

## CASE STUDY



## albelli

## Neo4j Helps Photo Organization Service Monetize Relationships in 1.2 Petabytes of Data

**INDUSTRY**

Software / Printing

**USE CASE**

Social Network /  
Graph-Based Search

**GOAL**

- Create a new photo storage and organizing platform to replace existing legacy platform
- Deliver a well-organized, easy-to-use and secure storage solution for all of a user's images, easing the process of selecting and creating photo products

**CHALLENGE**

- Microsoft SQL Server required cumbersome JOINS that were no longer working for given volume of data
- Complete development within a continuous deployment environment and short time to market

**SOLUTION**

Transferred and organized 1.2 petabytes of photos from on-premise to cloud where Neo4j holds all metadata-related information

**RESULTS**

- Allows for continued user growth without performance fears
- Frees up time to focus on the strategic goals of increasing monetization, gathering customer feedback and cutting costs

The albelli company provides software that syncs and organizes a huge volume of photos for over a million users. With Neo4j, the team provided a modernized tool that can continue to grow along with its user base, providing increased revenue and allowing for the pursuit of strategic monetizing goals.

**The Company**

Founded in 2003, albelli combines the opportunities of the new digital online era with the “traditional craftsmanship” of bookbinding and printing. Their online photo products software has won several awards throughout Europe and serves multiple markets in the region. They are the market leader in the photo book printing industry and are one of the fastest growing e-commerce companies in Europe. albelli also owns other photo products brands such as resnap, bonusprint, önskefoto and fotoknuden.

**The Challenge**

albelli provides users with the ability to sync photos from multiple sources to one location. The software automatically organizes the photos based on a variety of factors — such as date, time and location — into events that are displayed in a user-friendly timeline, keeping customers from having to sort through and organize thousands of photos. Customers have complete control over sharing options and can determine which albums are public or private.

After photos are uploaded and organized, users have the ability to purchase a variety of products that range from photo albums to calendars, mugs and canvas portraits.

Because of the large number of photos — there are over a million users with an average of 2,000 photos each — as well as the large size of the files in combination with all of the descriptors associated with each photo, the company was dealing with huge volumes of data.

The company was still using Microsoft SQL Server, a relational database that was cumbersome and slow because it required such a huge number of JOINS for relationship-based queries. It was becoming clear that they needed to add the capabilities of another database to the mix to overcome these data relationship challenges.

The development team started from scratch and recognized that their domain was made up of graph-like relations between photos and users, and they knew translating this graph-like domain to SQL Server wouldn't be a good fit. So, they decided to start a proof-of-concept project with Neo4j.

“If the average photo collection user has 2,000 photos, how do we then cope with tens of nodes on each of those photos and keep the product fast and flowable?” said Josh Marcus, Chief Technology Officer at albelli. “That's where we found our biggest technical challenge.”

## CASE STUDY



“The huge advantage with Neo4j was that we were able to focus on modeling our data and how to best serve our customers instead of agonizing how to structure tables and JOINS. It also required very little coding, so we were able to keep our focus on our customers.”

– Josh Marcus  
Chief Technology Officer,  
albelli

### The Strategy

The company had a nimble research and development team at its disposal — only five people. When they first began building the new data architecture, they explored a number of different databases. But they weren't just looking for a database to store data; they were looking for a way to leverage the relationships that existed between their data.

This led them right to Neo4j. They first heard about the software at a conference, and reached out to learn more. Once the decision had been made to use Neo4j, the team took advantage of all the training tools available — including the online training courses and coding academy — to become Neo4j Certified Professionals.

With a strong grasp of Cypher and Neo4j, the next step was to transition their huge volume of data into their new graph database.

### The Solution

At peak, the team used up to 700 EC2 instances running at maximum speed for one and a half months to get 1.2 petabytes of data — which included 500 million images plus customers data — transferred from on premise to cloud storage. The migration involved up to five heavily equipped Neo4j servers and saw over 10 billion messages sent across the system. During this process they also extracted the date, time and location information from each photo so that it could then be related to others.

“The huge advantage with Neo4j was that we were able to focus on modeling our data and how to best serve our customers instead of agonizing how to structure tables and JOINS,” said Marcus. “It also required very little coding, so we were able to keep our focus on our customers.”

In the end, their database boasted a total of one billion nodes, 4.1 billion properties and 2.6 billion relationships with Neo4j acting as the central database. Alongside Neo4j, albelli used other database technologies in their architecture, including Redis, DynamoDB, Aurora and Microsoft SQL Server – allowing them to take advantage of a polyglot persistence approach.

### The Results

In a little less than two months, the small albelli team was able to transfer their data into Neo4j and launch a faster and more responsive database for a growing customer base. This frees up a significant amount of time for them to focus on the future, which includes new ways to monetize, focus on gathering customer feedback and cutting costs.

Neo4j is the leader in graph database technology. As the world's most widely deployed graph database, we help global brands – including [Comcast](#), [NASA](#), [UBS](#), and [Volvo Cars](#) – to reveal and predict how people, processes and systems are interrelated.

Using this relationships-first approach, applications built with Neo4j tackle connected data challenges such as [analytics and artificial intelligence](#), [fraud detection](#), [real-time recommendations](#), and [knowledge graphs](#). Find out more at [neo4j.com](#).

**United Kingdom** [uk@neo4j.com](mailto:uk@neo4j.com)  
**France** [ventes@neo4j.com](mailto:ventes@neo4j.com)  
**Scandinavia** [nordics@neo4j.com](mailto:nordics@neo4j.com)  
**DACH** [vertrieb@neo4j.com](mailto:vertrieb@neo4j.com)  
**Southern Europe** [southern-europe@neo4j.com](mailto:southern-europe@neo4j.com)