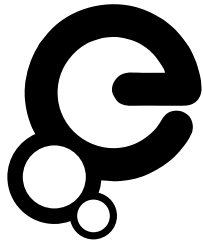**CASE STUDY**

Europeana

Neo4j Helps to Make Millions of European Artworks and Cultural Objects Accessible to All

INDUSTRY
Arts & Culture

USE CASE
Graph-Based Search

GOAL
Provide searchable information about Europe's millions of artworks and cultural objects

CHALLENGE
Traditional databases were too unwieldy to store and traverse across such large volumes of connected data

SOLUTION
Neo4j presents related data about artefacts that can be quickly and easily explored

RESULTS
– Neo4j helps users efficiently search through over 6 million artworks and cultural objects
– Europe's museums, libraries and other institutions are promoted to a global audience

From the Mona Lisa to Mozart's music, information about millions of art-works from across Europe is being showcased online by the Europeana Foundation. This 'connected' data can be readily explored using the Neo4j graph database, helping academics and enthusiasts worldwide to find out more about Europe's rich cultural heritage.

The Organisation

European culture is encapsulated in millions of paintings, books, music recordings and the like, dating back centuries – but until recently, the only way people could find out about these artefacts was to visit the particular institution housing them. Europeana, a non-profit organisation co-funded by the European Commission, was established in 2008 to help individuals find out about these works of art online. For the past eight years, Europeana's 60 staff have been cataloguing cultural objects from museums, libraries, universities and archives across the European Union in order to promote them to the wider world. They have now amassed searchable, digital information about over 53 million artefacts and made this available to users globally via a web portal and API.

The Challenge

The self-declared mission of the Netherlands-based Europeana Foundation is: "We transform the world with culture." To achieve this, the Foundation has encouraged over 3,000 European institutions to provide all their digital data about their artworks. By gathering this information and making it available online, Europeana has been able to promote the institutions and make their collections known to enthusiasts and researchers worldwide.

It is art appreciation for the digital age. "We are living in a digitised society," said Europeana Infrastructure Manager Matt Nader. "Going to a library and reading an old newspaper doesn't happen very often any more. But a digital version, it makes it much closer to you."

Europeana's team of around 10 developers have built databases of structured information about every artefact – the date it was created, by whom, where it's located and all digitised data available about it. Crucially, they have made the information accessible via multiple European languages so almost everyone can access it.

Nader gives the example of perhaps the world's best-known painting – Leonardo da Vinci's Mona Lisa, displayed in the Louvre museum in Paris. "If you are a researcher and you want to have original information about 'Mona Lisa, you have to go to the Louvre and arrange some time to be able to see the requested material. But we provide that, the original authenticated material, via our API or portal."

Europeana also wanted to encourage users to explore the artefacts in-depth, by creating connections between related items – like all information about Mona Lisa, or all works by the same painter or composer.

But at this point, said System Architect Yorgos Mamakis, the Foundation realised that using traditional databases to house such 'connected' data was impractical.

CASE STUDY

"We are really happy that we made the decision to use Neo4j. It most certainly fits our purpose 100%."

– Yorgos Mamakis
System Architect,
Europeana

"Our objective is to make as many connections between the cultural artefacts as possible," said Mamakis. "But we were missing a meaningful way to have a 'relation' and to go from one object to another via these relations because it was hidden somewhere in the data model. It's so memory-intensive that, considering the number of records that we have, would result in billions or trillions of triples in a standard semantic repository. And traversing over that or retrieving that sort of information would be extremely slow with standard hardware."

The Solution

Europeana found its answer in Neo4j, which it installed in 2014. As Mamakis said: "The most obvious solution was Neo4j, a graph database supporting everything we wanted out-of-the-box. Neo4j provided the relations traversal and the links we needed in a structured manner."

In an ongoing project, Europeana now presents over 6 million (12%) of its 53 million cultural objects and records in Neo4j. The database, which was easy to implement – "We did not find it complicated at all to work with Neo4j," Mamakis said – is becoming increasingly central to the Foundation's operations.

Mamakis explained: "We expect that, as the number of artefacts we have increases, more and more will end up in Neo4j so that it becomes one of our core systems. Ideally we will be able in the not so distant future to cut down on the other backends that we have."

Through Neo4j, Europeana offers visitors 'Similar Items' to encourage them to move between related pieces of information and find out more about artefacts, based on their own interests. Searches on Mona Lisa, for example, now turn up tens or even hundreds of results. Europeana also offers an 'Explore' button and hosts dozens of online 'Exhibitions' to encourage further discovery.

Mamakis explained: "Neo4j adds a benefit in the quality of our records and the user experience. It's the fact that the users get another way of browsing through the data. Now you don't just retrieve the object, but also the family of objects that is closely related to that – so you have another entry point to access new objects and potentially find more information about what you are interested in."

The Results

Europeana is a success: It regularly attracts around 250,000 online visitors and 1 million page views every month. Its Neo4j-based approach to making data easily accessible and searchable was also an instant hit with the institutions providing the information.

Mamakis said: "The more partners we convince to structure the data in such a way, the better. 12% are already present in the Neo4j backend which means that our data partners are convinced that Neo4j is useful. The more data in there the better for us, and the better experience for the user. This is really the bottom line for us."

Matt Nader said: "I think the advantage of having the Neo4j graph database is a huge step. We can find our objects very fast. Going step-by-step in a traditional relational database, the application would become very heavy, response time would be slow. Neo4j is responding very fast."

Mamakis concluded: "We are really happy that we made the decision to use Neo4j. It most certainly fits our purpose 100%."

About Neo Technology

Neo Technology is the creator of Neo4j, world's leading graph database. Neo4j is a highly scalable native graph database that leverages data relationships as first-class entities to help companies build intelligent applications that meet today's evolving connected data challenges including fraud detection, real-time recommendations, master data management, network security and IT operations.

Enterprises like Walmart, UBS, Cisco, HP, adidas and Lufthansa and hot startups like Medium, Musimap and Glowbl rely on Neo4j to harness the connections in their data.

United Kingdom uk@neotechnology.com

France ventes@neotechnology.com

Scandinavia nordics@neotechnology.com

DACH vertrieb@neotechnology.com

Southern Europe
southern-europe@neotechnology.com