

## CLINICAL KNOWLEDGE GRAPH

## Struktur fürs Datenchaos

Krankenhäuser generieren täglich riesige Datenmengen – von Diagnosen über Laborwerte bis hin zu Forschungsdaten. Graphdatenbanken helfen, diese Informationen intelligent zu vernetzen und so Zusammenhänge sichtbar und nutzbar zu machen.

Krankenhäuser stehen vor der Herausforderung, immer größere Datenmengen aus Medizin, Verwaltung und Forschung sinnvoll zu nutzen. Eine Graphdatenbank kann hier helfen, denn sie speichert Daten nicht als Tabellen, sondern als Netzwerk: Jeder Punkt (Knoten) steht für eine Information – etwa einen Patienten, einen Laborwert oder ein Medikament – und die Linien dazwischen (Kanten) zeigen, wie diese Informationen miteinander verbunden sind. Dadurch lassen sich Beziehungen in den komplexen Daten besonders schnell erkennen und abfragen.

Doch eine Graphdatenbank allein liefert noch kein Wissen, sie bildet nur die technische Grundlage. Erst, wenn die Daten inhaltlich miteinander verknüpft und in einen verständlichen Zusammenhang gebracht werden, entsteht daraus ein sogenannter Knowledge Graph. Dieser verbindet die gespeicherten Informationen zu einem Wissensnetz, in dem nicht nur Verbindungen sichtbar sind, sondern auch deren Bedeutung. Durch die Kombination von Graphtechnologie, künstlicher Intelligenz (KI) und maschinellem Lernen (ML) entstehen intelligente Wissenssysteme, die Management-Entscheidungen beschleunigen, Nachvollziehbarkeit ermöglichen, Prozesse effizienter machen und die Qualität der Versorgung verbessern können.

„Graphdatenbanken bringen große, unterschiedliche Datenquellen in eine gemeinsame Struktur. Der Knowledge Graph geht darüber hinaus, weil er aus diesen Verknüpfungen neue Erkenntnisse gewinnt und Muster sichtbar macht, die



„ **Klassische Datenbanksysteme stoßen an ihre Grenzen, wenn es darum geht, Beziehungen in großen Datenmengen schnell zu erkennen.**

vorher nicht bekannt waren“, erklärt Dr. Alexander Jarasch, Global Head of Pharma and Life Sciences bei Neo4j, das Prinzip. Die Neo4j-Graphdatenbank ist nativ, das bedeutet, dass sie Entitäten (z. B. Patienten oder Befunde) mitsamt

zusätzlichen Informationen und Beziehungen als Graphstrukturen speichert. Dadurch sind extrem schnelle Abfragen von Daten und deren Beziehungen möglich. Nicht native Graphdatenbanken legen die Informationen in Tabellen oder Dokumenten ab, wodurch Beziehungen erst zusammengeführt werden müssen. Das verlangsamt die Abfrage deutlich.

### Seltene Erkrankungen erkennen

Speziell für den klinischen Kontext wurde am Max-Planck-Institut für Biochemie in Neo4j ein Clinical Knowledge Graph (CKG) entwickelt. Er verbindet klinische und biologische Daten intelligent, um bessere Entscheidungen in Diagnostik, Therapie und Forschung zu ermöglichen. Der CKG kommt zum Beispiel im Dr. von Haunerschen Kinderspital der

© Chaosman\_Studio / stock.adobe.com

## Traditionelle Datenbanken versus Graphdatenbanken

Traditionelle Datenbanksysteme speichern Daten häufig in voneinander getrennten Silos in Form von Tabellen. Am Beispiel von Patientendaten würde dies bedeuten, dass Laborwerte, Diagnosen und verordnete Medikamente in separaten Datensätzen mit isolierten Einzelinformationen vorliegen. Ein Gesamtbild und die entscheidenden Zusammenhänge – etwa, dass ein bestimmter Patient mit einer spezifischen Diagnose ein entsprechendes Medikament erhält und relevante Blutwerte aufweist – können erst durch aufwendige Verknüpfungen hergestellt werden.

Graphdatenbanken verfolgen einen anderen Ansatz. Sie sind darauf spezialisiert, nicht nur die Daten selbst, sondern auch die Beziehungen zwischen ihnen als gleichwertige Informationseinheiten zu speichern und zu verwalten. Auf diese Weise entsteht eine vernetzte Datenstruktur, ein sogenannter Knowledge Graph (Wissensgraph), in dem der Kontext der Daten von Anfang an Bestandteil des Modells ist.

## „Mit Graphdatenbanken lassen sich medizinische und organisatorische Prozesse neu denken.“

Ludwig-Maximilians-Universität München – LMU zum Einsatz, um seltene Erkrankungen zu identifizieren. In der Kinderklinik werden dafür genomische

Daten, Blut- und Testwerte, individuelle Krankheitsverläufe, wissenschaftliche Fachliteratur und Arzneimittelinformationen im Clinical Knowledge Graph verknüpft. Jeder Patient ist als Knoten dargestellt, der mit anderen Knoten verbunden ist. Auf der Suche nach Verbindungen zwischen Proteinen, Genen und klinischen Symptomen navigieren Bioinformatiker in rund 16 Millionen Knoten und 220 Millionen Beziehungen. Gefördert wird das Projekt durch die Care-for-Rare-Foundation, die sich für Kinder mit seltenen Erkrankungen einsetzt und die Diagnosestellung sowie innovative Therapien unterstützt.

„Mit der klassischen Gen-Sequenzierung kämen wir aktuell nur bei 30 Prozent der Kinder zu einer Lösung“, erklärt Daniel Weiss, Projektleiter am Kinderspital. Da das Team die Daten von sehr vielen Kindern aus ganz Europa auswertet, die sich alle nur in wenigen Punkten unterscheiden, wird sprichwörtlich nach der Nadel im Heuhaufen gesucht. Dadurch, dass in dem Graph jedoch alle Informationen direkt miteinander verbunden sind, gibt es deutlich weniger Abfragen auf Datenbankebene als bei Anbietern, die die Daten nur miteinander vergleichen, statt sie zu verknüpfen. „Somit liegt das Ergebnis innerhalb von Millisekunden vor und wir haben sofort Zugriff auf alle weiteren Informationen zu der Erkrankung. Manchmal gibt es nur zwei Kinder, die identische Merkmale aufweisen, und nur aufgrund der Auswertung können wir diesen Kindern helfen oder vielleicht sogar ihr Leben retten“, zeigt sich Weiss begeistert.

### Einsatz in allen Krankenhausbereichen möglich

Ein Knowledge Graph kann grundsätzlich in allen Bereichen einer Klinik eingesetzt werden. So nutzt ein New Yorker Krankenhaus Graphtechnologie, um die Ausbreitung von Infektionen zu verfolgen. Die Ärzte erkennen im Knowledge Graphen, dass Patient A zu dem Zeitpunkt der Ansteckung im selben Raum wie Patient B war, welche Patienten sich in der Nähe aufhielten und welche mit

denselben Ärzten und Pflegekräften in Kontakt standen. Diese Einblicke ermöglichen tiefgehende Analysen dazu, wie Krankheiten sich ausbreiten und wer akut gefährdet sein könnte.

Ein Knowledge Graph eignet sich aber auch, um die MRT-Auslastung zu verbessern, indem er Engpässe erkennt und alternative Zeiten oder Geräte empfiehlt. Auch Belegungspläne lassen sich optimieren, wenn Aufnahmen, Entlassungen, Behandlungsdauer und Zimmerkapazitäten verknüpft werden. Denkbar wäre auch ein Einsatz in der Material- und Medikamentenlogistik. So könnte der Graph erkennen, dass Verbandmaterial knapp wird, und automatisch eine Bestellung auslösen. „Möglich ist eigentlich alles“, erklärt Jarasch.

### Internes LLM

Die Programmierung des Clinical Knowledge Graphs nach den Wünschen des jeweiligen Anwenders und das Einpflegen der Daten übernehmen am Haunerschen Kinderspital Bioinformatiker. Auch um den Abruf der Informationen kümmern sich derzeit noch die Informatik-Experten. Ab 2026 sollen die Münchner Ärzte dies – ohne zeitlichen Mehraufwand – jederzeit per Spracheingabe oder Textchat selbst tun können, auch wenn sie nicht besonders technikaffin sind. Möglich wird das durch ein Large Language Model (LLM), das nicht nur mit Daten aus dem Internet trainiert wurde, sondern auch Zugriff auf interne Daten aus den jeweiligen Krankenhäusern hat. „In unserem Fall wird das LLM sozusagen vor den Clinical Knowledge Graph geschaltet und übernimmt die Vermittlung zwischen den Ärzten und dem Graphen“, erläutert Weiss. Der Fachbegriff für die Zusammenarbeit des LLMs mit dem Graphen lautet Graph Retrieval Augmented Generation (GraphRAG). Das bedeutet, dass die Antworten des LLM mit den Informationen des Graphen angereichert werden.

„Ein öffentliches LLM wie zum Beispiel das, auf dessen Basis ChatGPT funktioniert, hat selbstverständlich keinen

Zugriff auf individuelle Patientendaten. Aber genau diese Daten möchten Ärzte oder Forschende mit natürlicher Sprache abfragen. Deshalb haben wir für das Kinderspital einen Chat auf Basis eines LLMs entwickelt, das sicher und datenschutzkonform auf den Krankenhausservern läuft“, erklärt Jarasch. „Das können wir für jedes andere Krankenhaus ebenso umsetzen.“

### Daten während des Patientengesprächs abrufen

Durch das interne LLM und die Chatfunktion können die Ärzte in Zukunft bereits während eines Patientengesprächs Auffälligkeiten in den Daten erkennen. „Ein Arzt kann Diagnosen, Medikamente, Laborwerte und vieles mehr direkt eingeben und bekommt sofort angezeigt, ob es ähnliche oder identische Fälle, Wechselwirkungen mit anderen Medikamenten oder neue Studien gibt. Daraus kann er dann eine individuelle Therapie ableiten“, so Jarasch.

Die Datenbank von Neo4j ist ein klassisches Lizenzmodell, das ein Unternehmen entweder selbst verwaltet oder aber in der voll gemanagten Cloud des Anbieters nutzt. Die Datenbank kann zudem auf jedem beliebigen Endgerät laufen, egal ob auf einem Computer oder Laptop. ■

Sonja Buske  
Freie Journalistin

