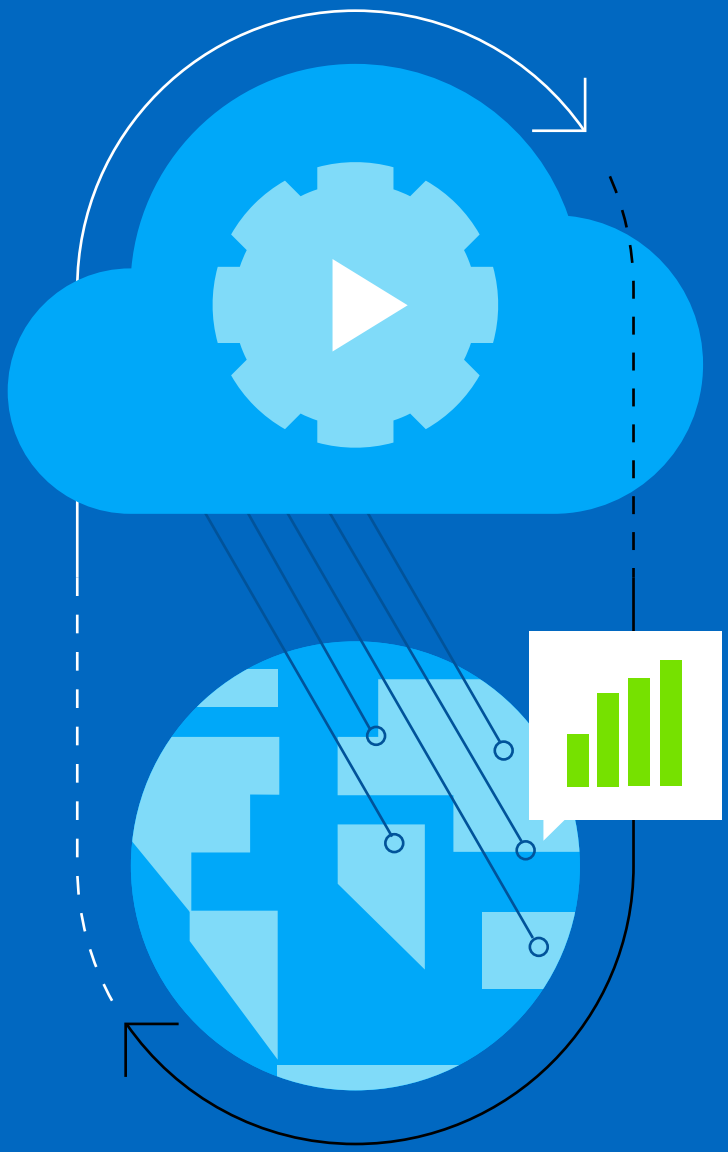
Progress®



2017 Data Connectivity Outlook

Discover the latest trends in our 4th annual database technology survey

REPORT

Table of Contents

Executive Summary

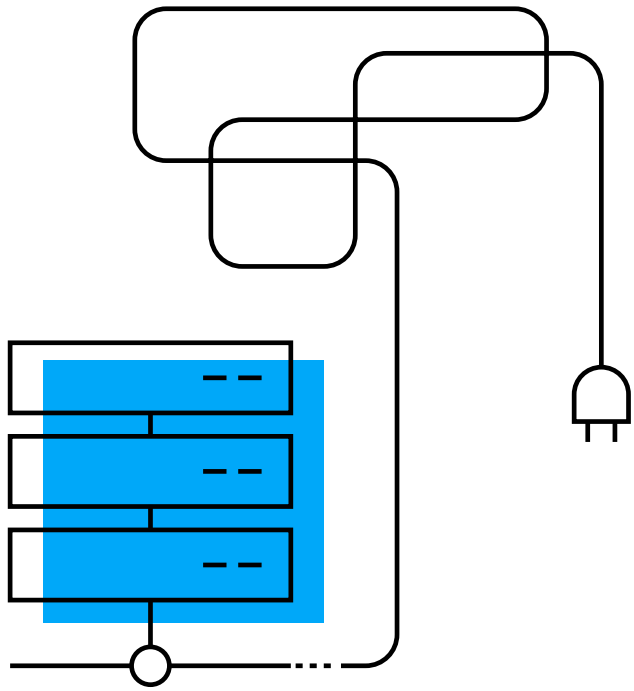
Progress is excited to present the results of its 4th annual Data Connectivity Outlook, based on 1,200 survey responses. Respondents included business and IT professionals in various roles, representing a range of industries and organization sizes across the globe.

This year's report demonstrates the impact of the changing landscape of disruptive data sources. Organizations must deal with data spread across a variety of sources—from relational to big data to SaaS. In fact, the survey reveals that SaaS data sources increased significantly from last year (62% to 79%). To overcome the challenge of exploding data sources, companies are adopting data access standards – SQL (ODBC, JDBC) and REST (OData) – to connect these disparate data sources faster.

Adoption of analytics is growing as businesses use data to drive decisions. Among the applications survey respondents use standards for, analytics and reporting are high priority areas (63%). Reflecting that trend, more ISVs are embedding analytics capabilities within business applications. While this strategy provides basic analytic functions, open analytics is growing rapidly as a complementary strategy to connect a cloud application's data to third-party tools and programming languages.

Regardless of how the data landscape has evolved, the survey shows that enterprises still have significant investments in on-premise RDBMS and big data interfaces. Using the cloud requires that they seamlessly integrate the two environments. However, many voiced concerns about accessing that data from the cloud. Traditional technologies for connecting these database silos are difficult to manage and neither scalable nor engineered for the cloud.

Clearly, a scalable, firewall-friendly solution is needed.



Who We Surveyed

1200
PARTICIPANTS

38%
EMEA

36%
NA

20%
APJ

7%
SA

Role

Developers
Architects
IT Professionals
Executives
Business Analysts
Product Managers

Industry

IT Services
Manufacturing
Financial Services
Education
Retail
Government/Military
Healthcare
Telecommunications

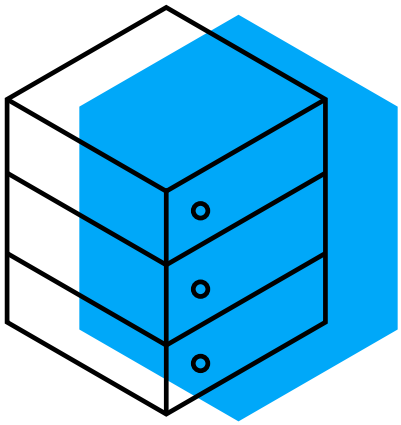
and more...

Data Continues to Spread Across a Variety of Sources

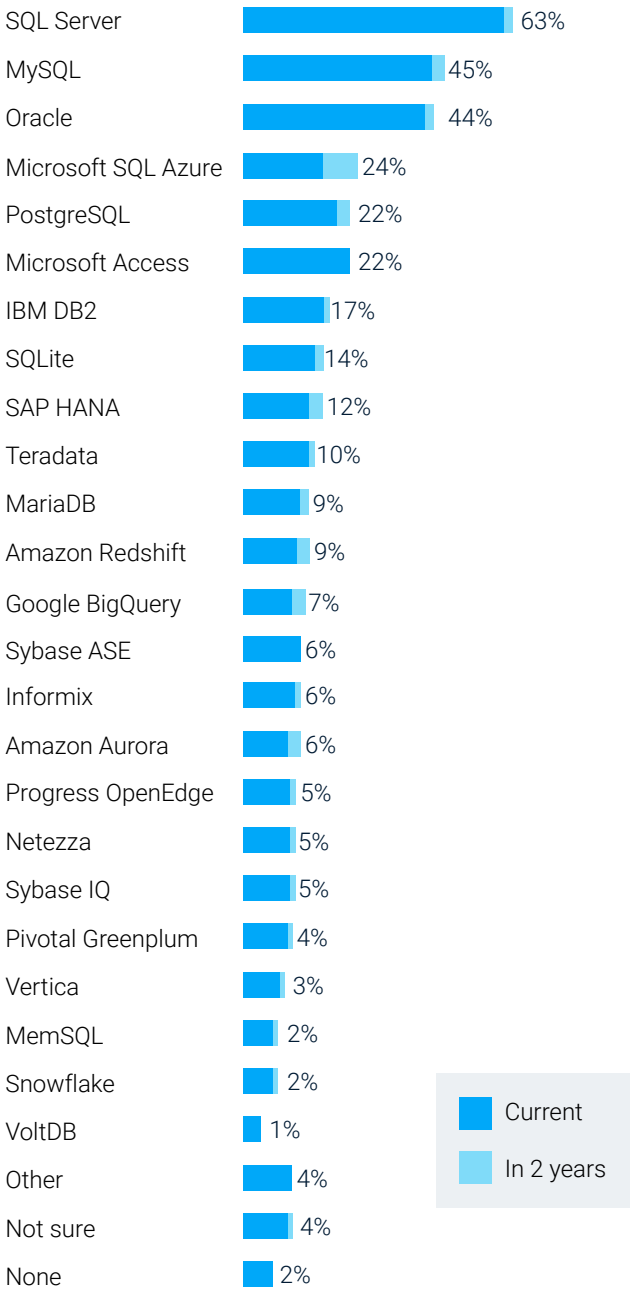
In the last few years of doing this survey, we've seen an explosion in the number and types of data sources, as well as a proliferation of third-party APIs. This is in keeping with trends seen across the industry. To illustrate, in 2005 there were 105 open APIs. Today there are over 16,000 and 40 more are being made public every week. We've also observed a growing number of SaaS applications, with every SaaS instantly becoming another database.

Relational Data Sources

Relational databases are still going strong in the enterprise. SQL Server, Oracle and MySQL maintain their position at the top of relational databases used from 2016 to 2017, with usage remaining relatively the same from year to year.



Which relational/analytics database technologies do you or your customers currently use? And which do you or your customers plan to adopt in the next two years?



SaaS Data Sources

From 2016 to 2017, SaaS adoption increased from 62% to 79%. More than 50% of respondents use two or more SaaS data sources, while more than 35% use three or more. This rapid growth in SaaS adoption is primarily driven by CRM, HRM and marketing applications. Of these, the top data sources are Office 365, Salesforce and Google Analytics, which increased the most (11%) since 2016.

As enterprises attempt to democratize SaaS data, the accessibility of a SaaS application will determine its success in the long run. Learn more about the importance of data connectivity for SaaS applications in our connectivity outlook for CRM leaders.

Relational databases remain critical to business processes, while adoption of SaaS data sources is increasing consistently, year over year.

As the number of SaaS apps grows, so does the need for connectivity between SaaS and on-premise databases to unlock and monetize the data.

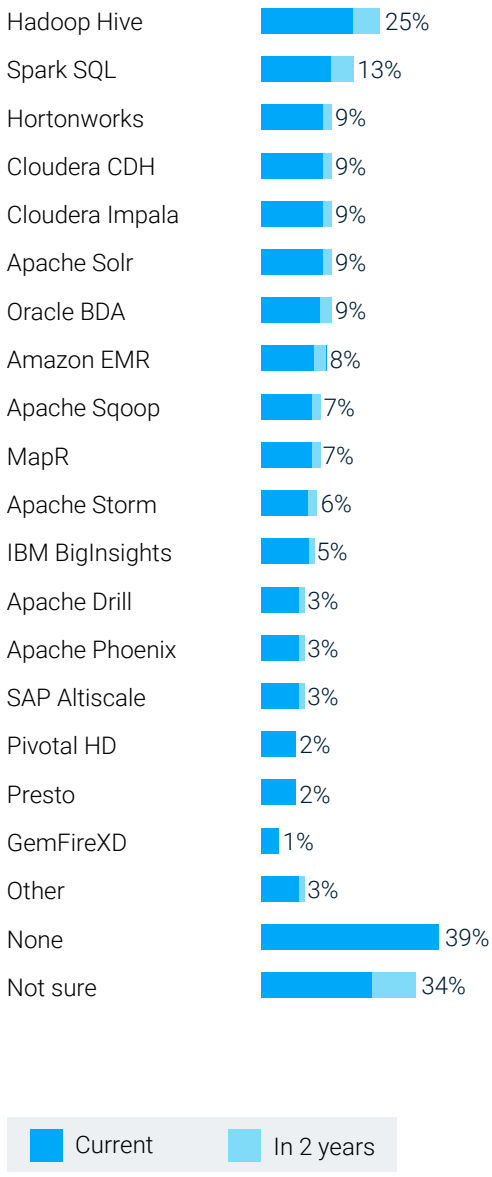
What SaaS datasources do you or your customers currently use? And which do you or your customers plan to adopt in the next two years?

Big Data Sources

Hadoop Hive remains the most popular big data interface for two years in a row, showing a 5% increase in adoption from 2016 to 2017. Adoption of Hadoop Hive is expected to increase 8% in the next two years.

Hortonworks is the most popular Hadoop Hive distribution, followed by Cloudera and Amazon EMR. SparkSQL is the most popular interface for Apache Spark, with a significant 6% expected increase in adoption in the next two years.

What big data platforms/interfaces do you or your customers currently use? And which do you or your customers plan to adopt in the next two years?



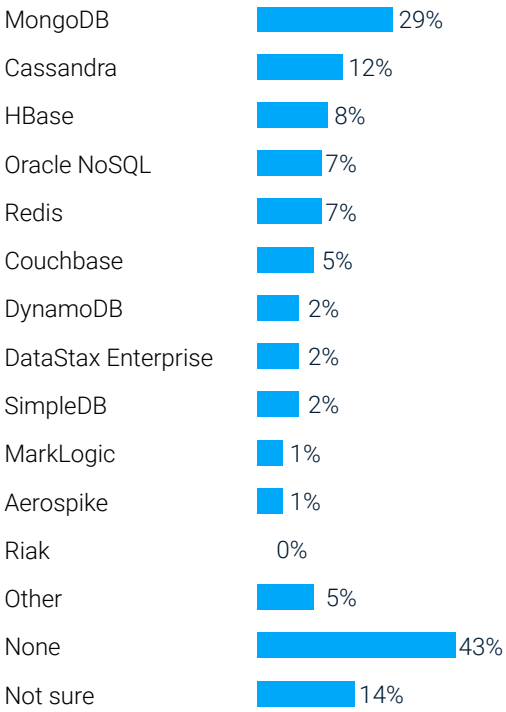
Big data saw 11% year-over-year growth compared to last year, thanks to constant disruptions in the big data ecosystem. These disruptive technologies with frequent release cycles make it impossible to continuously certify and maintain connectivity for hundreds of big data components and versions, which change monthly in many cases.

For that purpose, enterprises started relying on technology partners for day one support_ to all these versions.

NoSQL Data Sources

NoSQL databases have consistently increased in popularity for real-time and batch data analytics. More than 50% of those surveyed use NoSQL databases, increasing by 8% from last year. MongoDB remains the most popular NoSQL database for two years in a row, with adoption increasing 5% from last year. Enterprises are now looking to run analytics on this MongoDB data. In the wake of this new development, third party ODBC and JDBC drivers for MongoDB are becoming popular.

Which NoSQL database technologies do you or your customers currently use or support?



End-users continue to demand SQL access to NoSQL data for analysis.

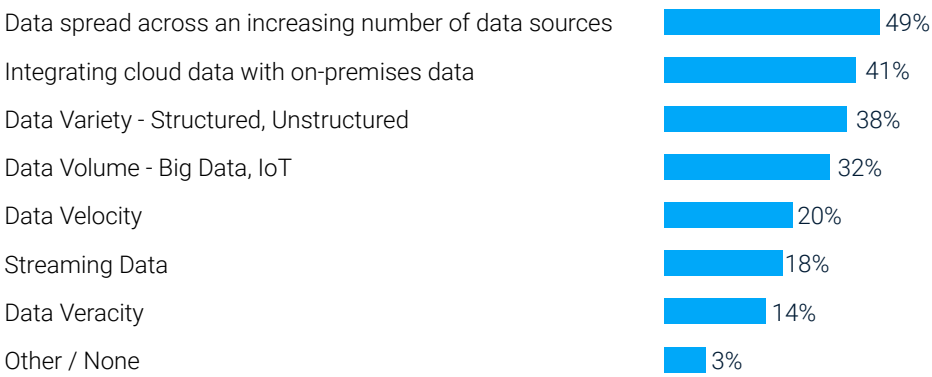
Spark SQL is a recommended interface for Cassandra/DataStax for interactive, real-time analytics.

Spark SQL usage has increased from 7% last year to 13% this year.

Increasing Number of Data Sources is the #1 Data Integration Challenge

To provide a unified view of the data for ad hoc analysis and business intelligence, most companies are combining data across several disparate sources. The biggest challenge is incorporating all relevant data across an increasing number of cloud, database and other third-party sources.

What are your biggest data integration challenges?



Data Integration can give you a true 360° view of your enterprise data, but it also introduces its own set of challenges, such as increasing data sources, hybrid connectivity and data variety. These issues make it difficult to scale adequately to handle the increasing need for self-service analytics.

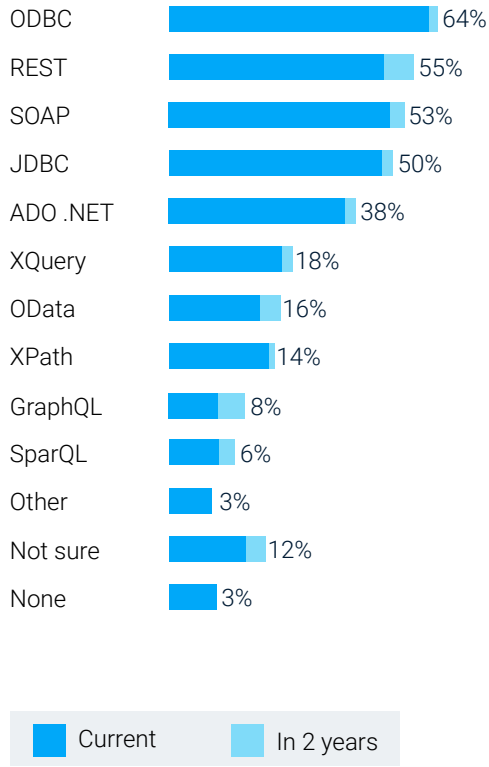
Standards: The Key to Faster Data Access Across Expanding Data Sources

Organizations are increasingly adopting standard SQL and REST data access standards to support more disparate data sources faster, while ensuring interoperability and compatibility with existing systems. OData is an industry standard REST API and is growing in popularity.

The proliferation of data sources and APIs continues to expand each year. Each new data source is unique and has its own API and requirements. You could develop against the APIs yourself, but that requires expertise in each database API, as well as ongoing maintenance and support.

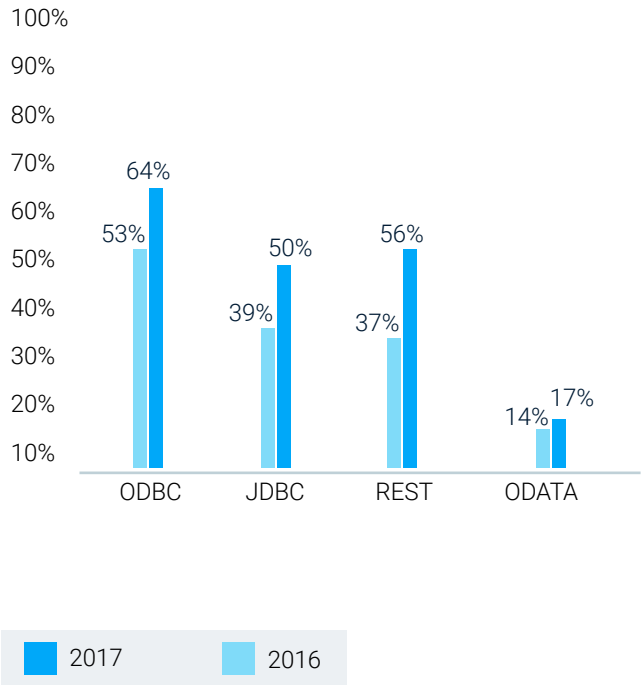
The chart shows use of data access standards is increasing, reaching almost 100% of respondents this year over 88% last year. More than 75% of survey respondents use two standards and more than 50% use three or more standards.

Which data access interfaces do you or your customers currently use? And which do you or your customers plan to adopt in the next two years?



More companies are relying on standards-based data connectivity specialists like Progress to stay on top of all these database and API changes so engineering resources can be focused on projects that differentiate the business, not data connectivity.

From 2016 to 2017, use of ODBC increased 11%, JDBC 11%, OData 3%, while REST increased significantly by 19%.



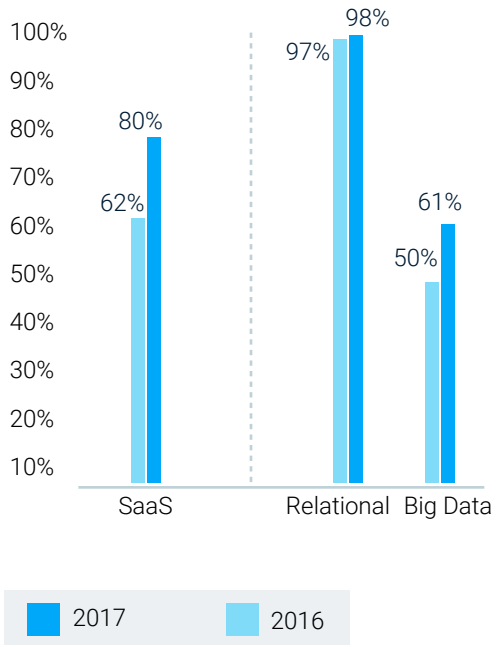
Each standard is closely linked to different trends in data consumption across various shapes, sizes and locations. Big data sources have been driving adoption towards SQL standards such as ODBC and JDBC.

And the OData REST API is gaining ground in hybrid and cloud environments because of its rich set of interoperable query capabilities to unify data access for disparate resources across the Internet.

Other popular standard APIs to query data are emerging such as GraphQL and ORDS. Learn more in this industry debate: OData vs GraphQL vs ORDS.

Increasing Cloud Adoption Makes Real-Time Hybrid Connectivity Critical

As noted earlier, SaaS adoption increased significantly since our 2016 survey. Growing SaaS adoption and the necessity of on-premises data (both relational and big data) highlight the need for real-time connectivity across the entire hybrid architecture.

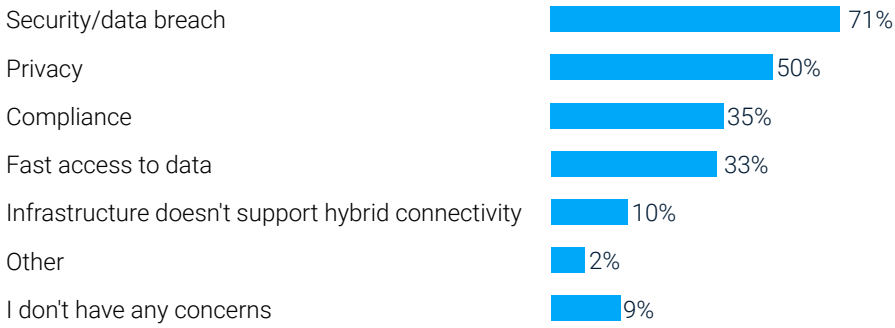


To overcome the challenge of exploding enterprise data sources and increasing SaaS adoption, the on-premises and cloud systems must be connected in a seamless way.

Hybrid connectivity leverages open standards to easily connect the expanding hybrid environment to get the most value out of the data.

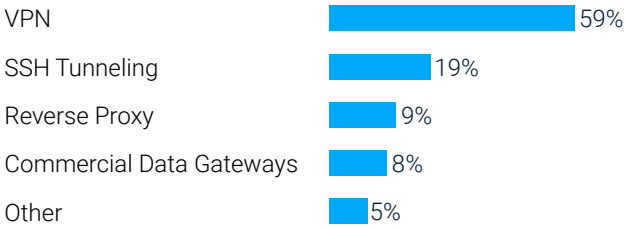
However, connectivity between SaaS platforms and on-premise systems behind the firewall is a complex issue. And security vulnerabilities, data privacy and compliance are top challenges.

What are your major concerns while exposing on-premises data beyond the firewall?



To solve the challenge of connecting cloud and ground, respondents are primarily using VPN, SSH tunneling and reverse proxy.

How do you currently traverse the firewall?



Enterprises today are rapidly adopting the cloud, while continuing to retain sensitive business data inside the firewall. Connecting these database silos (big or relational) using technologies such as VPN or SSH tunneling are difficult to manage and not engineered for accessing business data from the cloud.

VPN is a secure option for traversing the firewall, but it becomes very complex when you have to scale it. It also it involves plenty of manual effort because there are many points of failure to manage.

SSH tunneling requires administration, poses security risks to manage SSH keys, and doesn't scale well.

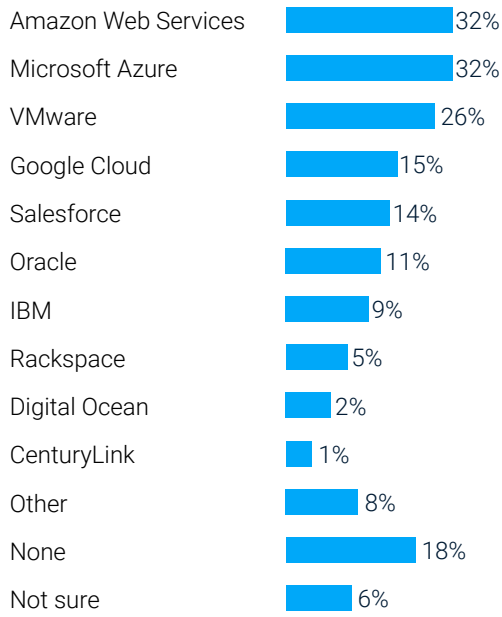
Commercial gateways such as our vendor-agnostic hybrid connectivity solutions are gaining in popularity because they are engineered for the cloud to support standard data access for both on-premises and hybrid scenarios—without changing the firewall configuration.

AWS and Microsoft Azure Are Leading Cloud Services Platforms

This surprising result shows use of Amazon Web Services (AWS) and Microsoft Azure to be the same. In most regions, such as North America, AWS is leading Azure. Further analysis revealed that we received more survey responses from Europe, where Azure is more popular.

In the Other category, many cloud providers were mentioned such as Linode, Red Hat OpenShift, and Microsoft OneDrive, indicating a diverse market.

Which cloud infrastructure do you or your customers currently use?



More and more organizations are adopting on-demand clouds to help their businesses scale and grow. For organizations with large investments in on-premises systems, hybrid connectivity is a necessary part of cloud adoption.

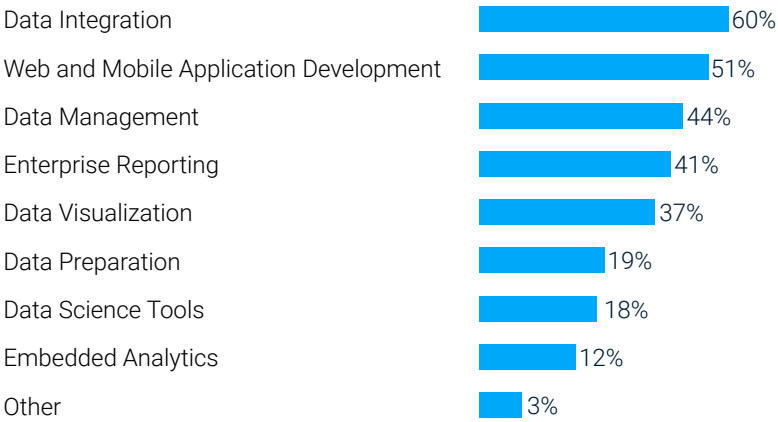
To help you deploy hybrid connectivity on cloud computing platforms such as AWS, Azure, VMware and Google Cloud, see our cloud and hybrid tutorials.

Open Analytics is Growing, Along with Embedded Analytics

Although standard interfaces are used more often for data integration projects (60%), app development (51%) and data management (44%), the total share of respondents using reporting and analytics tools is 63%. Adoption of analytics is growing as businesses use data to drive decisions. This trend is giving rise to open analytics as organizations seek to query cloud applications with their favorite analytics tool or programming language.

Open analytics is the integration of an open data access layer (such as ODBC, JDBC or OData) into business applications (such as CRM, ERP, Finance or HR) to be directly consumed by external analytics tools and popular programming languages.

Which types of applications do you connect to from standard SQL/REST interfaces?



In addition, out of those who selected embedded analytics, 45% are using data science tools, 59% are using data visualization tools, and 65% are doing enterprise reporting. This shows that embedded analytics alone will not serve the diverse analytics needs of an organization. Companies are thus adopting both embedded analytics and external analytics tools such as Tableau, PowerBI, Qlik or MicroStrategy for advanced analytics.

	No Embedded Analytics	Embedded Analytics
Data Science Tools	14%	45%
Data Visualization	35%	59%
Enterprise Reporting	38%	65%

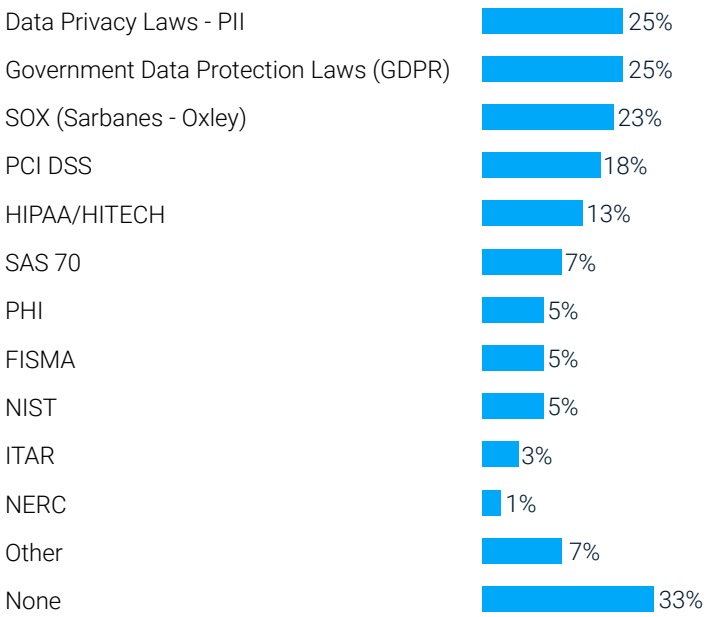
One of the latest trends among ISVs is embedding analytics functionality in their solutions to give users data, dashboards and reporting. But users also want to perform complex analysis on aggregated data using BI tools and programming languages. With open analytics, customers can use the tool or language of their choice to query cloud applications. These strategies work together to meet the diverse needs of business users, data analysts and data scientists.

The Complex Landscape of Data Protection Laws

In an era in which technology increasingly facilitates the exchange of information, organizations have to comply with more data security standards and laws than ever before. More than 66% of respondents say they must comply with one or more standards.

Industry-specific regulations such as Sarbanes-Oxley, PCI DSS and HIPAA, and government regulations such as PII and GDPR, top the list of data protection/privacy laws. Respondents wrote in many more, included in the Other category, such as PIPEDA in Canada, FERPA in the US and standards required by the EU. The landscape of privacy and protection laws around the world continues to evolve and grow more complex.

Which security/data standards are you required to comply with?



The different country or industry compliance requirements should be considered as you decide on your data connectivity strategy so you don't put compliance and security at risk.

To help you comply with requirements such as GDPR, we recently launched a self-hostable hybrid connectivity solution called Hybrid Data Pipeline. This can be hosted anywhere: any cloud, any on-premises and any country. You can now manage your own security while leveraging our innovative hybrid connectivity solution.

Key Takeaways

- Relational databases are still going strong
- SaaS adoption increased from 62% to 79% since last year
- Hadoop Hive is the most popular big data interface and is expected to grow 8% in the next two years
- More than 50% use NoSQL, with MongoDB being the most popular
- The number one integration challenge is an increasing number of data sources
- Almost 100% use data access standards, and adoption is increasing
- OData, the industry standard REST API, is growing in popularity
- The need to connect cloud and ground data is rife with concerns about security and compliance
- Amazon Web Services and Microsoft Azure are leading cloud service providers in an increasingly diverse market
- 63% use reporting and analytics tools, giving rise to open analytics
- More than 66% must comply with data protection laws

Learn More About Our Solutions

About Progress DataDirect

Progress® DataDirect® delivers powerful data connectivity for applications running on-premise or in the cloud. Get connected to the vast data landscape using a single standard interface with SQL or REST. DataDirect connectivity solutions support the full range of relational, cloud, NoSQL and Big Data sources across trusted industry standards, including ODBC, JDBC, ADO.NET and OData (REST).
Learn more at www.progress.com/datadirect-connectors

About Progress

Progress (NASDAQ: PRGS) offers the leading platform for developing and deploying mission-critical business applications. Progress empowers enterprises and ISVs to build and deliver cognitive-first applications, that harness big data to derive business insights and competitive advantage. Progress offers leading technologies for easily building powerful user interfaces across any type of device, a reliable, scalable and secure backend platform to deploy modern applications, leading data connectivity to all sources, and award-winning predictive analytics that brings the power of machine learning to any organization. Over 1700 independent software vendors, 80,000 enterprise customers, and 2 million developers rely on Progress to power their applications.
Learn about Progress at www.progress.com or +1-800-477-6473.

Worldwide Headquarters

Progress, 14 Oak Park, Bedford, MA 01730 USA
Tel: +1 781 280-4000 Fax: +1 781 280-4095
On the Web at: www.progress.com
Find us on facebook.com/progresssw twitter.com/progresssw youtube.com/progresssw
For regional international office locations and contact information,
please go to www.progress.com/worldwide

Rev 2017/06 | 170522-0070

Progress®