Seven benefits of using open source data science

Accelerate and scale
AI innovation with
IBM Watson Studio and
Anaconda Repository for
IBM Cloud Pak for Data

Industry leaders are building and scaling AI

Build and scale AI with trust and transparency

Forbes reports that 75% of enterprises will use AI-based software to discover operational and experiential insights to guide innovation.[1]

AI and open source AI are at an inflection point. And as the demand grows for AI models and apps, so too does the need to scale and achieve ROI sooner. Industry leaders are accelerating and scaling AI implementations by shortening time, effort, costs and decisioning.

Savings can range from low six-figures to high seven-figures or more. Projects are complete in hours or days, not months. With AI and decision optimization from IBM Watson® and open source innovation from Anaconda, this could be you.

Imagine using a platform that unifies data and AI projects across multiple tools and clouds. It automates AI development, includes governed open source data libraries, and virtualizes data to minimize the need to move it. The result is that your organization can gain seven important benefits. Read on for a closer look.

Seven benefits of using open source data science

01
Go live fast

02
Simplify decisions

03
Expand a talent pool

04
Save costs

05
Avoid lock-in

06
Synchronize apps and AI

07
Show results

01

Go live fast

It’s critical in these times that businesses are agile, and waiting weeks for a model means it’s likely to be out-of-date by the time you get it. This affects your AI development and outcomes—and your bottom line.

Slash waiting time by using AI to build AI. This is called AutoAI, a feature of IBM Watson Studio in Cloud Pak® for Data. Point to the column you need to predict in your dataset, and AutoAI automatically builds, tests, and ranks optimum machine learning models. This can cut months and weeks down to days or even minutes. And AutoAI can generate Python code, enabling you to further refine your model. You can also tap curated, governed open source data packages in the new Anaconda Repository for IBM Cloud Pak for Data to save additional time.

→ Watch AutoAI in action in an on-demand webinar featuring Wunderman Thompson Data

→ Speed up open source data science innovation with Anaconda

02

Simplify decisions

Leading organizations are turning artificial intelligence into decision intelligence. They are using AI to predict, and a different tool, decision optimization, to prescribe what to do with those predictive insights. With prediction and optimization in a data and AI platform, you can accelerate time to value and minimize errors when planning inventory, allocating resources, or scheduling talent.

For instance, a paper manufacturer used the ability to predict demand and prioritize orders to save USD 50 million in excess inventory and waste. The ability to predict *and* optimize is included in IBM Watson Studio. Adding it to a multicloud AI platform with access to open source data science has multiple benefits. You can build predictions with your Python notebooks and optimize decisions from those predictions by using a natural language interface in the cloud of your choice.

→ Learn how and why to combine prediction and optimization in this white paper

→ Read how analyst firm ESG validated IBM Watson Studio's ability to predict and optimize

03

Expand a talent pool

The data science talent shortage is well-known. What's new is that automation tools have made it easier for anyone to take part in AI development. GUI-based data science tools and AutoAI help analysts and subject matter experts turn their ideas into models and prepare data much faster. Beginners can be onboarded quickly.

These tools, plus access to open source packages, also help developers become data science powerhouses. This has become critical. Gartner predicts that application developers and software engineers will be the main implementers of AI in two to five years.[2]

→ See how AutoAI works and learn 10 ways it speeds value in this ebook

→ Discover how IBM Watson Studio Desktop can save data scientists significant data prep time

04

Save costs

Data science and AI don’t come cheap. You have to implement solutions, maintain data, move data, and invest in hardware. The cost of training data scientists and delivering AI and data science models can be staggering.

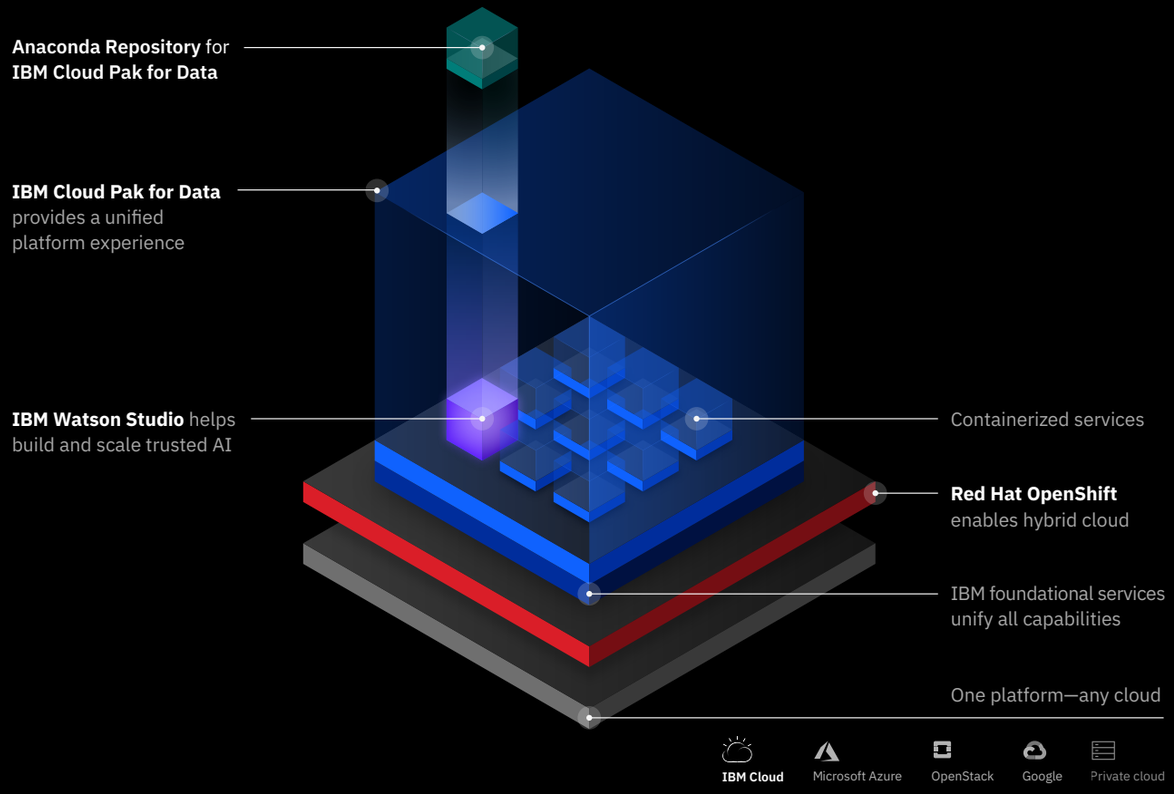
IBM Cloud Pak for Data helps you minimize these costs and others. Based on a foundation of Red Hat® OpenShift®, it enables you to reduce operational costs for data access by 70%, achieve as much as USD 14.4 million in container management efficiencies, and experience other cost savings of up to USD 1.2–3.4 million.[3] Having a governed, curated library of open source packages, without having to build from scratch, reduces costs as well. And being able to apply decision optimization to your predictive insights can result in additional savings.

→ Save on IT costs

→ Read Forrester’s Projected Total Economic Impact™ of IBM Cloud Pak for Data

IBM Cloud Pak for Data

You have everything you need to build and scale AI



05

Avoid lock-in

You need the flexibility to host your data and models on multiple clouds or behind your firewall as needed. So you shouldn't be locked in to a single cloud vendor. That introduces unnecessary risk. IBM Cloud Pak for Data enables you to choose your own hybrid multicloud approach.

IBM Cloud Pak for Data also enables you to use data and AI services from a broader ecosystem, including those from open source communities. Instead of relying on specific tools or skills, you can balance investments that are proportional to your business need.

You can also increase open source adoption securely with Anaconda Repository for IBM Cloud Pak for Data, which centrally stores open source data packages favored by your team.

→ Read the paper on open source and AI

→ Get the solution brief on Anaconda Repository for IBM Cloud Pak for Data

06

Synchronize apps and AI

Syncing app development with model development can ensure that models are successfully deployed into modern applications and that the results are fed back into retraining and app updates. This is called ModelOps.

Using models to drive transformation will separate leaders from laggards, observes 451 Research.[4] One model-driven competitive edge is to use prescriptive insights from decision optimization teamed with predictive insights from AI models. Combining the two is a capability of IBM Watson Studio. Another edge is to use the curated, open source tools and libraries in Anaconda Repository to streamline AI tasks.

Models can be pushed from a data science team to a DevOps team in a regular deployment and update cycle, aligned with continuous integration and continuous delivery (CICD) to suit business needs.

→ Read the 451 Research white paper

→ Watch the Multicloud ModelOps webinar with Forrester

07

Show results

Industry pioneers are demonstrating the value of being able to automate, predict and optimize while combining open source and proprietary tools. These capabilities can enable you to optimize workforce scheduling, build supply and demand plans using what-if analysis against millions of variables, and allocate resources when demand spikes. With the right data platform, you can meet challenges like these more easily than ever.

Furthermore, by extending IBM Cloud Pak for Data with IBM Watson Studio, you can use out-of-the-box industry accelerators and decision optimization to turn your ideas into results at the speed of business. In addition, you can decrease the TCO of analytics by being able to integrate data and AI projects across multiple clouds, and get greater benefit from investments in open source tools.

→ Read Forrester's Total Economic Impact™ of IBM Cloud Pak for Data

→ Benefits of modernizing your predictive and prescriptive analytics

More information

Power up your open source data science on a multicloud data and AI platform

→ **Check out IBM Cloud Pak for Data**

→ **Learn about Anaconda Repository for IBM Cloud Pak for Data**

→ **Explore IBM Watson Studio**

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
July 2021

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.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. 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.

01 Gill Press, "Top Artificial Intelligence (AI) Predictions For 2020 From IDC and Forrester," Forbes, November 22, 2020

02 Hype Cycle for Artificial Intelligence, 2019, Published 25 July 2019

03 Infographic, Save on IT costs with IBM Cloud Pak for Data, ibm.com

04 451 Research, AI and ModelOps with Intelligent Automation