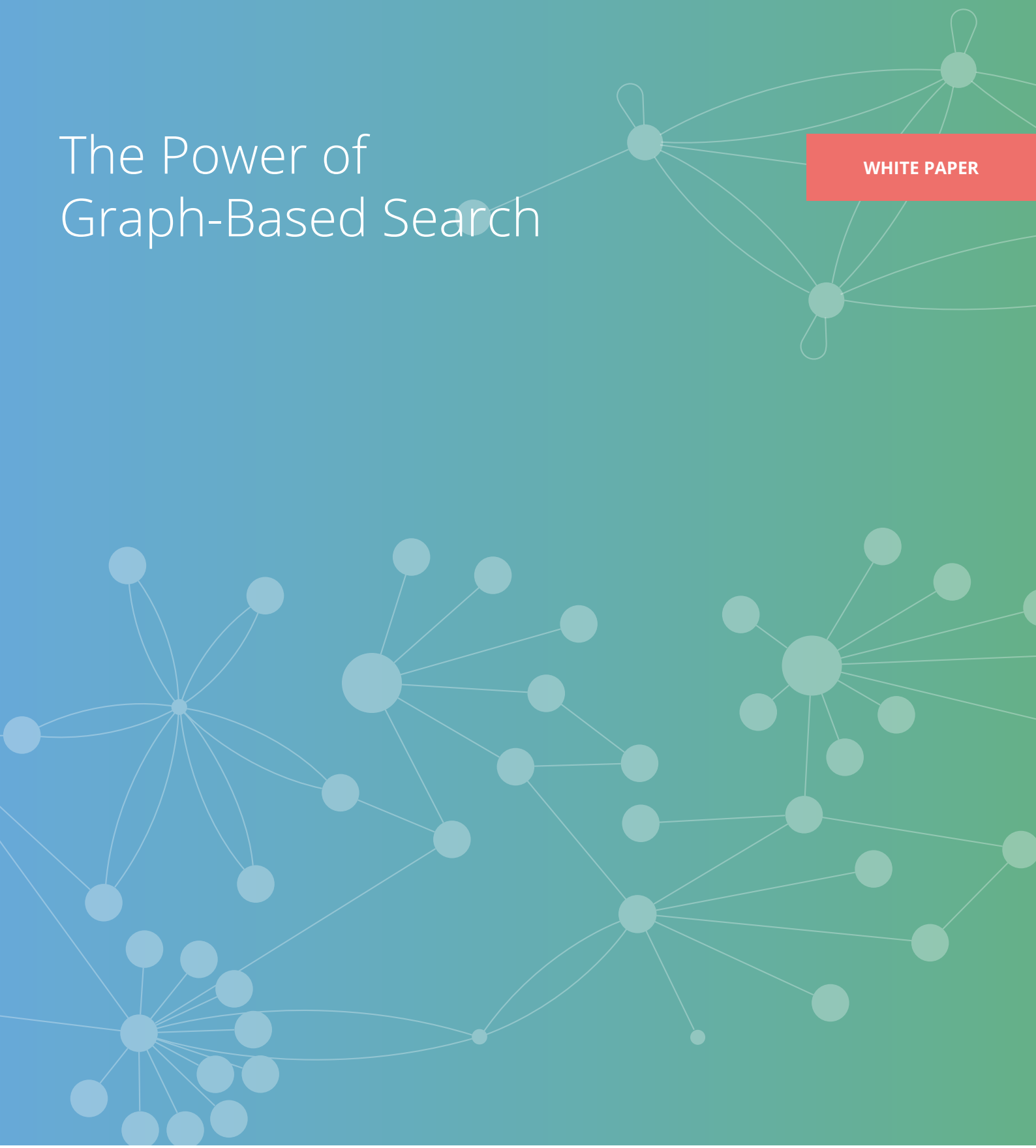
neo4j

The #1 Platform for Connected Data

The Power of Graph-Based Search

WHITE PAPER



WHITE PAPER

TABLE OF CONTENTS

The Power of Graph-Based Search

APRIL 2015

Made famous by Facebook and Google, 'graph-based search' is providing major business benefits for a growing number of Neo4j's corporate customers.

Facebook and Google are two of the most renowned businesses in history, and rightly so: with more than 1 billion users,[1] Facebook has been hugely successful in its mission to help people find other people. And with more than 1 billion monthly visitors,[2] Google has been just as successful in its mission to help people find information.

Facebook's database of people and Google's database of information have one crucial thing in common: they were both built using graph technology – providing a major endorsement for Neo, the world's most successful provider of corporate graph databases.

And in recent years, both Google and Facebook have realized they could make much better use of their huge swathes of searchable content, and have each launched new services built round 'graph-based search' to exploit these commercial opportunities.

This in turn has highlighted the business benefits that Neo4j's own graph-based search and discovery features offer to corporate customers.

So what exactly is graph-based search, and what can it do for you?

The Rise of Graph-Based Search

In their early days, both Facebook and Google offered users basic ways to access their data – the standard approach of 'listing' or 'keyword' search, where you type in a word or phrase you're interested in and get back a list of all the web pages and documents that include said keyword.

This method is essentially plain pattern recognition – and many people will be familiar with the cumbersome process of repeatedly redefining your search terms until you finally hit on something of interest.

1. http://newsroom.fb.com/news/2013/01/introducing-graph-search-beta/
2. http://searchenginewatch.com/sew/news/2081332/google-hits-billion-monthly-unique-visitors-mark

"If we were to do that in any other system it would be impossible. Thanks to graph technology, we can easily add new datasets and build new experiences around them.

By having everything in a graph database, we don't have to know in advance what kinds of questions we're going to want to ask. It's the challenge of answering the questions that come up ad hoc where Neo4j really, really shines."

Kurt Freytag,
Head of Product,
CrunchBase

But realizing the limitations of this, in 2012 Google launched 'Knowledge Graph'[3] and in 2013 Facebook followed suit with its 'Graph Search' service[4], both of which provide a better way to search data by helping users with more contextual information.

Knowledge Graph is a database that enhances Google's search engine results with semanticsearch information gathered from a wide variety of sources.

Likewise, Facebook's Graph Search enables users to combine search phrases to get more structured and more localized search results, rather than simply using standard keywords and getting the results that match those words.

The key to their enhanced search capability is that the first search takes into account the entire structure of connected data available. And because graph systems understand the way data is related, they return much more precise and richer results.

Graph-based search, powered by systems like Neo4j, is able to deliver relevant information that you may not have specifically asked for – offering a more proactive and targeted search experience that allows you to quickly triangulate onto the data points that are of the most interest to you.

In essence, graph-based search is intelligent: you can ask much more precise and useful questions and get back the most relevant and meaningful information, whereas traditional keyword-based search delivers results that are more random, diluted and lower-quality.

Graph-based search is also much quicker: you can query all of your connected data in real time, then hone in on the answers provided and launch new real-time searches prompted by the insights you have discovered.

It's more of a natural 'conversation' with your data, rather than a series of one-off searches. It's search and discovery, rather than search and retrieval.

Graph-Based Search in Action

But if Neo4j's graph-based search and discovery is superior in nature, what actual business benefits does this provide?

The experience of Neo4j users shows that it has allowed them and their customers to discover new business insights and launch new products and services:

CrunchBase

Take Neo4j user CrunchBase. Launched in 2007, CrunchBase offers profiles of over 650,000 high-tech companies, executives and investors and all the activities connecting them, including funding, job changes, product launches, successes and failures.

This data is used monthly by millions of consumers, analysts, investors, venture capitalists, recruiters and job seekers – and one of CrunchBase's unique selling points is that Neo4j allows it to add any new data and handle unpredictable searches on it.

3. http://www.google.co.uk/insidesearch/features/search/knowledge.html
4. http://newsroom.fb.com/news/2013/02/graph-search-some-favorite-searches/

"You never know what people are going to want to know, so you have to be able to repackage the information in different ways on-thefly. The spontaneity of it and the ability to repackage things on-the-fly – we can do that kind of thing with Neo4j. People want to do their own discovery.

Graph-based search makes things possible that we only dreamed of before. We can attempt some previously very very difficult things without breaking a sweat."

Evan Stein,
Founder,
Decibel

Head of Product, Kurt Freytag, explained: "Neo4j enabled us to take our foundation of data and extend it much further, to do all sorts of interesting business intelligence, data analytics and expose all this information in much more interesting ways on the front-end to users."

This has resulted in CrunchBase launching new products and services, such as 'Graph Insights', which enables site visitors to ask complex questions about any organization and get back answers in real time – a situation where traditional relational and schema-less databases would grind to a halt.

"If we were to do that in any other system it would be impossible," Kurt said. "Thanks to graph technology, we can easily add new datasets and build new experiences around them.

"By having everything in a graph database, we don't have to know in advance what kinds of questions we're going to want to ask. It's the challenge of answering the questions that come up ad hoc where Neo4j really, really shines."

Decibel

Another Neo4j user, Decibel, has achieved similar business benefits by providing data that can be repeatedly and instantly queried in unpredicted ways. Company founder Evan Stein calls this 'conversationality'.

London-based Decibel caters for music lovers, offering detailed data on over 1 million albums of every type of music, including the tracks, composers, musicians, what else they've recorded, where and when, who else recorded those tracks and a myriad more other information.

Evan wanted to expose this "thicket" of inter-connected data in such a way that it could be quickly searched by customers in any way they chose.

"You never know what people are going to want to know, so you have to be able to repackage the information in different ways on-the-fly," he said. "The spontaneity of it and the ability to repackage things on-the-fly – we can do that kind of thing with Neo4j. People want to do their own discovery."

Evan explained: "Neo lets us set up the data is a way that's really natural. The primary benefit is 'conversationality' – because in the world of computers, you ask certain questions and the engineer says 'Bah no, can't do it, I'll have to batch it overnight and give you a report'.

"Well, there's no freedom in that. What you get with Neo4j is freedom – the ability to be conversational about what you're doing, real discovery.

"People go from idea to idea. It's not necessarily linear. If the answer comes back and you have the freedom to follow your train of thought, then it's much more natural, it's much more fun and you're not just playing with data. It's as close to thinking as you can get."

Crucially, Decibel found that Neo4j could deliver this service at acceptable speeds, unlike conventional databases.

"The real difference compared to a relational database is we can give people an answer within their lifetime – because sometimes you'd have to wait a full minute or two minutes, and two minutes in the world of computers is an eternity, so that's just as bad as not answering them," Evan said.

"Graph-based search makes things possible that we only dreamed of before. We can attempt some previously very very difficult things without breaking a sweat."

BUSINESS BENEFITS OF NEO4J

Neo4j makes possible advance search-and-discovery because:

1. Companies can structure their data exactly as it occurs and carry out searches based on this structure. Neo4j provides a database model and language that supports this.

2. Users can get fast, accurate search results in real-time. Neo4j assigns a variety of rich metadata to content for rapid search and retrieval.

3. Companies can easily change their data and its structure, and add a wide variety of new data. The builtin flexibility of Neo4j's data model allows this.

Conclusion

For businesses that have huge volumes of products or content, graph-based search provides a better way to make this data available to users, as Google's Knowledge Graph and Facebook's Graph Search have demonstrated.

Graph-based search offers numerous competitive advantages – including better customer experience, more targeted content, and increased revenue opportunities. And there are many areas where it is valuable, including customer support portals, content portals, product catalogs and social networks.

Now businesses seeking graph-based search capabilities can get them off-the-shelf from Neo4j, the world's most widely used graph database.

Neo4j makes possible advanced search-and-discovery because:

1. Companies can structure their data exactly as it occurs and carry out searches based on this structure. Neo4j provides a database model and language that supports this.

2. Users can get fast, accurate search results in real-time. Neo4j assigns a variety of rich metadata to content for rapid search and retrieval.

3. Companies can easily change their data and its structure, and add a wide variety of new data. The built-in flexibility of Neo4j's data model allows this.

In contrast, information held in a relational database is much more inflexible to future change: if you want to add new kinds of content or make structural changes, you are forced to re-work the relational model in a way that you don't need to do with the graph model. The graph model is much more easily extensible.

Companies have adopted Neo4j because it's 1,000 times faster than relational databases for working with connected data. And its graph-based search features have enabled firms such as CrunchBase and Decibel – and many others – to discover new business insights, launch new products and services, and attract new customers.

It's a transformational discovery that these companies have made: the business advantage of using Neo4j's graph-based search and discovery.

References

1. http://newsroom.fb.com/news/2013/01/introducing-graph-search-beta/
2. http://searchenginewatch.com/sew/news/2081332/google-hits-billionmonthly-unique-visitors-mark
3. http://www.google.co.uk/insidesearch/features/search/knowledge.html
4. http://newsroom.fb.com/news/2013/02/graph-search-some-favoritesearches/

About Neo4j

Neo4j is an internet-scale, native graph database that leverages connected data to help companies build intelligent applications that meet today's evolving challenges including machine learning and artificial intelligence, fraud detection, real-time recommendations and master data. As the #1 Platform for connected data, Neo4j has over three million downloads, the world's largest graph developer community, and over thousands of graph-powered applications in production.

The world's most sophisticated organizations worldwide, from enterprises like Walmart, eBay, UBS, Cisco, HP, adidas and Lufthansa to hot startups like Medium, Musimap and Glowbl, use Neo4j to harness the connections in their data.

United Kingdom uk@neo4j.com
France ventes@neo4j.com
Scandinavia nordics@neo4j.com
DACH vertrieb@neo4j.com
Southern Europe southern-europe@neo4j.com