



# FullContact

## LOCATION

18th & Blake  
(LoDo)

## IDEAL TEAM SIZE

4

## I Know That Face

### About Us:

FullContact is solving the world's contact information problem. 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 80 people located in LoDo, walking distance to Union Station and the 16th Street Mall.

### About You:

You enjoy solving hard problems and figuring out how to make things work. You are passionate about algorithms, machine learning, and mathematical models. You think computers can help to create and strengthen connections between people.

### The Project:

In 2013, a talented Field Session team helped FullContact build what evolved into an important component of our tech stack: a library to identify if two images are visually the same even when byte-wise distinct. Now, we are looking for a new team to help build the next step in our image-processing backend: a system based on machine learning and data analysis techniques to create a database of faces for on-the-fly recognition. This system will start by learning the faces of FullContact's employees and some day scale up to operating on all the millions of profiles in our identity database.

Nearly every social network on the internet allows users to upload an image file to visually represent them. Users often use the same image, or variations thereof, on each network they have an account on, thereby forming a consistent visual identity. This isn't always the case, however - an image chosen for LinkedIn may be a professional image of the user in business attire, whereas their picture for Twitter may feature the user "dressed down" in a more casual setting.

The common element between these two cases is the user's face. This is often the strongest signal we have as humans when recognizing another person, but is currently not something FullContact has the technical capability to leverage when we are connecting all the components of an individual's public persona. FullContact's collection of billions of profile images has millions of faces in it, but we need a way to discover who those faces belong to and search for them again when we find a new image of a person.

The goal of this Field Session project is to build a prototype of an automated service to identify a user by means of a new image of their face and then retrieve their associated data from FullContact's Person API. The project will leverage OpenCV for image processing and facial recognition via eigen/fisherfaces, and Elasticsearch for fuzzy-querying previously unseen images against the training dataset; we would also be excited to work with students to incorporate any other technologies that would help solve this difficult problem and make this project a success.

## What Success Looks Like

A successful project will be able to:

- Train an automated system to identify, in real time, a person chosen from a limited group based on an image of their face
- Retrieve the FullContact profile of an identified person (social handles, profile images, job title, etc.)
- Stretch Goal: Identify a person based on an image of them where their face is not the prominent feature/focus
- Stretch Goal: Register and later identify a person who was previously unknown to the system

## Skills Required

- You must be willing to solve (really) hard problems
- You must enjoy working with algorithms
- A basic understanding of Linear Algebra (MATH332) will be helpful

## Student Benefits

- You will end the session with the reward of solving a hard problem
- You will be exposed to cutting-edge technology
- You will be working in a collaborative environment with other smart engineers
- You will experience the business and engineering methodologies that FullContact employs

## Summer Internship:

A limited number of paid summer internships may be available upon course completion. Also, about  $\frac{1}{3}$  of our full time engineering staff are graduates of CSM and participated in field session projects that we hosted.