

#### Does the platform have the ability to accurately explain AI model outcomes and safeguard fairness by minimizing model bias and drift?

The ability to deploy AI with continuous model evaluation is what keeps leading-edge organizations ahead. Beyond helping with AI model creation and management, the platform you choose needs to empower your teams to make trusted decisions.

Choose a platform that has services such as [Explainable AI](#) to help you continuously monitor and evaluate the status, fairness, quality and drift of your AI models. Your teams should be able to quickly assess and mitigate AI risks, and furthermore have the capabilities to explain any potential biases. It's crucial that your organization not only operationalizes AI but uses it fairly and responsibly when putting models into operation.

#### Can the platform help improve business outcomes through pre-built AI models?

There will likely be instances when you want to craft custom AI models to address specific outcomes. However, if your platform has pre-defined AI workflows and industry solutions built in, you can automate complex processes and optimize your employees' time instead of building each model from the ground up.

Look for a platform that provides pre-trained AI models that help you quickly tap into organizational data and improve efficiencies in areas such as finance, business operations, customer care and supply chains. The key is to find a solution that has the breadth of industry capabilities you need to address your specific goals. With pre-defined AI workflows, such as the ones provided through [IBM Watson](#), you can infuse AI into your projects to help shape and predict future outcomes. You can also continue to scale the expertise of your data team and make smarter decisions based on deeper insights.

## 4. Choosing a provider

### The IBM solution: IBM Cloud Pak for Data

[IBM Cloud Pak® for Data](#) is a comprehensive cloud-native data and AI platform that helps modernize data management, DataOps, governance, analytics and AI so you can drive outcomes for your business faster. It's the ideal analytics platform for enterprises.

Built on Red Hat OpenShift foundations, IBM Cloud Pak for Data delivers an open information architecture with integrated capabilities from IBM and IBM partners, including open source, to help you achieve your AI aspirations while enabling governance and data protection.

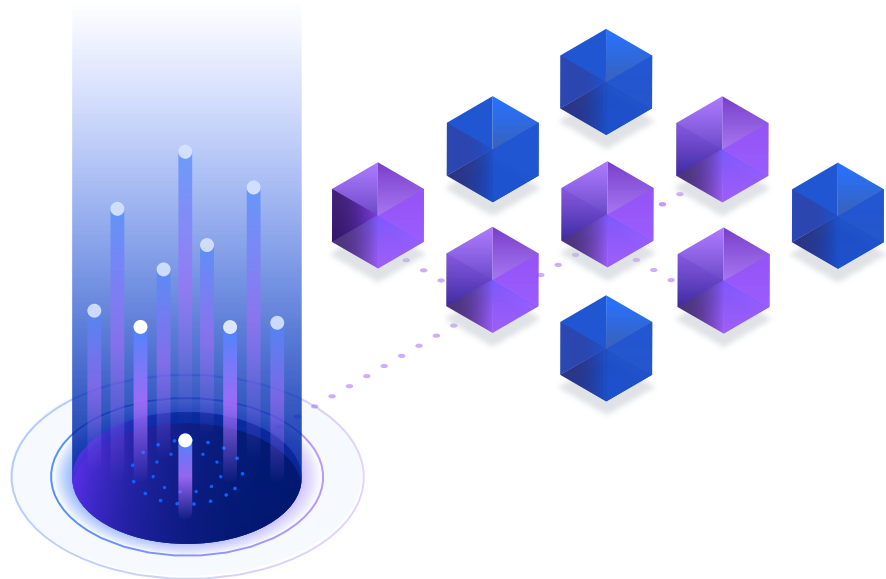
Flexible deployment and consumption models for the platform include:

- [IBM Cloud Pak for Data](#): A client-managed software platform that runs on any cloud
- [IBM Cloud Pak for Data System](#): A preconfigured, hyper-converged system that combines storage, compute, networking and software
- [IBM Cloud Pak for Data as a Service](#): A “pay-as-you go” subscription model of IBM Cloud Pak for Data services, fully managed on the IBM Cloud

### Next steps

Learn more about the value of a unified data and AI platform and its associated ROI by taking a look at Forrester’s [Projected Total Economic Impact of IBM Cloud Pak for Data](#) report.

To discuss any questions you may have about the qualities of a true data and AI platform and how your organization may be able to benefit from one, [schedule a free consultation](#) with an IBM Data and AI expert.





© Copyright IBM Corporation 2020

IBM Corporation  
New Orchard Road  
Armonk, NY 10504

Produced in the United States of America  
December 2020

IBM, the IBM logo, IBM Cloud, IBM Watson, and IBM Cloud Pak are trademarks or registered trademarks of International Business Machines Corporation, in the United States and/or other countries. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on [ibm.com/trademark](http://ibm.com/trademark).

Red Hat® and OpenShift® are registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

It is the user's responsibility to evaluate and verify the operation of any other products or programs with IBM products and programs. THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

Statement of Good Security Practices: IT system security involves protecting systems and information through prevention, detection and response to improper access from within and outside your enterprise. Improper access can result in information being altered, destroyed, misappropriated or misused or can result in damage to or misuse of your systems, including for use in attacks on others. No IT system or product should be considered completely secure and no single product, service or security measure can be completely effective in preventing improper use or access. IBM systems, products and services are designed to be part of a lawful, comprehensive security approach, which will necessarily involve additional operational procedures, and may require other systems, products or services to be most effective. IBM DOES NOT WARRANT THAT ANY SYSTEMS, PRODUCTS OR SERVICES ARE IMMUNE FROM, OR WILL MAKE YOUR ENTERPRISE IMMUNE FROM, THE MALICIOUS OR ILLEGAL CONDUCT OF ANY PARTY.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.

AL73DZ1K