

The top 10 big data frameworks used in the enterprise

Big data and machine learning programs are expanding across organizations, but finding tech talent to helm these initiatives is a challenge, according to a Qubole report.

By Alison DeNisco Rayome | September 11, 2018, 6:00 AM PST



Big data continues its rapid growth in the enterprise across all industries, according to a Tuesday report from Qubole and Dimensional Research. Organizations are using big data to drive IT projects, improve sales, and enhance customer service. And they're increasingly tapping big data frameworks to address the challenge of reaping the full value of the information gleaned in their business.

No single software framework dominates the big data landscape, the report found after surveying 401 data professionals with big data responsibility in large enterprises.

More about Big Data

- [GDPR: A cheat sheet](https://www.techrepublic.com/article/eu-general-data-protection-regulation-gdpr-the-smart-persons-guide/)
(<https://www.techrepublic.com/article/eu-general-data-protection-regulation-gdpr-the-smart-persons-guide/>)
- [Amazon Neptune is here: 6 ways customers use the AWS graph database](https://www.techrepublic.com/article/neptune-is-here-6-ways-customers-use-the-aws-graph-database/)
(<https://www.techrepublic.com/article/neptune-is-here-6-ways-customers-use-the-aws-graph-database/>)
- [4 tips for making data cleanup easier and more](#)

However, 25% of organizations are using homegrown approaches for big data processing.

SEE: [The Power of IoT and Big Data](http://www.techproresearch.com/downloads/the-power-of-big-data-and-iot/)

(<http://www.techproresearch.com/downloads/the-power-of-big-data-and-iot/>)

(Tech Pro Research)

No framework is ubiquitous, but there are a few standouts. Here are the top 10 big data frameworks, according to the report:

1. Spark (31%)
2. Hive (17%)
3. HBase (17%)
4. MapReduce (15%)
5. Presto (13%)
6. Kafka (13%)
7. Impala (11%)
8. Storm (11%)
9. Flink (9%)
10. Pig (6%)

efficient

(<https://www.techrepublic.com/article/tips-for-making-data-cleanup-easier-and-more-efficient/>)

- [Walmart's next healthcare move: Using data to identify bad doctors](https://www.techrepublic.com/article/next-healthcare-move-using-data-to-identify-bad-doctors/)
(<https://www.techrepublic.com/article/next-healthcare-move-using-data-to-identify-bad-doctors/>)



Master machine learning with elastic compute resources

Smart businesses are discovering that cloud-based machine learning platforms provide scalable resources and dozens of helpful tools.

SPONSORED BY [AWS/INTEL](#)

While many of these numbers are in the same range, the report also noted that some increased in use over 2017, while others decreased. Spark, HBase, Presto, Kafka, Impala, Flink, and homegrown approaches all rose in popularity in 2018, while Hive, MapReduce, Storm, and Pig all dropped in usage.

Businesses may be prioritizing big data initiatives, but talent shortages remain a major issue, the report found. Three-quarters (75%) of respondents said they faced a headcount shortfall in engineers, scientists, and operators who can deliver big data value. While 79% of businesses

said they want to increase their data team headcount in the next year, 83% said that it's very difficult finding data professionals with the necessary skills and experience.

Organizations experience several other challenges when it comes to big data, the report found. The most common big data roadblocks named were a lack of experience that slows progress (44%), keeping up with new data sources (42%), constantly evolving use cases (41%), too many manual tasks (38%), and the volume of data (34%).

As more companies look to implement machine learning programs across a wide range of use cases, strong big data practices become essential, the report noted. Top priorities for machine learning initiatives in the next year include improving data security and threat protection, optimizing the customer experience, and leveraging predictive maintenance, according to the report.

The big takeaways for tech leaders:

- *Spark, Hive, and HBase are the most popular big data software frameworks used in the enterprise. — Qubole, 2018*
- *75% of data professionals said they faced a headcount shortfall in engineers, scientists, and operators who can deliver big data value. — Qubole, 2018*

Keep up to date on all of the newest big data trends. Click here to subscribe to the TechRepublic Big Data essentials newsletter.

SUBSCRIBE

Also see

- [How big data is changing farming \(PDF download\)](https://www.techrepublic.com/resource-library/downloads/how-big-data-is-changing-farming-pdf-download/) (<https://www.techrepublic.com/resource-library/downloads/how-big-data-is-changing-farming-pdf-download/>) (TechRepublic)
- [Shell drills for data insights](http://www.zdnet.com/article/shell-drills-for-data-insights/) (<http://www.zdnet.com/article/shell-drills-for-data-insights/>) (ZDNet)
- [Cheat sheet: How to become a data scientist](https://www.techrepublic.com/article/cheat-sheet-how-to-become-a-data-scientist/) (<https://www.techrepublic.com/article/cheat-sheet-how-to-become-a-data-scientist/>) (TechRepublic)
- [AI and big data converge to improve your airport customer experience](http://www.zdnet.com/article/examining-the-technology-behind-your-airport-experience/) (<http://www.zdnet.com/article/examining-the-technology-behind-your-airport-experience/>) (ZDNet)
- [How to avoid big data project failures: Your 5-step guide](https://www.techrepublic.com/article/how-to-avoid-big-data-project-failures-your-5-step-guide/) (<https://www.techrepublic.com/article/how-to-avoid-big-data-project-failures-your-5-step-guide/>) (TechRepublic)





Image: iStockphoto/Ralwel

RELATED TOPICS:

[DATA CENTERS](#)[DIGITAL TRANSFORMATION](#)[INTERNET OF THINGS](#)[CXO](#)[ARTIFICIAL INTELLIGENCE](#)[DATA MANAGEMENT](#)

About Alison DeNisco Rayome

Alison DeNisco Rayome is a Staff Writer for TechRepublic. She covers CXO, cybersecurity, and the convergence of tech and the workplace.

Recommended

Promoted Links by Taboola

These Twins Were Named "Most Beautiful In The World," Wait Till You See Them Today

[Give It Love](#)

Minneapolis, Minnesota: This Unbelievable, Tiny Company Is Disrupting a \$200 Billion Industry

[EverQuote Insurance Quotes](#)

You're Probably Overpaying at Amazon – This Genius Trick Will Fix That

[Honey](#)

Why IT needs to factor drones into its big data plans

Why MongoDB and Cloudera prove big data is now mainstream in the enterprise

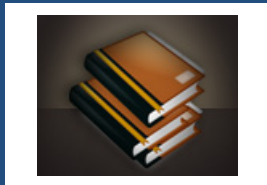
Salesforce update uses automation to limit manual tasks, speed sales cycle

ADD YOUR COMMENT

WHITE PAPERS, WEBCASTS, AND DOWNLOADS

Downloads // From Select Hub

Workday HCM vs. Dayforce vs. Oracle HCM



Find out which of these top human capital management (HCM) systems is best for your needs. Get the pros and cons for each across 20 key business and functional requirements.

Get recommendations, selection tools and software pricing based on your specific needs. More than 160 HRIS system evaluations are available.

CONTINUE

Downloads // From Select Hub

Top Supply Chain Management Software for 2018

CONTINUE

White Papers // From Commvault

5 Essential Cloud Data Recovery Lessons

FIND OUT MORE

White Papers // From TechRepublic.com

Tech Pro Research: IT Leader's Guide to Deep Learning

LEARN MORE

Downloads // From Select Hub

Top HR Software for 2018 - Expert Reviews & Pricing - Free Analyst Report

SPECIAL OFFER

WE'LL GET
YOU THERE,
SECURELY.

LEARN HOW

NETSCOUT.

FEATURED CONTENT



AI for Business

Amazon & Intel help you
innovate with AI



The Power of AI

Amazon & Intel: From
Insight to Action with AI

EDITOR'S PICKS



Can Russian hackers be stopped? Here's why it might take 20 years



The new commute: How driverless cars, hyperloop, and drones will change our travel plans



Exomedicine arrives: How labs in space could pave the way for healthcare breakthroughs on Earth



How Sephora is leveraging AR and AI to transform retail and help customers buy cosmetics



RECOMMENDED

Sponsored Links by Taboola

These Twins Were Named "Most Beautiful In The World," Wait Till You See Them Today

Give It Love

Minneapolis, Minnesota: This Unbelievable, Tiny Company Is Disrupting a \$200 Billion Industry

EverQuote Insurance Quotes

FREE NEWSLETTERS, IN YOUR INBOX

Tech News You Can Use

We deliver the top business tech news stories about the companies, the people, and the products revolutionizing the planet.

Delivered Daily

 **SUBSCRIBE**

Best of the Week

Our editors highlight the TechRepublic articles, galleries, and videos that you absolutely cannot miss to stay current on the latest IT news, innovations, and tips.

Delivered Fridays

 **SUBSCRIBE**



Software-Defined WAN for Dummies

Best Practices
Benefits
ROI

Free eBook

CLOUDGENIX

