Vendor Landscape: Graph Databases

Leverage Graph Databases To Succeed With Connected Data

by Noel Yuhanna
October 6, 2017

Why Read This Report

Whether you are doing customer analytics, fraud detection, or risk assessment, you need connected data. While linking a few data points or data sources is often simple, creating connections with significant amounts of heterogeneous data requires a radically new approach. Enterprise architects should read this report to learn how the new and emerging graph database technology offers the ability to link pieces of data easily to create more meaningful and actionable insights.

This report was originally published as a market overview in 2015; we've updated it to reflect new vendors and a broader set of emerging use cases.

Key Takeaways

Connected Data Insights Are Critical For Success
Business users are demanding faster, more real-time, and integrated data to support various business insights. Enterprise architects are looking for solutions that can deal with connected data without writing code.

Graph Helps Connect Data Easily
Graph databases offer a simplified way to connect silos of data using SQL-like queries that do not require programming experts.

The Graph Database Vendor Landscape Is Growing Rapidly
Forrester expects the graph database market to have significant success in the coming years as organizations leverage more connected data to support actionable insights. This report highlights 10 graph database vendors and open source projects.

Vendor Landscape: Graph Databases

Leverage Graph Databases To Succeed With Connected Data

by Noel Yuhanna
with Gene Leganza and Jun Lee
October 6, 2017

Table Of Contents

Related Research Documents

The Forrester Wave™: Big Data Fabric, Q4 2016

The Forrester Wave™: Big Data NoSQL, Q3 2016

The Forrester Wave™: Document Stores, Q3 2016

Share reports with colleagues.
Enhance your membership with Research Share.

Forrester Research, Inc., 60 Acorn Park Drive, Cambridge, MA 02140 USA
+1 617-613-6000 | Fax: +1 617-613-5000 | forrester.com

Connected Data Is Creating New Business Opportunities

It's all about connected data! Connecting data helps companies answer complex questions, such as "Is David's credit card purchase a fraud, given his buying patterns, the type of product that he is buying, the time and location of the purchase, and his likes and dislikes?" or "From the thousands of products, what is Jane likely to buy next given her buying behavior, products she has reviewed, her purchasing power, and other influencing factors?" Developers could write Java, Python, or even SQL code to get answers to such complex questions, but that would take hours or days to program and in some cases might be impractical. What if business users want answers to such ad hoc questions quickly, with no time for custom code or with no access to the technical expertise needed to write those programs? While organizations have been leveraging connections in data for decades, the need for rapid answers amid radical changes in data volume, diversity, and distribution has driven enterprise architects to look for new approaches.

Today, enterprises want to leverage various big data sources to:

> **Extract deeper and actionable meaning quickly.** Today, many firms do not have advanced data-literate staff, such as data scientists, who can use data preparation and advanced analytics tooling to find deeper, hidden insights from data.[1] Even if they do, taking the time to write complex code prevents organizations from responding quickly to micro trends or minimizing customer churn when rapid action is called for.

  "We would love to support more actionable insights that focus on improved business outcomes; however, due to lack of sufficient technical expertise we cannot develop these. We are now looking at ways to automate them or using technologies that require no coding." (Director of IT, North American telco)

> **Integrate all customer data to drive new customer experiences.** An increasing focus on the customer experience is driving firms to integrate data from such diverse sources as internal systems of record and social media feeds. That integration is complex, as it is typically not stored in a single repository or even one data center, thus requiring big data or new approaches to support messy data.[2] Finding connections in increasingly distributed customer data presents an architectural challenge.

  "We are experiencing tremendous data explosion. We are finding that integrating all customer data to drive new customer experiences is becoming a challenge, especially as we leverage large, complex data sets that's on-premises and in the cloud." (VP IT, North American retailer)

> **Navigate connected data to deliver contextual intelligence.** In many cases, firms must be able to act upon on connected data in real time, as in the case of fraud detection or recommendation engines. Forrester's customers have reported that traversing a complex landscape of highly connected data and providing contextual insights using traditional technologies has been extremely challenging.

"We have very large amounts of complex and diverse data sets in our organization across platforms; some are legacy. As a result, today we are finding it hard to support contextual insights to help with churn, new market opportunities, competitive intelligence, and customer insights." (Enterprise architect, North American insurance company)

› **Provide self-service capabilities for business users to connect data.** Any delay in getting connected data to decision makers can cost businesses revenue and affect customer stratification. Self-service insights have become critical for businesses to succeed in the era of digital business, with business users and analysts connecting and leveraging connected data themselves through automation, machine learning, and artificial intelligence technologies.[3]

"We are heading towards complete self-service data platform capability over the next three to four years, where business users, analysts, and partners will be able to leverage data and insights they need themselves. We still have a way to go, but it's become a big initiative for us to gain competitive advantage and lead ahead." (VP, global data services, financial services firm)

Traditional Ways Of Connecting Data Have Never Been Simple

There are many ways to connect data to create insights, but they vary significantly in the time, effort, and expertise required. The scenarios for combining data are complex: Data could be of various types — structured and nonstructured — with variable frequency, some persisted and others streamed at regular or irregular intervals. Compounding the complexity in increasing data volume and variety is an increasing trend toward distributed data. The data might be on-premises in one or multiple data centers, in multiple clouds, or even at the edges. Enterprises use the following traditional approaches to connect data:

› **Relational databases and data warehouses.** Databases or data warehouses are ideal technologies to store, process, and access connected data because they can organize data into tables, columns, materialized views, and partitions. However, while storing data into these repositories is easy, connecting them through a semantic layer in an automated manner still requires manual effort.

› **Hadoop/data lakes.** Similar to databases, Hadoop-based data lakes can store lots of data, but big data integration is neither trivial nor automated. Developers have to write Python, Java, or C code to create data mapping or leverage data discovery and preparation tools, which still require time and effort to make sensible connections to data that's relevant for the business.

› **Data virtualization.** While data virtualization builds a semantic layer, data architects have to physically and logically map data to create connections. This works well for system-of-record data sets where there is some knowledge about the data; however, the complexity kicks in when introducing new, unknown data such as geolocation, clickstream, or social data and rapid access and integration is needed.

› **Manually through application programs.** Most organizations that connect data to deliver either fraud detection, customer 360, recommendation engines, or social networking use cases are writing C#, Python, Scala, and other programming languages to make those connections, in addition to machine learning capabilities found in various open source and commercial programs. However, this takes time and effort — and a team of developers to work together to deliver these sophisticated applications.

Graph Is The Fastest Way To Connect Data As Complexity Grows

While relational databases have existed for decades, graph database is a new and emerging hot category that focuses on connected data to help organizations ask more complex questions easily without writing complex programs or requiring deep technical expertise. Graph databases use nodes (things), edges (relationships), and properties (key-value pairs) to store and access connected data. Each data point is a node; the relationship between two data points is called an edge. Nodes are very similar to the objects familiar to object-oriented developers. Graph databases have some of the flexibility of key-value stores but also offer full support for relationships.

Forrester defines graph databases as:

> *Optimized database technologies to store, manage, and access connected data to answer complex question and deliver new insights.*

Two Different Approaches To Graph Exist, With Some Commonality On Use Cases

There are two types of graph technologies, property graph and RDF triple, both of which focus on storing, processing, and accessing connected data but differ from query languages, storage format, and traversal approaches (see Figure 1). Typically, the scope of RDF triple is on a standardized model for data interchange on the web, whereas property graph centers on the graph data and query capabilities. Property graph databases are more optimized for graph traversal, especially when the network becomes more complex. Specifically:

› **Property graphs are "whiteboard friendly."** When business experts draw a picture of their domain on a whiteboard, it's often naturally a property graph (see Figure 2). This helps bridge the gap between the business and the tech organization and reduce the time between the design (logical modeling) and implementation (physical modeling) of new and modified models. Typical property graph use cases include social networking, recommendation engines, real-time recommendations, and fraud detection. Leading vendors include DataStax, Neo4j, and OrientDB.

› **RDF triple stores were originally intended for the Semantic Web.** Created in 1999, RDF databases were designed for web pages and World Wide Web metadata interchange to support the Semantic Web (see Figure 3). They don't store data as a graph model, but they do offer optimized graph query capabilities. An RDF triple store uses SPARQL as the query language. While RDF has

similar types of use cases as property graph, the most common ones include healthcare analytics, master data management (MDM), knowledge management, and risk analytics. Leading vendors include IBM, MarkLogic, and Oracle.

FIGURE 1 Differences Between Property And RDF Triple Store Graph Databases

| Features	RDF triple store	Property graph
Foundation	Created in 1999 by W3C, which had the mission to develop global web standards	Created by Neo Technology in 2000 to target broader use cases
Scope	A standard model for data interchange on the web	Centers on database storage and query capability
Focus	Edge-centric	Property-centric
Industry footprint	Leading verticals include healthcare, government (intelligence), financial services, pharma, and publishing	Leading verticals include retail, banking, insurance, investment, eCommerce, and high-tech
Query language	SPARQL	Cypher, GSQL, GraphLog
Storage format	Rows of RDF triples	Weighted graphs, hypergraphs, and undirected graphs
Traversal	Not fully optimized	Optimized to find the shortest path

FIGURE 2 Property Graph Model
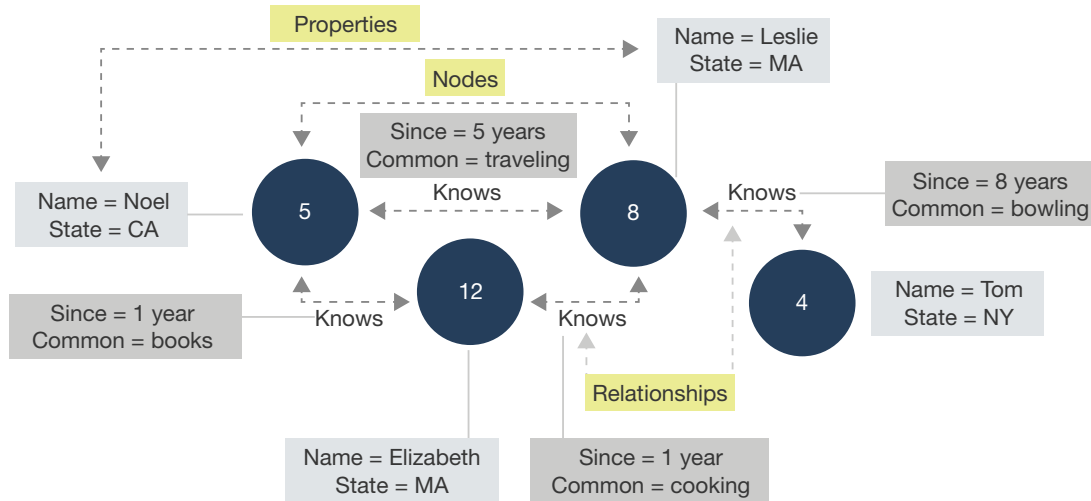
FIGURE 3 RDF Triple Store Model
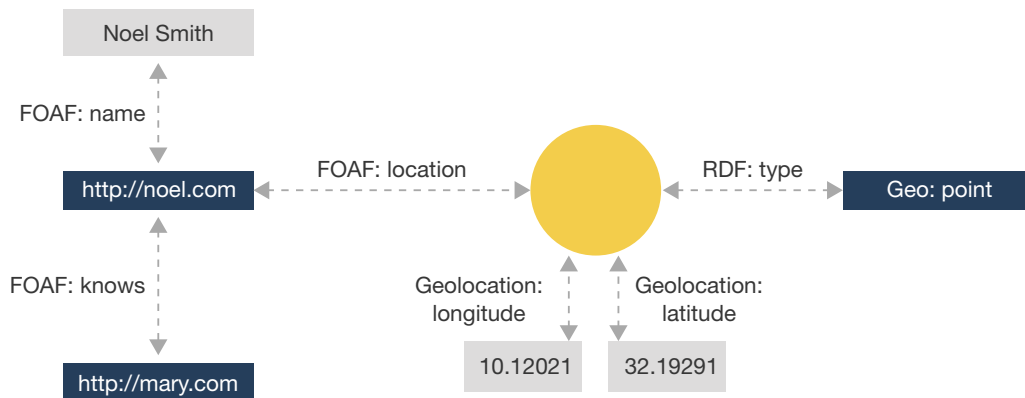
Multiple Graph Query Languages Exist, With Some Standardization

Query languages are a critical component of any graph solution; however, no standardized language currently exists among all graph solutions.

› **Most RDF solutions leverage SPARQL.** SPARQL is a semantic query language to access data in RDF format. It was made a standard by the RDF Data Access Working Group of the World Wide Web Consortium.

› **Property graph solutions fall into two major camps — Gremlin and Cypher.** Gremlin is a graph traversal language developed by Apache TinkerPop of the Apache Software Foundation. It is an imperative and declarative language that links traversal operators to form path-like expression. Vendors that support Gremlin include ArangoDB, DataStax, IBM, Microsoft, Oracle, and OrientDB. On the other hand, Cypher is a declarative graph query language that allows for expressive and efficient querying and updating of property graph. Cypher was originally developed by Neo4j for its graph database but opened up through the OpenCypher project in October 2015.

Vendors that are adopting Cypher language include AgensGraph, Neo4j, SAP, and TIBCO Software. While the two camps offer choices, Gremlin remains the most widely used language among vendors. Based on customer feedback, we find that Gremlin is better for simple cases, but when it comes to complex queries with many properties, Cypher can outperform Gremlin.

Graph Use Cases Expand To Support More Complex Requirements

Graph technology gives customers new insights and intelligence that were previously impossible or hard to find or detect. While the use cases for graph continue to expand, today we find enterprises using it for:

› **Recommendation engines.** Graph databases connect the dots from internal or external information gathered about the customer to deliver intelligence, such as what John and Mary like or dislike and what products and services they are likely to buy. While nongraph technologies can support recommendation engines, graph creates faster time-to-value. For example, Walmart leverages Neo4j graph for real-time recommendations to understand online shopper behavior and the relationship between customers and products as well as to capture new interests shown by the customer's online visit.[4]

› **A 360-degree view of the business.** Major retail stores, midsized retailers, and eCommerce sites have started to exploit graph technology in a big way. Customer data stored and processed through graph helps them identify what new products to launch or what gaps exist in packaging, pricing, and delivering products in various regions. For example, a large retailer in North America leverages graph to understand its product performance periodically, minimize customer churn, and ensure that inventory is optimized across various regions.

› **Fraud detection.** Detecting fraud in real time is extremely challenging, requiring complex algorithms and intelligence to avoid false positives. For example, a consumer might use the same credit card twice to fill up with gas on the same day. Is it fraud or simply someone filling up both of her cars? Today, the graph database is one of the top technologies used to detect fraud, uncover fraud rings, and identify sophisticated scams including money laundering. For example, RTE, one of Ireland's prominent broadcasting companies, leverages OrientDB as part of its investigation unit to link various data. When tasked to investigate public representatives and their purported assets, RTE needed to form links between seemingly unstructured and disconnected data to uncover undisclosed assets and expose their respective representatives.

› **Master data management.** Master data has varied hierarchies that run top-down and diagonally and which are often built on top of a relational model that is complex, slow, and not very agile. Traversing complex, interconnected data sets in real time is often challenging — but adding context is even harder. A graph database, although relatively new to MDM, puts master data into a more intuitive, intelligent, and integrated context. MDM systems with graph databases provide a unified view of core business data and deliver intelligence about product performance and customer insights. For example, Novetta markets and supports an entity analytics solution using graph technology. This firm's solution provides an advanced business view of data, validates data sets, builds business cases, and test-drives the application. This delivers on a cyber and discovery analytics solution that detects adversarial personas, networks, signals, and technologies.[5]

› **Social networking.** Social media apps are all about connections. How do you build the next big social app that offers smart recommendations about whom to follow? Relationships are important in social networking, and connecting users is critical. Social networks generally evolve very quickly. Graph databases can traverse social networks and related data very quickly, too, delivering powerful recommendations and predictive insights. For example, social media companies such as Facebook, LinkedIn, and Twitter all leverage some sort of graph processing within their platform to identify friends and families across the globe.

The Graph Database Market Is Gaining Momentum

Forrester has seen the number of graph inquires double over the past 12 months, reflecting enterprises looking at graph architectures and products closely to address their connected data challenges. In addition, in a recent Forrester survey, 51% of global data and analytics technology decision makers either are implementing, have already implemented, or are upgrading their graph databases (see Figure 4). We expect the graph database market to grow significantly as organizations look to new approaches in dealing with silos of data. There are now more than 10 graph database vendors and open source projects, and we're likely to see startups and new vendors expanding their focus on graph either as part of their embedded solution or as a standalone offering (see Figure 5). In addition, we are likely to see vendors such as Microsoft, MongoDB, SAP, and Teradata further integrate graph technology into their data platforms to deliver more integrated big data analytics and graph solutions. Here are 10 of the top graph database vendors today:

› **ArangoDB natively supports graph as part of its multimodel database.** ArangoDB is a privately held German company founded in 2014. It offers an SQL-like query language called ArangoDB Query Language (AQL) in addition to support for TinkerPop/Gremlin. It runs on various hardware and operating systems including Kumbernets and DC/OS. ArangoDB supports full ACID compliance on single-instance and single-document transactions in cluster configuration. Besides transactions, it supports operational and analytical workloads. Customers use ArangoDB for a single view of everything, fraud detection, real-time recommendations, and the internet of things (IoT). ArangoDB is available as open source under an Apache 2 license and in a commercial version that includes advanced security, availability, and performance features. One of the largest

deployments, in terms of data size, has 1.5 billion edges and vertices, with 1.5 TB of data on a six-node HA cluster.[6] Customers include AboutYou, Egress, Karlsruher Institute of Technology, Liaison, and MakeMyTrip.

› **Cray offers a viable RDF graph option that runs on Cray computers.** Cray, headquartered in Seattle, is known for its supercomputer infrastructure and its supporting solutions for data storage and analytics, including a graph product called Cray Graph Engine (CGE). CGE is available as a standalone software package for Cray XC supercomputers or bundled with hardware and other big data analytics packages in the Urika-GX platform, first released in September 2016. However, CGE is currently available on Cray platforms only. Earlier this year, Cray partnered with the Markley Group to provide a Urika-GX offering in a private-cloud setting with CGE preinstalled and ready to use. CGE is an in-memory, vertically scalable graph database that delivers high performance for complex pattern-matching queries. It integrates with Spark, HDFS, and Lustre's parallel file system. Enterprises use CGE for cyber-reconnaissance, life sciences, social network analysis, customer intelligence, and financial services compliance applications. Customers using CGE include Deloitte, Mayo Clinic, Noblis, ORNL, and many government agencies.

› **DataStax ramps up its DSE Graph to enable new distributed graph use cases.** DataStax, a key contributor to the popular open source Apache Cassandra project, offers the DSE Graph product. In 2015, DataStax recognized it was time to bring distributed graph technologies to the enterprise through DSE and acquired Aurelius, the creators of the open source graph database Titan. With DSE Graph, DataStax offers enterprises multimodel capabilities to store, process, and access various data formats and types in real time to support broader and emerging use cases for transactions and operational analytics. DSE Graph is available as a commercial software offering and through a subscription to DSE Max. Customers use DSE Graph for real-time customer 360, product catalogs and complex list management, real-time fraud detection, cybersecurity, asset management, internet-of-things (IoT) analytics, and personalization and recommendation engines. DSE Graph continues to gain momentum with customers who like its ability to scale out in a multi-data-center, multicloud, and hybrid environment as well as its visual operations, search, and advanced security functionality.

› **Open source FlockDB offers a good alternative to commercial solutions.** While FlockDB does not have visibility among enterprises, it's an open source distributed graph database that Twitter built to store large and complex relationships and later released to the community. Since no commercial vendor supports FlockDB, businesses should leverage it with caution, especially for use cases that might require 24x7 production support. We have found FlockDB to be very suitable when there is a team of developers who get acquainted with the code and customize it to support complex graph applications that commercial products don't offer. FlockDB is suitable for set operations requiring horizontal scalability and low-latency access to applications such as large social networking maps, complex routing, fraud detection, or system monitoring.

› **IBM Graph adds graph capabilities to address the growing market.** IBM Graph is a new cloud graph database offering that became generally available in August 2016. IBM Graph is a fully managed multitenant cloud service that enables users to create one or more property graphs in their respective instance. It utilizes TitanDB and Apache TinkerPop configured with a highly available Cassandra and ElasticSearch back end. IBM Graph supports eventually consistent transactions. IBM sells IBM Graph as an independent offering on IBM Bluemix only, provisioned on-demand to support Graph DBaaS scenarios.[7] Customers use IBM Graph for personalization and real-time recommendation engines, network and IT operations, investment portfolio management, forensics and fraud detection, MDM, and identity and access management. IBM also recently released a managed JanusGraph service on the Compose platform, which will be available on IBM Bluemix later this year.[8] It is TinkerPop compliant and provides facility for self-service and elastic resource scaling.

› **Neo4j continues to dominate the graph database market.** Neo4j released version 1.0 of the open source Neo4j property graph database in 2010 and since then has continued to deliver innovation and expand the product to support broader use cases. Today, Neo4j is used by many of the world's largest companies in industries such as technology, travel, financial services, retail, telecommunications, and government. The most common use cases include MDM, data lineage and customer 360, real-time recommendation engines, fraud detection, network and IT operations, and knowledge graphs. Neo4j offers a community edition that's wholly open source, while the enterprise edition includes some closed source add-ons. Customers that use Neo4j include Airbnb, Cisco Systems, eBay ShopBot, Hewlett Packard Enterprise, Lufthansa, Marriott, Orange, Telenor, Telia Zone, UBS, University of Washington, and Walmart. One company leverages Neo4j to process more than 1 billion transactions per day across 3 billion node graphs.

› **Oracle offers choices with RDF and property graph for on-premises and cloud.** Oracle has been offering RDF triple graph store as part of Oracle Database for more than a decade. It introduced a distributed property graph with parallel in-memory graph analytics for Apache HBase and Oracle NoSQL database in 2015 as well as a distributed property graph with parallel in-memory graph analytics for Oracle Database in 2016. These products are available on Database Cloud Service and Big Data Cloud Service, respectively. The Oracle property graph and RDF graph products are also embedded in a growing number of Oracle Cloud solutions, such as Adaptive Intelligent Apps and Social Network Cloud. Customers use Oracle RDF to support linked open-data publishing, semantic search, semantic metadata, MDM, and knowledge management systems, while property graph use cases include fraud detection, customer 360, recommendation engines, and blockchain analysis. Customers include Eli Lilly, European Union Publications Office, Galica Bank, Italian National Statistics Agency, Novartis Pharma, Roche Pharma, and Thompson Reuters.

› **OrientDB challenges Neo4j and other established graph players.** OrientDB is the key contributor and supporter of OrientDB, an open source NoSQL graph database written in Java that was first released in 2010. OrientDB supports both schemaless and schema-based modes and uses SQL to query structured and unstructured data on top of Gremlin and SPARQL. Customers often mention its multimodel engine, ease of use, reliable performance, and small footprint as core strengths.

OrientDB has a fully ACID-compliant graph database to support transactional and operational use cases. Key use cases for OrientDB include recommendation engines, real-time fraud detection, MDM, and social networks. OrientDB's impressive list of customers includes Accenture, Cisco Systems, CenturyLink, Comcast, FOX Sports, Ericsson, Sonatype, and Warner Bros.

› **MarkLogic offers a viable alternate when using moderate-sized graph applications.** MarkLogic uses RDF, but property graphs can be persisted using GraphSON or GraphML. It doesn't require a data model to get started. Customers typically load the data they have — in XML or JSON — and then build the graph over the documents. MarkLogic's graph queries are composable with all the other query types available — structure queries, search, and aggregates. Customers use MarkLogic graph to support MDM, operational data hubs, content discovery, data governance, cybersecurity, and fraud detection. It runs on AWS and Azure cloud for customers who want a graph platform that's easy to provision and manage. Customers include KPMG for its regulatory, tax, and compliance reporting; OECD to support an operational data hub for information sharing and knowledge creation; and Pearson for a digital educational platform that leverages MarkLogic with Semantics to connect data in course assessments and students' relationships to instructional content as a whole.
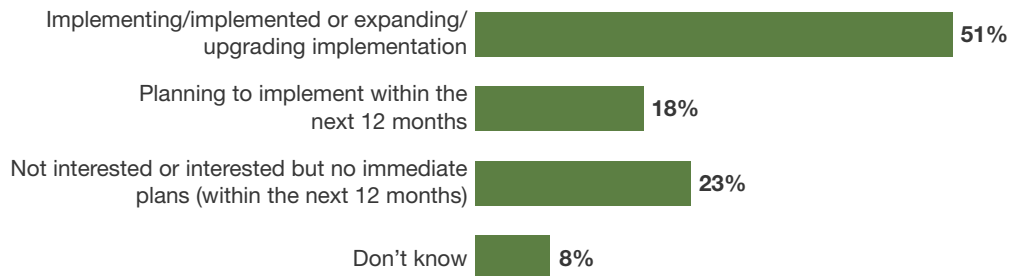
› **TIBCO Software has joined the game, promising innovation to support new use cases.** With a history of supporting graph-related use cases for customers since 2010 by leveraging various TIBCO products, the vendor released TIBCO Graph Database (TGDB) in 2016. Based on a property graph model, TGDB is derived from big data and relationship management problems that require transactional integrity. TGDB is currently offered as an independent product in either open source or commercialized form; the code is available on GitHub. The top use cases for TGDB include geospatial/routing/matching, finding the shortest route, MDM, the IoT, targeted recommendations based on customer data, and AI applications. TGDB is currently not supported in the public cloud, but this support is likely in the near future. The product supports a subset of SQL, which allows applications to filter the desired list of nodes and edges. In addition, an open API is available for customers and partners wanting flexibility in accessing graph data, and the Cypher query language will likely be supported in a future release.

FIGURE 4 Most Organizations Are Starting To Leverage Graph Database

"What are your firm's plans to use the following data management technologies?"
(Graph databases)

Implementing/implemented or expanding/ upgrading implementation	51%
Planning to implement within the next 12 months	18%
Not interested or interested but no immediate plans (within the next 12 months)	23%
Don't know	8%

Base: 2,106 global data and analytics technology decision makers

Source: Forrester Data Global Business Technographics® Data And Analytics Survey, 2017

FIGURE 5 Overview Of Graph Database Vendors

Vendor	Initial release	Graph model	Cloud support	Query languages supported	Open source/ commercial/ both	Standalone product/ bundled with solution
ArangoDB	2011	Property	Amazon Web Services (AWS), Azure, and DigitalOcean	ArangoDB Query Language, Gremlin	Both	Standalone, but part of multimodel database (DB) offering
Cray	2016	RDF	Yes (Urika-GX offering includes Cray Graph Engine for Private Cloud)	SPARQL	Commercial	Standalone on Cray XC computers and bundled with big data packages with Urika-GX platform
DataStax	2016	Property	Cloud enabled; soon to be available as a managed service	Gremlin, DseGraphFrames, SparkSQL	Commercial	Standalone; can be purchased as an add-on to existing cluster or as a bundle with DSE Max
FLockDB	2010	Property	None	Java	Open source	Standalone
IBM Graph	2016	Property	IBM Bluemix	Gremlin	Commercial but built on TinkerPop and Titan	Standalone in the cloud
Microsoft	2017	Property	Azure Service	Gremlin and APIs	Commercial	Graph is part of CosmosDB (multimodel DB)

FIGURE 5 Overview Of Graph Database Vendors (Cont.)

Vendor	Initial release	Graph model	Cloud support	Query languages supported	Open source/ commercial/ both	Standalone product/ bundled with solution
MarkLogic	2016	RDF	AWS and Azure Marketplace	SPARQL	Commercial	Graph is part of the MarkLogic Server
Neo4j	2007	Property	AWS Marketplace and Azure	Cypher, Gremlin, Procedures in Java	Both	Standalone product
Oracle	2005	Property and RDF	Part of Oracle Cloud Database offering	RDF: SPARQL and PL/SQL Property: PGQL, SparkSQL, Gremlin, and PL/SQL	Commercial	Standalone and bundled solutions with Oracle Database and Hbase/ NoSQL
Orient DB	2010	Property	AWS, Azure Marketplace, Century Link, Digital Ocean, and Jelastic	Gremlin, GraphQL, SParkQL, and SQL	Both	Standalone solution
TIBCO Software	2016	Property	None	Subset of SQL, APIs, Cypher (planned)	Both	Standalone

Recommendations

Invest In Graph Database To Succeed In Your Digital Business Strategy

Today's enterprises are spending more time looking for answers to complex business questions: Why is there an increase in customer churn in Chicago? Is a customer claim fraudulent? What is the next best product a customer will buy from us? What are the risks involved in giving a loan or credit to a consumer? A graph database allows organizations to think differently and create new intelligence-based business opportunities that would otherwise be difficult to develop. The bottom line? Enterprise architects who investment in graph database will be ready to use connected data quickly to create new actionable insights, respond quickly to competitive threats, and help their firms grow faster than their competitors by delivering new and innovative products and services. You should:

› **Invest in graph technology for connected data.** Without graph database technology, it can be painfully slow to build and deploy analytics, personalization, and next-best-offer analysis, causing you to miss an opportunity to sell products and services to a customer at the right time. Consider a graph database to speed your most critical processing for delivering breakthrough customer experiences.

› **Target the right data set for graph.** Not all business data is suitable for graph because the data may not be connectable or may be more optimized in a different data format, such as JSON. Do some initial discovery to find out if data patterns could be linked to create insights, that'll be a good starting point to leverage graph.

› **Use graph for all kinds of operational intelligence and predictive analytics.** Graph can support more than just simple reports that require connected data; it can also support complex requirements such as fraud detection, risk analytics, operational intelligence, social networks, complex data search, IoT analytics, recommendation engines, and advanced analytics.

› **Train your existing staff to leverage graph.** Graph technology requires different skills, especially when setting up the database and server, models, and queries that exploit graph. Training your existing staff is critical to succeed with graph.

› **Ensure that data security is nailed down in the graph platform.** While most graph technologies offer security and governance as part of the platform, you must nail down data security from the start, especially when dealing with sensitive data such as credit card information or healthcare-related data.

Engage With An Analyst

Gain greater confidence in your decisions by working with Forrester thought leaders to apply our research to your specific business and technology initiatives.

Analyst Inquiry

To help you put research into practice, connect with an analyst to discuss your questions in a 30-minute phone session — or opt for a response via email.

Learn more.

Analyst Advisory

Translate research into action by working with an analyst on a specific engagement in the form of custom strategy sessions, workshops, or speeches.

Learn more.

Webinar

Join our online sessions on the latest research affecting your business. Each call includes analyst Q&A and slides and is available on-demand.

Learn more.

Forrester's research apps for iOS and Android.
Stay ahead of your competition no matter where you are.

Supplemental Material

Survey Methodology

The Forrester Data Global Business Technographics® Data And Analytics Survey, 2017 was fielded between February and April 2017. This online survey included 3,378 respondents in Australia, Brazil, Canada, China, France, Germany, India, New Zealand, the UK, and the US from companies with 100 or more employees.

Forrester Data Business Technographics ensures that the final survey population contains only those with significant involvement in the planning, funding, and purchasing of business and technology products and services. Research Now fielded this survey on behalf of Forrester. Survey respondent incentives include points redeemable for gift certificates.

Please note that the brand questions included in this survey should not be used to measure market share. The purpose of Forrester Data Business Technographics brand questions is to show usage of a brand by a specific target audience at one point in time.

For Technographics Clients: How To Get More Technographics Data Insights

Forrester Data Global Business Technographics Data And Analytics Survey, 2017 of 3,378 global data and analytics decision makers includes many additional questions and parameters by which you can analyze the data contained in this report.

We can provide additional insights about the consumers highlighted in this report:

› Who they are (e.g., demographics, lifestyle, and interests).

› What they do (e.g., digital, mobile, social behaviors).

› Affiliations they have (e.g., brands used, products owned).

› How they feel (e.g., attitudes, interests).

Endnotes

[1] Eighty-five percent of global business and technology decision makers and influencers consider developing or acquiring new skills as at least a middle priority initiative for their IT organizations over the next 12 months. Source: Forrester Data Global Business Technographics Priorities And Journey Survey, 2017.

[2] Fifty-nine percent of global data and analytics decision makers say that they are implementing, have implemented, or are expanding their use of big data technology to support customer experience analytics. Source: Forrester Data Global Business Technographics Data And Analytics Survey, 2017.

[3] Fourteen percent of global data and analytics decision makers stated that "lack of end user self-service capabilities" is one of the biggest challenges their firm is facing when orchestrating its business intelligence strategy. Source: Forrester Data Global Business Technographics Data And Analytics Survey, 2017.

[4] Source: "Neo4j at Walmart," Neo4j (http://neo4j.com/users/walmart/).

[5] See the Forrester report "Brief: The MDM Crossroad — To Graph Or Not To Graph?"

[6] TB: terabyte; HA: high availability.

[7] DBaaS: database-as-a-service.

[8] JanusGraph is the open-source fork of TitanDB. IBM is a founding member of the project under The Linux Foundation with Google, Hortonworks, Expero, and GRAKN.AI.

We work with business and technology leaders to develop customer-obsessed strategies that drive growth.

PRODUCTS AND SERVICES

› Core research and tools
› Data and analytics
› Peer collaboration
› Analyst engagement
› Consulting
› Events

Forrester's research and insights are tailored to your role and critical business initiatives.

ROLES WE SERVE

Marketing & Strategy Professionals	**Technology Management Professionals**	**Technology Industry Professionals**
CMO	CIO	Analyst Relations
B2B Marketing	Application Development & Delivery	
B2C Marketing	› Enterprise Architecture	
Customer Experience	Infrastructure & Operations	
Customer Insights	Security & Risk	
eBusiness & Channel Strategy	Sourcing & Vendor Management	

CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.