

# **Graph-DB: Beziehungen visualisieren, Informationen entdecken**

**Frank Ferro  
Trivadis GmbH  
Frankfurt/Main**

## **Schlüsselworte**

Graph-Datenbank, Big Data, Oracle Spatial and Graph, Visualisierung, Panama-Papers, Proof-of-Value, Prototyp, Einsatzgebiete

## **Einleitung**

In der IT dreht sich immer mehr um Daten. Spätestens seit Begriffe wie Big Data, Data Scientist etc. auch im Umfeld von Fachbereichen Einzug gehalten haben, sind Daten in aller Munde. Sogar Aussagen wie „Daten sind das heutige Öl“ werden oft zitiert. Doch weit wichtiger wie Daten sind für Unternehmen Informationen. Ein beträchtlicher Anteil der tagtäglich verarbeiteten Informationen ist nicht wegen des Umfangs interessant, sondern wegen der inhärenten Verknüpfungen, die darin enthalten sind. Die Panama Papers sind der beste Beweis.

## **Mehrwert durch dne Einsatz von Graph-Datenbanken**

In den Datenbanken sind in erster Linie Daten abgelegt. Die meisten Unternehmens-Anwendungen funktionieren nach dem gleichen Prinzip: Es wird gezielt nach Datensätzen wie Kunden, Bestellungen, Rechnungen gesucht. Dazu werden die Entitäten aus der Datenbank in Oberflächen zur Anzeige bzw. Bearbeitung angezeigt. Meist sind dazu auch bestimmte Sachbearbeiter für bestimmte Kunden zuständig. Somit kann der Sachbearbeiter sich gezielt auf die Daten konzentrieren. Jeder Sachbearbeiter hat damit aber auch einen ganz eingeschränkten View auf die Kundenwelt. Zusammenhänge können somit leider nicht erkannt werden. Soziale Netzwerke haben dahingehend einen ganz anderen Ansatz. Hier steht im Vordergrund wer mit wem wie vernetzt ist. Erst dann werden Daten zu einer Person angezeigt. Hier dominieren ganz klar die Beziehungen.

Damit wird klar, daß ein beträchtlicher Anteil der tagtäglich verarbeiteten Informationen nicht wegen des Umfangs interessant ist sondern wegen der inhärenten Verknüpfungen, die darin enthalten sind. Diese Verknüpfungen machen den eigentlichen Wert solcher Informationen aus. Wenn vernetzte Informationen in relationalen Datenbanken gespeichert werden sollen, werden die Verbindungen normalerweise ignoriert, denormalisiert oder zusammengefasst, um dem Datenmodell zu genügen. Dabei geht die Informationsfülle verloren, die in anderen Datenbanken und Datenmodellen erhalten wird. Graphdatenbanken und das Graphdatenmodell sind insbesondere dazu geeignet, die wertvollen inhärenten Verknüpfungen zwischen Informationen abzubilden. Will man z. B. in relationalen Datenbanken die Personendetails von in Beziehung stehenden Personen auswerten, so bedarf es einem Join der Tabelle auf sich selbst. Bei Navigation über mehrere Beziehungsknoten muß der Join entsprechend oft wiederholt werden. Das macht sich bereits ab einer Navigationstiefe von 3 zeitlich deutlich bemerkbar.

Die Präsentation graphbasierter Repräsentation ist ein wichtiger Aspekt um die angesprochene Informationsfülle effektiv und intuitiv nutzbar zu machen. Hier liegt es nahe eine grafische Darstellung mittels Graphen einzusetzen. Ein entscheidender Vorteil solcher graphbasierter Visualisierung besteht in der Ausdrucksstärke. Lineare, textuelle Präsentationen von nichtlinearen Strukturen stellen eine zusätzliche Abstraktionsebene dar. Diese Abstraktion kann dem Benutzer durch Verwendung grafischer und insbesondere graphbasierter Darstellungen erspart werden, wodurch sich

mehr Informationen im gleichen Zeitraum übermitteln lassen. Darüber hinaus sind bildhafte Informationen weitgehend sprachunabhängig. Komplexe, vernetzte Strukturen sind daher von uns Menschen in bildhafter Form leichter zu erkennen, leichter zu interpretieren, leichter zu bearbeiten, und auch leichter im Gedächtnis zu behalten. Bilder bzw. Grafiken ermöglichen uns Menschen schon seit Langem aus Informationen schnell und einfach, wertvolles Wissen zu gewinnen. Bereits vor Jahrhunderten wurden Stammbäume erstellt. Ein Stammbaum ist dabei nichts anderes als ein Graph. Mit Aufkommen der IT wurde der Mensch „umgeschult“. Spätestens mit Excel haben viele Menschen nur noch in Spalten und Zeilen (also relational) gedacht. Die Graph-Datenbank unterstützt wieder das „normale“ Denken.

Werfen wir beispielsweise einen Blick auf die Finanzindustrie: Eine natürlich Person hat einen Bevollmächtigten. Die Vollmacht selbst hat wieder Attribute (gültig seit, Betrags-Einschränkung, ect). Eine Organisation hat Geschäftsführer, hat Vertretungsberechtigte, hat wirtschaftlich Berechtigte. Eine Überweisung ist eine Transaktion von einem Kunden zu einem anderen Kunden. Egal wo wir auch hinschauen, überall sind Beziehungen im Spiel und kann als Graph betrachtet werden. Und insbesondere die Navigation im Graphen ist kindelreich. Algorithmen wie kürzester Weg von einem Knoten zu einem anderen Knoten oder gibt es eine Verbindung zwischen zwei Knoten und wenn ja über wie viele Zwischenknoten sind in den Abfragesprachen einfach vorhanden.

Mit Hilfe von Graph-Datenbanken rückt plötzlich die Beziehung in den Vordergrund. Und meist sind mit der Beziehung auch weitere Attribute abgelegt. Wie Kunden untereinander vernetzt sind, ist aus mehreren Gründen interessant:

- Sind hochzufriedene Kunden miteinander vernetzt?
- Wie stark ist ein Kunde vernetzt, welcher eine Beschwerde eingereicht hat?
- Wie ähnlich ist die Produktnutzung bei vernetzten Kunden?
- Wie stark sind hochrisiko Kunden (aus Sicht Anti Financial Crime) miteinander vernetzt?
- Wie stark ist die Vernetzung von z. B. Bevollmächtigten, die selbst nicht Kunden sind, mit Kunden?
- Wie viele Betreuer haben vernetzte Kunden?

Gemeinsam mit einem großen Bankinstitut haben wir einen Proof-of-value durchgeführt, um anhand der Kundendaten den Vernetzungsgrad zu visualisieren. Mit diversen Fachbereichen diskutieren wir gemeinsam die unterschiedlichen Einsatzgebiete und die damit verbundenen Mehrwerte.

Erfahren Sie wie und wo sich relationale Datenbanken und Graphen-Datenbanken sinnvoll ergänzen lassen. Und wir sprechen hier ganz klar von einer Ergänzung und keiner Ablösung.

Ebenso stellen wir den entstandenen Prototypen vor, der einen ersten Eindruck über die Visualisierungsmöglichkeiten als auch über analytischen Auswertungsmöglichkeiten bietet.

**Kontaktadresse:**

Frank Ferro  
Trivadis GmbH  
Lyoner Str. 15  
D-60528 Frankfurt/Main

Telefon: +49 (0) 162 29 10 217  
E-Mail frank.ferro@trivadis.com  
Internet: www.trivadis.com