

Team 096 Project Proposal

1. OpenSource
- 2.

Our project aims to aggregate the findings of several databases related to criminal activity and present it in a way that makes relationships between different entities or events clear. This application can take the place of pin boards you may see detectives use in investigation to map different **objects** to each other. An object can be a person, IP address, gang, etc. An object can be imported into a board with either a manual entry, or a lookup of a person, event, IP address, etc in one or more publicly accessible databases (facilitated by our application). While importing an object into the board, we will have an algorithm that suggests adding **relationships** between the new object and objects on the board. The user also has the opportunity to enter relationships between objects on the board. The visual representation will look much like a pinboard used by detectives. Objects will be represented by nodes and relationships by edges.

- 3.

The intention of this project is to replace evidence/pin boards used by investigators to visualize and draw connections between existing evidence. In addition to all the pros of a physical evidence board, our application will scan various databases (the actual databases may be specified by the user) to suggest connections between objects already on the board not recognized by the human user. This allows investigators to utilize all the information they have in the most efficient manner.

We plan to accomplish this in several ways:

- We will gather a list of publicly available databases on objects (people/known criminals, IP addresses, devices, etc).
- We have users log into the website (this is required to reasonably save their work)
- Users can create a new board and add an object.
 - If the object is, for example, a known criminal, we can query information on that person from a database. Users can enter or edit any field they want and that is the representation that will be saved in our database.
- Once added, we will have an algorithm that suggests potential relationships between the new object and existing objects in the board.

- The user can choose to accept or reject the suggestion or add their own.
- Objects will be represented by nodes on the board and relationships by edges.
- Users will be able to compute and view the overlap between different boards they or others in their organization create to see if different investigations are actually related.

4.

We will use some tools to help us visualize data. For example, the [Network Chart from Python Graph Gallery](#) would transform the relations between objects that come out from the database into a visualized graph and it is easier for people to understand. Moreover, we will implement an algorithm to evaluate the suggested relationships between different objects and allow users to input their own relationships into the database. For example, we would suggest to the users that 2 objects have potential relationships and the user will make their choice to accept the suggestion or discard it, the user will also be able to add their own information to the database. We can also compute the overlap between separate boards (separate investigations) for users.

5.

This project is useful for law enforcement agencies who want to find potential suspects, witnesses, or links to other criminals. The basic functions include lookup of criminals, which lists their criminal connections, and also the relationship between criminals who do not have a direct connection. Using a node tree to visualize the connections between criminals is a complex feature, whereas it is simple to list attributes (ip addresses, location, etc.) of a criminal. There are currently many websites available that have criminal lookups (such as <https://illinoiscourtrecords.us/criminal-court-records/>). Ours is unique in that we specialize in the relationships between different criminals and how they are related to other objects. We also show relationships between separate investigations within the same organization.

6. Dataset 1:

<https://catalog.data.gov/dataset/crimes-one-year-prior-to-present/resource/e3a0a89d-cab5-4280-b6a5-20a1781139c3>

Dataset 2:

<https://catalog.data.gov/dataset/nypd-arrest-data-year-to-date>

Dataset 3:

<https://catalog.data.gov/dataset/crime-data-from-2020-to-present/resource/5eb6507e-fa82-4595-a604-023f8a326099>

We will use the crime datasets published by the United States Government (<https://data.gov/>) as the backbone to support the data query. These data sources are all in csv format. The data size is all above 100000 criminal events for each data source. These data sources basically capture the information of the location of the event, the reported time, the occurring time, the event type, the criminal's name, the victim's name, etc

7. Users of the website will be able to do the following:
 1. Login their account
 2. View the about us page
 3. Enter an existing board
 4. Create a new board
 5. Add objects to a board
 - a. They can search for terms in the search box while checking off boxes for databases we will search.
 - b. Once we populate the fields, users can edit any field they want before saving an object to the board
 6. Add relationships between objects via our suggestions or entered edges.
 7. View the overlap between different boards via a score we compute.
2. Work Distribution
 - a. Recommendation Algorithm for relationship suggestion: Alex
 - b. Frontend Page layout/infrastructure: Philip James
 - c. Frontend Flowchart/board visual: Chinmay
 - d. User authentication: Philip
 - e. Backend API Design: Boyang
 - f. API implementation: All
 - g. Creation and querying of our databases: Philip, Alex

- h. Connecting to and querying external databases:
Boyang, Chinmay
- i. Overlap algorithm: Philip

Page sketches:

