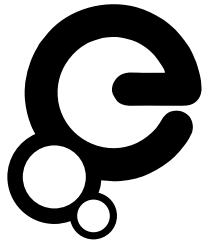


FALLSTUDIE



Europeana

Neo4j macht Millionen europäischer Kunstwerke und kultureller Objekte für alle zugänglich

BRANCHE

Kunst und Kultur

ANWENDUNGSFALL

Graph-basierte Suche

ZIEL

Durchsuchbare Informationen zu Millionen von europäischen Kunstwerken und Kulturobjekten

HERAUSFORDERUNG

Herkömmliche Datenbanken waren für die Speicherung und Traversierung so großer vernetzter Datenbestände nicht geeignet

LÖSUNG

Neo4j bildet komplexe Daten so ab, dass sie schnell und mühelos durchsucht werden können

ERGEBNISSE

- Neo4j ermöglicht es den Nutzern, über 6 Millionen Kunstwerke und Kulturobjekte effizient zu durchsuchen
- Europäische Museen, Bibliotheken und andere Einrichtungen werden einem globalen Publikum zugänglich gemacht

Von der Mona Lisa zu Mozarts Symphonien – Informationen zu Millionen von Meisterwerken aus ganz Europa werden jetzt von der Europeana Foundation im Internet „zur Schau gestellt“. Diese vernetzten Daten erschließen sich anhand der Neo4j Graphdatenbank Forschern und Enthusiasten auf der ganzen Welt, die sich mit dem reichhaltigen kulturellen europäischen Erbe beschäftigen.

Die Organisation

Die europäische Kultur kommt in Millionen von Gemälden, Büchern, Partituren usw. zum Ausdruck, die über viele Jahrhunderte zurückreichen – doch bis vor kurzem musste man an zahlreichen Orten Station machen, wollte man sie zu Gesicht bekommen. Europeana, eine gemeinnützige Organisation, zu deren Gründern auch die Europäischen Kommission gehört, wurde 2008 ins Leben gerufen, um Menschen auf der ganzen Welt die Chance zu geben, diese Werke im Internet kennen zu lernen und zu erkunden. Über die letzten acht Jahre haben die 60 Mitarbeiter von Europeana eigens zu diesem Zweck Kulturobjekte aus Museen, Bibliotheken, Universitäten und Archiven in der EU katalogisiert. Inzwischen sind Informationen zu über 53 Millionen Artefakten zusammengestellt worden, die nun für alle, die Interesse haben, über ein Webportal und ein API zugänglich sind.

Die Herausforderung

Die Europäische Stiftung mit Sitz in den Niederlanden charakterisiert sich selbst so: „Wir verwandeln die Welt mit Kultur.“ In diesem Sinne hat die Stiftung mehr als 3000 europäische Institutionen gebeten, die digitalen Daten zu den in ihrem Besitz befindlichen Kunstwerken beizusteuern. Mit dem Erfassen und Bereitstellen dieser Datenbestände im Internet macht Europeana diese Einrichtungen weltweit bekannt und präsentiert ihre Sammlungen Kunstfreunden und Wissenschaftlern auf allen Kontinenten.

Das ist Wertschätzung der Kunst im digitalen Zeitalter. „Wir leben in einer digitalisierten Welt“, meint Matt Nader, Infrastruktur-Manager bei Europeana. „Besuche in einer Bibliothek und das Lesen alter Zeitungen sind selten geworden. Aber mit einer digitalen Version kommt eine unmittelbare Nähe zustande.“

Das 10-köpfige Team von Entwicklern bei Europeana hat zu jedem einzelnen Objekt Datenbanken mit strukturierten Informationen erfasst: Entstehungsdatum, Künstler, Aufbewahrungsort – und alle weiteren verfügbaren Details. Außerdem können die Informationen in verschiedenen europäischen Sprachen abgerufen werden.

Als Beispiel nennt Nader das vielleicht berühmteste Gemälde der Welt, Leonardo da Vincis „Mona Lisa“, das im Louvre in Paris hängt. „Wenn ein Wissenschaftler Zugang zu den Originalinformationen zur ‚Mona Lisa‘ haben möchte, musste er oder sie bisher dem Louvre einen Besuch abstatten, um Einsicht in das Material zu bekommen. Bei uns kann jetzt das Material über unser API oder unser Portal gesichtet werden.“

Europeana wollte Nutzer dazu anregen, die Artefakte gründlich zu recherchieren; deshalb sollten Relationen zwischen Informationen hergestellt werden, die in einer relevanten Beziehung zueinander stehen, also Informationen zur ‚Mona Lisa‘ oder zu einem bestimmten Maler oder Komponisten.

Doch an diesem Punkt musste das Team feststellen, dass keine der traditionellen Datenbanken für diese Menge vernetzter Daten geeignet war.

„Unser Ziel war es, möglichst viele Verbindungen wie zwischen den kulturellen Artefakten herzustellen“, sagt Yorgos Mamakis, einer der System-Architekten. „Aber uns fehlte die Möglichkeit, eine

FALLSTUDIE



„Wir sind wirklich zufrieden mit unserer Entscheidung zugunsten von Neo4j. Es passt 100% für unseren Zweck.“

– Yorgos Mamakis
System Architect, Europeana

„Relation“ von einem Objekt zu einem anderen abzubilden, weil diese irgendwo im Datenmodell versteckt war. Denn der Vorgang ist angesichts der riesigen Anzahl von Datensätzen so speicherintensiv, dass es zu Milliarden oder gar Billionen von Einträgen in einem herkömmlichen Triple-Store käme. Solche Mengen zu traversieren und derartige Informationen abzurufen, wäre mit Standard-Hardware extrem langsam.“

Die Lösung

Die Lösung zu diesem Problem fand Europeana mit der Graphdatenbank Neo4j, die 2014 installiert wurde. Dazu Mamakis: „Die offenkundige Lösung war Neo4j, eine Graphdatenbank, bei der die Funktionalität, die wir uns wünschten, bereits eingebaut war. Neo4j bot uns die Traversierung der Relationen und die Relationen, auf die es uns ankam, in einer strukturierten Weise.“

In dem derzeit laufenden Projekt präsentiert Europeana jetzt mehr als 6 Millionen (12%) seiner 53 Millionen Kulturobjekte und Datensätze in Neo4j. Die Datenbank, deren Implementierung sehr einfach war, gewinnt für die Projekte der Stiftung zunehmend an Bedeutung. „Wir fanden es überraschend einfach, mit Neo4j zu arbeiten“, sagt Mamakis.

Mamakis führt weiter aus: „In dem Maße, wie die Menge der Artefakte zunimmt, wird immer mehr in Neo4j gespeichert, sodass diese Datenbank zu einem unserer Schlüsselsysteme wird. Idealerweise werden wir in der nicht allzu weiten Zukunft in der Lage sein, die Anzahl der anderen Backends, die wir noch haben, zu reduzieren.“

Dank Neo4j legt Europeana Besuchern auf ihre Anfragen „ähnliche“ Objekte vor, sodass sie miteinander verbundene Informationen mühelos sichten und je nach Interesse konsultieren können. Eine Suche im Zusammenhang mit Mona Lisa kann also Dutzende oder gar Hunderte von Ergebnissen zutage fördern. Europeana bietet auch einen „Explore“-Button sowie zahlreiche Online-Ausstellungen, die zu weiteren Entdeckungen anregen.

Mamakis führt weiter aus: „Mit Neo4j gewinnen unsere Datensätze an Wert und die Benutzererfahrung ist eine ganz andere. Man hat jetzt völlig neue Möglichkeiten, die Daten zu durchsuchen. Es wird nicht nur einfach ein Objekt abgerufen, sondern mit ihm die ganze Familie, die damit verknüpft ist. Das heißt, man hat einen ganz anderen Einstieg, um auf neue Objekte zuzugreifen und kann deshalb potenziell viel mehr Informationen zum Thema finden, für das man sich interessiert.“

Das Ergebnis

Europeana ist ein Erfolg: Es zieht regelmäßig um die 250 000 Online-Besucher und 1 Million Seiten-Ansichten jeden Monat an. Die Art und Weise, wie Neo4j den mühelosen, agilen Zugang zu Informationen und ein flexibles Durchsuchen ermöglicht, überzeugte auch die teilnehmenden Institutionen sofort.

Dazu Mamakis: „Je mehr Partner wir davon überzeugen können, ihre Daten in dieser Weise zu strukturieren, desto besser. 12% davon sind schon jetzt im Neo4j-Backend präsent, was heißt, dass unsere Datenpartner mit uns einig gehen, dass Neo4j eine nützliche Funktionalität bereitstellt. Je mehr Daten vorliegen, desto besser für uns - und desto besser auch die Benutzererfahrung. Und das ist letztendlich, worum es uns geht.“

Matt Nader meint dazu: „Ich bin ganz sicher, dass für uns die Neo4j Graphdatenbank ein riesen-großer Schritt ist, der viele Vorteile bietet. Wir können unsere Objekte in kürzester Zeit finden. Das schrittweise Verfahren in einer herkömmlichen relationalen Datenbank würde die Anwendung sehr schwerfällig machen und die Reaktionszeiten ganz einfach zu lang. Neo4j hingegen reagiert unglaublich schnell.“

Mamakis schloss mit dieser Bemerkung: „Wir sind wirklich zufrieden mit unserer Entscheidung zugunsten von Neo4j. Es passt 100% für unseren Zweck.“

Info zu Neo Technology

Neo Technology ist Entwickler der weltweit führenden Graphdatenbank Neo4j. Neo4j ist eine hoch skalierbare native Graphdatenbank, die Relationen zwischen Daten als Einheiten erster Klasse nutzt. Dies versetzt Unternehmen in die Lage, Applikationen zu entwickeln, die selbst so schwierige Problemen wie Betrugsaufdeckung, Echtzeit-Empfehlungen, Stammdaten-Management, Netzsicherheit und IT-Transaktionen meistern können.

Großunternehmen wie Walmart, UBS, Cisco, HP, adidas und Lufthansa, aber auch „heiße“ Startups wie Medium, Musimap und Glowbl verlassen sich auf Neo4j, um aus den Beziehungen in ihrem Datenbestand den optimalen Nutzen zu ziehen.

Großbritannien uk@neotechnology.com
Frankreich ventes@neotechnology.com
Skandinavien nordics@neotechnology.com
Deutschland vertrieb@neotechnology.com
Südeuropa
southern-europe@neotechnology.com