



## Graph Exploration With Neo4j

*Client Contact:* Jack Kelly ([jack@fullcontact.com](mailto:jack@fullcontact.com))

*Team Size:* 3 - 4

*Location:* Tabor Center, Downtown Denver

*Skillset:* Java/Scala, RESTful APIs, Graphs

### About Us:

FullContact is a platform for contact management and insights. Our mission is to make relationships better. We're a 2011 TechStars Boulder graduate that provides APIs and applications that allow businesses and consumers to keep their contact information up-to-date automatically. We're a team of nearly 100 people across the world in Colorado, Latvia, Israel, and India, headquartered in an office in LoDo (Lower Downtown Denver). Our engineering department in Denver has 15 members focused on APIs, infrastructure, and data processing.

### The Problem:

At the core of FullContact's ability to make relationships better is a massive graph database of contact information. This database was originally conceived as part of a 2011 field session project, and currently houses over two billion vertices, each representing a fragment of a contact record. We do our best to connect these fragments as we pull them from multiple sources for analysis, but sometimes we get it wrong.

When we get things wrong, maybe our database conflates two individuals when it shouldn't. Or maybe it cleaves a person's identity into two, or three, or more pieces. In either case, we need to inspect the graph to discover what data is causing issues. But because our database is optimized for querying into these two billion vertices, we lack a good way to visualize and inspect the graph for anomalies.

## The Project:

We want you to research, design, and build a prototype graph visualization and investigation platform using Neo4j ( [neo4j.com/product](https://neo4j.com/product) ) and related technologies. You'll build software capable of passing the JSON-based responses from our existing graph query server to a Neo4j database with the intent of making some subgraph of the larger identity graph available for investigation. We're particularly interested in some of the graph visualization tools ( [neo4j.com/graph-visualization-neo4j](https://neo4j.com/graph-visualization-neo4j) ) powered by Neo4j.

We'll be working with you to assist with the research and design components as necessary, but it will be up to you to handle putting all the pieces together in the end.

### What Success Looks Like

- The relevant subgraph for a queried piece of contact data (e.g., [jack@fullcontact.com](mailto:jack@fullcontact.com)) can be loaded into a Neo4j database
- The loaded subgraph can be investigated in some visual format
- *Extra:* You set up the prototype as a standalone service in our AWS infrastructure for collaborative graph exploration

### Skills Required

- Scala (preferred) or Java (okay) programming language
- git version control
- Basic HTTP concepts
- Basic database concepts
- Able to analyze and discuss graphs

### Student Benefits

- Reimbursement for commute expenses

## Summer Internship:

Many of our full time engineering staff are graduates of CSM and participated in field session projects that we hosted.

A limited number of paid summer internships may be available upon course completion.