

## Spectacular Cranes Project Report

Grant Gasser, Kartik Kesavabhotla, Sundaram Thevar

### [Project Link](#)

While speaking with Rajiv after the seminar on the Chicago Police Database, we learned that the Invisible Institute has noticed that allegation files tend to change over time, including information such as the category of the allegation, location, accused officers and more. We decided to make the theme of our project about tracking how the category of allegation category changes over time. While doing this, we also wanted to see if the new category was completely unrelated to the initial category or if it was much less severe than the initial category. We also wanted to see if we could identify any “cover charges” that the department frequently uses to either justify severe action or mask the severity of an allegation.

The first step for us was data integration. Since our theme is looking at how allegations have been changed over time (e.g. the category of allegation being changed to something less serious). This requires us to access data not currently available in CPDB. Professor Rogers and the Invisible Institute provided us with access to database files and we found

**08\_28\_2018\_case\_info** to be useful for answering these questions. This table contained allegation information such as the category of the allegation and the location. We were able to join this table with tables from CPDB which allegations had their category changed and which officers were involved.

### Data Integration and Relational Analytics

While integrating the data, we encountered an early challenge: of all of the allegations that had their categories changed (the strings are different), many of the new categories *meant* the same thing as the old category. This makes it really difficult to tell when allegations were changed to something completely different (possibly less severe) or when the category was just renamed because of different nomenclature. An example of this is in the figure below:

| old_category_desc                           | old_category   | current_category               |
|---|--|--------------------------------|
| 1 NO ARREST                                 | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Use Of Force                   |
| 2 NO ARREST                                 | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Use Of Force                   |
| 3 ARRESTEE-DURING ARREST                    | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Use Of Force                   |
| 4 UNNECESSARY PHYSICAL CONTACT -ON DUTY     | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Domestic                       |
| 5 USE OF PROFANITY                          | GROUP 01 - VERBAL ABUSE                                    | Verbal Abuse                   |
| 6 USE OF PROFANITY                          | GROUP 01 - VERBAL ABUSE                                    | Verbal Abuse                   |
| 7 MISCELLANEOUS                             | GROUP 08 - CRIMINAL CONDUCT                                | Criminal Misconduct            |
| 8 MISCELLANEOUS                             | GROUP 08 - CRIMINAL CONDUCT                                | Criminal Misconduct            |
| 9 IMPROPER SEARCH OF PERSON                 | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 10 NO ARREST                                | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Use Of Force                   |
| 11 SEARCH OF PREMISE WITHOUT WARRANT        | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 12 SEARCH OF PREMISE WITHOUT WARRANT        | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 13 UNNECESSARY PHYSICAL CONTACT -ON DUTY    | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Domestic                       |
| 14 UNNECESSARY PHYSICAL CONTACT -ON DUTY    | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Domestic                       |
| 15 NEGLIGENCE OF DUTY                       | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 16 NEGLIGENCE OF DUTY                       | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 17 INVENTORY PROCEDURES                     | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 18 INVENTORY PROCEDURES                     | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 19 MISCONDUCT DURING ISSUANCE OF CITATION   | GROUP 07 - TRAFFIC (NOT BRIBERY/EXCESSIVE FORCE)           | Traffic                        |
| 20 REPORTS                                  | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 21 REPORTS                                  | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 22 ARRESTEE-DURING ARREST                   | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Use Of Force                   |
| 23 ARRESTEE-DURING ARREST                   | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Use Of Force                   |
| 24 SEARCH OF PREMISE WITHOUT WARRANT        | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 25 SEARCH OF PREMISE WITHOUT WARRANT        | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 26 DOMESTIC ALTERCATION/INCIDENT - OFF DUTY | GROUP 05 - OFFICE OF PROFESSIONAL STANDARDS INVESTIGATIONS | Domestic                       |
| 27 SEARCH OF PREMISE WITHOUT WARRANT        | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 28 SEARCH OF PREMISE WITHOUT WARRANT        | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 29 SEARCH OF PREMISE WITHOUT WARRANT        | GROUP 03 - IMPROPER SEARCH                                 | Illegal Search                 |
| 30 NEGLIGENCE OF DUTY                       | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 31 NEGLIGENCE OF DUTY                       | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |
| 32 NEGLIGENCE OF DUTY                       | GROUP 10 - OPERATION/PERSONNEL VIOLATIONS (ON DUTY)        | Operation/Personnel Violations |

**Fig. 1: Looking at old\_category\_desc and current\_category, towards the bottom one can see that an allegation category of "SEARCH OF PREMISE WITHOUT WARRANT" is now represented as "Illegal Search." They essentially mean the same thing, and is certainly not evidence of officers or investigators changing the allegation to make it seem less serious.**

Unsure about how to solve this problem at the time, we continued with an aggregate analysis of the allegation categories from the old data to the new data.

Below is a glance at the distributions of allegations in each the old and new databases to give you a sense of how things look. It gives one a gross overview of shifts in the distributions as well as language used (Fig 2). The relative distributions are similar within each database however the ranking of allegations has changed.

## Creating the Mapping

Having had a chance to make some superficial comparisons between what we were dealing with in the old and the new databases with regard to these allegation designations, the next step was trying to figure out how we were in fact going to determine/define allegation change. We had to have a systematic method for doing this on a large scale. Our approach was as follows: for a given allegation category for a record in the old dataset, what is the most common new category name it has in CPDB? This approach is not perfect, so we took some liberty with the

idea and figured that although we may miss out on some suspicious allegations, this would allow us to narrow down to a subset of allegations that suspiciously had their allegations changed.

If we could define the "**mapping**" using this most common idea, then we have some idea about how the category names have changed over time. If we saw a record where the mapping is violated (the new category name is not what is usually is given the old category), then we may have a change that is more suspicious (as opposed to the new category string being different but meaning the same thing as the old category). This could reveal some foul play such as changing the allegation category to something less "severe" or something completely unrelated.

This is a key part of our project, so we'll give a quick example to make it easy to understand (Fig 2):

| old_category""                         | new_allegation_name""                                | num_allegation_name_given_old | percent_allegation_name_given_old_category | allegation_mapping_boolean |
|--|--|-------------------------------|--|----------------------------|
| Null                                   | "Excessive Force / On Duty - No Injury"              | 2                             | 100.00%                                    | TRUE                       |
| "Absent Without Permission"            | "Seat Belts"   | 1                             | 0.68%                                      | FALSE                      |
| "Absent Without Permission"            | "Leaving Assignment (District, Beat, Sector, Court)" | 2                             | 1.36%                                      | FALSE                      |
| "Absent Without Permission"            | "Failure To Ensure - Civil Rights"                   | 2                             | 1.36%                                      | FALSE                      |
| "Absent Without Permission"            | "Inadequate / Failure To Provide Service"            | 2                             | 1.36%                                      | FALSE                      |
| "Absent Without Permission"            | "Use Of Profanity"                                   | 1                             | 0.68%                                      | FALSE                      |
| "Absent Without Permission"            | "Absent Without Permission"                          | 139                           | 94.56%                                     | TRUE                       |
| "Abuse Of Authority"                   | "Use Of Official Position"                           | 1                             | 1.23%                                      | FALSE                      |
| "Abuse Of Authority"                   | "Miscellaneous"                                      | 6                             | 7.41%                                      | FALSE                      |
| "Abuse Of Authority"                   | "Neglect Of Duty"                                    | 3                             | 3.70%                                      | FALSE                      |
| "Abuse Of Authority"                   | "Reports"  | 2                             | 2.47%                                      | FALSE                      |
| "Abuse Of Authority"                   | "Abuse Of Authority"                                 | 62                            | 76.54%                                     | TRUE                       |
| "Abuse Of Authority"                   | "Altercation / Disturbance - Other"                  | 2                             | 2.47%                                      | FALSE                      |
| "Abuse Of Authority"                   | "Misuse Of Department Records"                       | 2                             | 2.47%                                      | FALSE                      |
| "Abuse Of Authority"                   | "Association With Felon"                             | 3                             | 3.70%                                      | FALSE                      |
| "Accident - Fatal"                     | "Neglect Of Duty"                                    | 16                            | 50.00%                                     | TRUE                       |
| "Accident - Fatal"                     | "Miscellaneous"                                      | 14                            | 43.75%                                     | FALSE                      |
| "Accident - Fatal"                     | "D.U.I. - Off Duty"                                  | 2                             | 6.25%                                      | FALSE                      |
| "Accident - Injury"                    | "Miscellaneous"                                      | 10                            | 26.32%                                     | FALSE                      |
| "Accident - Injury"                    | "D.U.I. - Off Duty"                                  | 28                            | 73.68%                                     | TRUE                       |
| "Accident - No Injury"                 | "D.U.I. - Off Duty"                                  | 6                             | 100.00%                                    | TRUE                       |
| "Accident - Property Damage"           | "D.U.I. - Off Duty"                                  | 16                            | 80.00%                                     | TRUE                       |
| "Accident - Property Damage"           | "Intoxicated Off Duty"                               | 2                             | 10.00%                                     | FALSE                      |
| "Accident - Property Damage"           | "Illegal Arrest / False Arrest"                      | 2                             | 10.00%                                     | FALSE                      |
| "Ageism"                               | "Eeo Investigations"                                 | 6                             | 100.00%                                    | TRUE                       |
| "Altercation / Disturbance - Neighbor" | "Improper Search Of Vehicle"                         | 2                             | 1.20%                                      | FALSE                      |
| "Altercation / Disturbance - Neighbor" | "Inadequate / Failure To Provide Service"            | 2                             | 1.20%                                      | FALSE                      |
| "Altercation / Disturbance - Neighbor" | "Association With Felon"                             | 4                             | 2.40%                                      | FALSE                      |
| "Altercation / Disturbance - Neighbor" | "Altercation / Disturbance - Neighbor"               | 148                           | 88.62%                                     | TRUE                       |

Notice in the chart above, we have **old\_category** and **new\_allegation\_name**. This shows all of the new category names that an allegation with an old category could have taken on. So allegations with "Absent Without Permission" as the old category changed to 6 different allegation categories in the new data: "Seat Belts, Leaving Assignment, Failure to Ensure - Civil Rights, Inadequate / Failure to Provide Service, Use of Profanity, and Absent Without Permission." To the right, we can see the distribution of how allegations with "Absent Without Permission" changed. One column shows the raw number and the other columns shows the

percentage. For each old category, we defined its **true mapping** as the most common change. Thus, the true mapping for “Absent Without Permission” is well, “Absent Without Permission,” since that was the new category 94.56% of the time. We use **allegation\_mapping\_boolean** to denote this. Thus, when **allegation\_mapping\_boolean** is false, this shows a violation of the mapping, which allows us to narrow down the amount of allegations we need to look at that might be suspicious.

Another example above is that “Abuse of Authority” in the old data maps to “Abuse of Authority” in the new data. In this case, just like the previous case, the strings are the same. Of course, this is not always true. Interestingly, 7% of these allegations changed to “Miscellaneous,” which is vague and unrelated to “Abuse of Authority.”

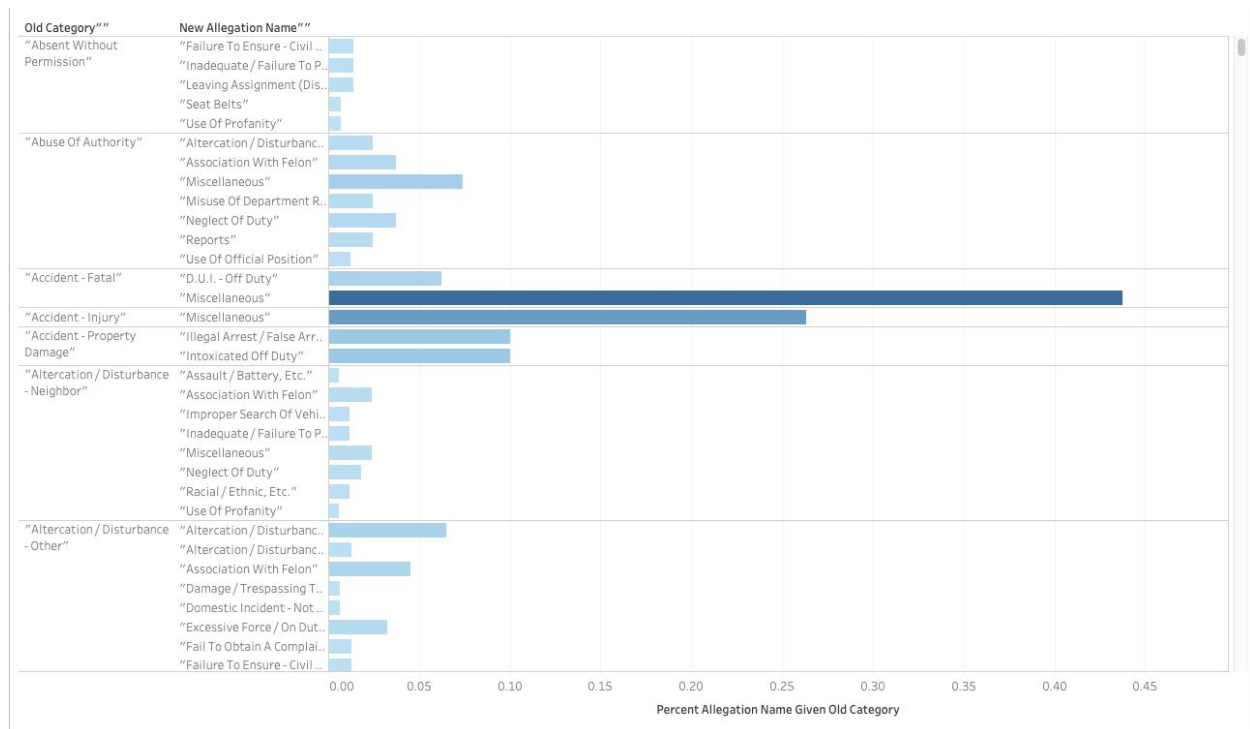
Below we show the **true mapping** of each allegation/location in the old database to the most frequently used allegation/location in the new database (Fig 4). In other words, ‘Choked’ in the old database was most frequently mapped to ‘Excessive Force/On Duty - Injury’ in the new database. One can see the **location** mapping here as well.

| Allegation Mapping                               |   | Location Mapping                                  |                                      |
|--|---|---|--------------------------------------|
| Old Category                                     | Allegation Mapping                                | Old Location Code                                 | Location Mapping                     |
| "Accident - Injury"                              | "D.U.I. - Off Duty"                               | "ABANDONED BUILDING"                              | "Public Way - Other"                 |
| "Accident - No Injury"                           | "D.U.I. - Off Duty"                               | "AIRCRAFT"  | "Airport"                            |
| "Accident - Property Damage"                     | "D.U.I. - Off Duty"                               | "AIRPORT BUILDING NON-TERMINAL - NON-SECURE AREA" | "Airport"                            |
| "Ageism"   | "Eeo Investigations"                              | "AIRPORT BUILDING NON-TERMINAL - .."              | "Public Way - Other"                 |
| "Altercation / Disturbance - Neighbor"           | "Altercation / Disturbance - Neighbor"            | "AIRPORT BUILDING NON-TERMINAL - .."              | "Airport"                            |
| "Altercation / Disturbance - Other"              | "Altercation / Disturbance - Other"               | "AIRPORT EXTERIOR - NON-SECURE AREA"              | "Airport Exterior - Non-Secure Area" |
| "Altercation / Disturbance - Traffic"            | "Altercation / Disturbance - Traffic"             | "AIRPORT EXTERIOR - SECURE AREA"                  | "Airport"                            |
| "Altering Medical Documents"                     | "Misuse Of Department Records"                    | "AIRPORT PARKING LOT"                             | "Airport"                            |
| "Amphetamines"                                   | "Positive Drug Screen - Originated From"          | "AIRPORT TERMINAL LOWER LEVEL - .."               | "Airport"                            |
| "An Act To Circumvent Criminal Prosecution"      | "Association With Felon"                          | "AIRPORT TERMINAL LOWER LEVEL - .."               | "Airport"                            |
| "Arrest, Improper Procedures"                    | "Court Attendance Irregularities"                 | "AIRPORT TERMINAL MEZZANINE - N.."                | "Airport"                            |
| "Arrestee-After Arrest, Prior To Lockup"         | "Arrest, Improper Procedures"                     | "AIRPORT TERMINAL UPPER LEVEL - .."               | "Airport"                            |
| "Arrestee-During Arrest"                         | "Excessive Force / On Duty - No Injury"           | "AIRPORT TERMINAL UPPER LEVEL - S.."              | "Airport"                            |
| "Arrestee-Lockup/Deletion"                       | "Excessive Force / On Duty - Injury"              | "AIRPORT/AIRCRAFT"                                | "Airport"                            |
| "Arson"  | "Excessive Force / Off Duty - Injury"             | "ALLEY"   | "Public Way - Other"                 |
| "Assault / Battery, Etc."                        | "Arson"   | "APARTMENT"                                       | "Private Residence"                  |
| "Association With Felon"                         | "Assault / Battery, Etc."                         | "APPLIANCE STORE"                                 | "Public Way - Other"                 |
| "Auto Theft"                                     | "Association With Felon"                          | "ATHLETIC CLUB"                                   | "Public Property - Other"            |
| "Bonding/Booking/Processing"                     | "Damage / Trespassing To Property"                | "BANK"  | "Public Way - Other"                 |
| "Bribe, Failure To Report"                       | "Bonding/Booking/Processing"                      | "BAR OR TAVERN"                                   | "Tavern/Liquor Store"                |
| "Burglary"                                       | "Bribe, Failure To Report"                        | "BARBERSHOP"                                      | "Public Way - Other"                 |
| "Choked"   | "Burglary"  | "BOAT/WATERCRAFT"                                 | "Public Property - Other"            |
| "Citation"                                       | "Excessive Force / On Duty - Injury"              | "BOWLING ALLEY"                                   | "Other Business Establishment"       |
| "Civil Suits - Third Party"                      | "Improper Processing / Reporting / Filing"        |   | "Public Way - Other"                 |
| "Closed Hand Strike (Punch)"                     | "Unnecessary Display Of Weapon / Carrying Weapon" | "CAR WASH"  | "Public Way - Other"                 |
| "Cocaine"  | "Excessive Force / On Duty - Injury"              | "CEMETARY"  | "Other Private Premise"              |
| "Coercion - Miscellaneous"                       | "Positive Drug Screen - Other Physical"           | "CHA APARTMENT"                                   | "Public Way - Other"                 |
| "Coercion - Threat Of Arrest/Charges"            | "Gang Affiliation"                                | "CHA HALLWAY/STAIRWELL/ELEVATOR"                  | "Public Way - Other"                 |
| "Coercion - Violence"                            | "Improper Search Of Vehicle"                      | "CHA PARKING LOT/GROUNDS"                         | "Public Way - Other"                 |
| "Compensatory Time"                              | "Gang Affiliation"                                | "CHURCH/SYNAGOGUE/PLACE OF WORSHIP"               | "Public Way - Other"                 |
| "Conduct An Improper / Inadequate Investigation" | "Compensatory Time"                               | "CLEANING STORE"                                  | "Other Business Establishment"       |
| "Conduct Unbecoming"                             | "Conduct An Improper / Inadequate Investigation"  |   | "Public Way - Other"                 |
| "Court Attendance Irregularities"                | "Neglect Of Duty"                                 | "COLLEGE/UNIVERSITY GROUNDS"                      | "Other Business Establishment"       |
| "Criminal Sexual Assault"                        | "Court Attendance Irregularities"                 | "COMMERCIAL / BUSINESS OFFICE"                    | "Other Business Establishment"       |
|  | "Criminal Sexual Assault"                         | "CONVENIENCE STORE"                               | "Other Business Establishment"       |

**Fig 3: Mapping table from old database to new database for allegation and location.**

Many of the allegations map to the same allegation in the new database but there are a significant number of changes in terminology. Some don't make obvious sense 'Not at home' to 'Seat Belts'. In fact, all of the following map to 'Seat Belts' : Not at home, Medical Roll Violation, Failure to Provide Documentation, Failure to Appear at Mss, Failure to receive approval for travel, timely notification to unit, and working while on the medical /IOD. **Seat Belts** seems to be a common cover for a certain set of allegations.

Below we show a small visual representation of the "other" mappings for each allegation and location. In other words, some allegations map to a distribution of allegations in the new database and while we can define the most common (and hence the likely 'unchanged' form of the allegation), we can also learn from how dispersed the changes were for a particular allegation (Fig 4).



**Fig 4: Distribution of mappings when the allegation was considered to be changed (violations of the mapping).**

Understanding these findings and taking this step set the stage for us to take a deeper look into what's driving these changes or associated with these changes (for instance, the officers involved).

By looking at the allegation category changes *by officer*, we were able to see the distribution of how allegations changed for particular officers (Fig 5). For example, Terry Abbate had CLOSED HAND STRIKE PUNCH change to several different **new** categories such as Miscellaneous, Fail to Obtain Complaint Register Number, etc. These seem like categories much less severe than CLOSED HAND STRIKE PUNCH! This can also disguise repeat offenders of certain misconduct, hiding systemic problems that a particular officer may have.

| id | first_name | last_name   | old_category                          | new_cat                                    | old_cat |
|----|------------|-------------|---------------------------------------|--|---------|
| 1  | Jeffery    | Aaron       | MISCELLANEOUS                         | Seat Belts                                 | 2       |
| 2  | Karina     | Aaron       | CLOSED HAND STRIKE PUNCH              | Excessive Force On Duty No Injury          | 30      |
| 3  | Anthony    | Abbate      | CLOSED HAND STRIKE PUNCH              | Unnecessary Display Of Weapon On Duty      | 6       |
| 4  | Terry      | Abbate      | CLOSED HAND STRIKE PUNCH              | Fail To Obtain A Complaint Register Number | 8       |
| 5  | Terry      | Abbate      | CLOSED HAND STRIKE PUNCH              | Miscellaneous                              | 8       |
| 6  | Terry      | Abbate      | CLOSED HAND STRIKE PUNCH              | Unnecessary Display Of Weapon On Duty      | 4       |
| 7  | Terry      | Abbate      | CLOSED HAND STRIKE PUNCH              | Association With Felon                     | 4       |
| 8  | Zaid       | Abdallah    | MISCELLANEOUS                         | Prisoners Property                         | 10      |
| 9  | Aziz       | Abdelmajeid | EXCESSIVE FORCE ON DUTY NO INJURY     | Excessive Force On Duty Injury             | 36      |
| 10 | Aziz       | Abdelmajeid | MISCELLANEOUS                         | Excessive Force On Duty No Injury          | 12      |
| 11 | Aziz       | Abdelmajeid | DOMESTIC ALTERCATION PHYSICAL ABUSE   | Domestic Incident Not Physical             | 8       |
| 12 | Moulay     | Abdullah    | MISCELLANEOUS                         | Miscellaneous                              | 5       |
| 13 | Moulay     | Abdullah    | INADEQUATE FAILURE TO PROVIDE SERVICE | Use Of Profanity                           | 1       |
| 14 | Moulay     | Abdullah    | MISCELLANEOUS                         | Sexual Orientation                         | 1       |
| 15 | Moulay     | Abdullah    | INTOXICATED ON DUTY                   | Fail To Obtain A Complaint Register Number | 1       |
| 16 | Moulay     | Abdullah    | ARRESTED DURING ARREST                | Racial Ethnic Etc                          | 1       |
| 17 | Jason      | Abjero      | MISCELLANEOUS                         | Miscellaneous                              | 2       |
| 18 | Henry      | Abrams      | NEGLECT OF DUTY                       | Indebtedness To City                       | 2       |

**Fig 5: Allegation changes by officer.**

## Graph Analytics

Using a graph analytics we tried to see what we could glean from this officer and investigator information. We used the mapping table joined back with the allegation data to construct our graph. Each vertex is an officer. Two vertices have an edge if they were listed on the same allegation (remember, only allegations where the mapping was violated, which means the allegation category was changed in a potentially suspicious way). This is basically just the co-accusal analysis but this only looks at allegations where the category changed in the police database.

Using a PageRank approach, we identified the most "important/influential" vertices in the graph. These are officers that are most frequently connected to other officers where the associated allegations have their category changed (Fig 6).

| id    | officer_fname | officer_lname | complaint_percentile | pagerank           |
|-------|---------------|---------------|----------------------|--------------------|
| 21309 | Eric          | Olson         | 94.46720123291016    | 14.6843363535723   |
| 28758 | Jesse         | Torres        | 43.90639877319336    | 14.484175385091378 |
| 17176 | Ricardo       | Mancha        | 83.57720184326172    | 14.033271002320387 |
| 27725 | Mark          | Struke        | 82.25520324707031    | 13.656611151876517 |
| 25980 | Roberto       | Sena          | 88.70120239257812    | 12.517037105678986 |
| 21941 | Daniel        | Passarelli    | 89.08329772949219    | 12.004761240422239 |
| 32086 | Shane         | Jones         | 97.64240264892578    | 11.045369399725754 |
| 11686 | Timothy       | Hawkins       | 87.7688980102539     | 10.949908884745193 |
| 31997 | Emerico       | Gonzalez      | 92.91960144042969    | 10.855702489793506 |
| 20177 | Nathaniel     | Myles         | 93.80229949951172    | 10.103644595649541 |
| 18384 | Kathleen      | Mc Lean       | 94.97540283203125    | 10.018741684885711 |
| 12440 | Terrance      | Hollis        | 27.52669906616211    | 9.624971127388415  |
| 32159 | Jorge         | Martinez      | 99.81659698486328    | 9.515512072442359  |
| 29379 | Artemio       | Vargas        | 74.81559753417969    | 9.110292972755232  |
| 15273 | Karl          | Kruger        | 88.70120239257812    | 9.067880632233914  |
| 29612 | Frank         | Villareal     | 99.2876968383789     | 8.949792463004878  |
| 13903 | Julie         | Joyce         | 71.59449768066406    | 8.797037105983549  |
| 27962 | John          | Swarbrick     | 77.83419799804688    | 8.618667197571575  |
| 3998  | Justin        | Carrillo      | 75.12129974365234    | 8.440248247900932  |
| 7273  | Daniel        | Dowling       | 84.99870300292969    | 8.247455360497561  |

only showing top 20 rows

**Fig 6: PageRank approach for displaying the top 20 officers connected in allegation changes.**

As one can see, the officers who have a high pagerank when allegations are changed also tend to have some of the highest complaint percentiles (often 90+). So there is something to be said in the sense that officers who are fielding some of the highest complaint percentiles are generally the same officers involved in changed allegations.

## Machine Learning & Text Analytics

Our latest analysis used NLP and text analytics to compare allegation category strings where the **mapping** was violated with cosine similarity to try to find the allegations where the category changed to something completely unrelated. We also compared the sentiment (a proxy for severity) of the old and new categories to see which allegations changed most in severity.



The main thing we found, which is consistent with a pattern we had noticed in previous analysis, is that the category: MISCELLANEOUS is commonly used to mask the content of allegations. Of course, the word miscellaneous is boring and uninformative, so it doesn't look bad at all.

|    | officer_id | officer_fname | officer_lname  | complaint_percentile | crid    | old_category       | new_allegation_name                            | allegation_mapping_boolean |
|----|------------|---------------|----------------|----------------------|---------|--------------------|--|----------------------------|
| 2  | 10322      | Joel          | Gonzalez       | 89.8819              | 1012611 | MISCELLANEOUS      | Inadequate Failure To Provide Service          | False                      |
| 3  | 31291      | Willard       | Wright         | 52.8214              | 1012622 | EEO INVESTIGATIONS | Miscellaneous                                  | False                      |
| 4  | 16729      | Jerry         | Luke           | 37.9466              | 1012622 | EEO INVESTIGATIONS | Miscellaneous                                  | False                      |
| 5  | 20033      | Kevin         | Murphy         | 80.4570              | 1012622 | EEO INVESTIGATIONS | Miscellaneous                                  | False                      |
| 6  | 31291      | Willard       | Wright         | 52.8214              | 1012622 | EEO INVESTIGATIONS | Miscellaneous                                  | False                      |
| 7  | 16729      | Jerry         | Luke           | 37.9466              | 1012622 | EEO INVESTIGATIONS | Miscellaneous                                  | False                      |
| 8  | 20033      | Kevin         | Murphy         | 80.4570              | 1012622 | EEO INVESTIGATIONS | Miscellaneous                                  | False                      |
| 9  | 13604      | Brian         | Johnson        | 97.5202              | 1012629 | CHOKED             | Excessive Force On Duty No Injury              | False                      |
| 10 | 7233       | John          | Dougherty      | 92.7369              | 1012629 | CHOKED             | Excessive Force On Duty No Injury              | False                      |
| 11 | 13604      | Brian         | Johnson        | 97.5202              | 1012629 | CHOKED             | Excessive Force On Duty No Injury              | False                      |
| 12 | 7233       | John          | Dougherty      | 92.7369              | 1012629 | CHOKED             | Excessive Force On Duty No Injury              | False                      |
| 13 | 30336      | Kenneth       | Weller         | 82.4004              | 1012648 | MISCELLANEOUS      | Excessive Force On Duty No Injury              | False                      |
| 14 | 15657      | Timothy       | Lange          | 64.2047              | 1012648 | MISCELLANEOUS      | Excessive Force On Duty No Injury              | False                      |
| 16 | 27997      | Brian         | Swiatkowski    | 98.5896              | 1012654 | PUSHPULLGRAB       | Use Of Profanity                               | False                      |
| 17 | 5328       | Markee        | Cooper         | 95.2734              | 1012654 | PUSHPULLGRAB       | Use Of Profanity                               | False                      |
| 18 | 12177      | John          | Higgins        | 88.5216              | 1012654 | PUSHPULLGRAB       | Use Of Profanity                               | False                      |
| 19 | 27997      | Brian         | Swiatkowski    | 98.5896              | 1012654 | PUSHPULLGRAB       | Unnecessary Physical Contact On Duty No Injury | False                      |
| 20 | 5328       | Markee        | Cooper         | 95.2734              | 1012654 | PUSHPULLGRAB       | Unnecessary Physical Contact On Duty No Injury | False                      |
| 21 | 12177      | John          | Higgins        | 88.5216              | 1012654 | PUSHPULLGRAB       | Unnecessary Physical Contact On Duty No Injury | False                      |
| 22 | 27997      | Brian         | Swiatkowski    | 98.5896              | 1012654 | PUSHPULLGRAB       | Fail To Obtain A Complaint Register Number     | False                      |
| 23 | 5328       | Markee        | Cooper         | 95.2734              | 1012654 | PUSHPULLGRAB       | Fail To Obtain A Complaint Register Number     | False                      |
| 24 | 12177      | John          | Higgins        | 88.5216              | 1012654 | PUSHPULLGRAB       | Fail To Obtain A Complaint Register Number     | False                      |
| 25 | 23190      | Hakeem        | Qazi           | 80.8987              | 1012689 | PUSHPULLGRAB       | Excessive Force On Duty No Injury              | False                      |
| 26 | 909        | Lindbergh     | Askew          | 78.5694              | 1012690 | MISCELLANEOUS      | SecondarySpecial Employment                    | False                      |
| 27 | 1360       | Donna         | Barnes Simmons | 49.6573              | 1012691 | MISCELLANEOUS      | SecondarySpecial Employment                    | False                      |
| 28 | 1586       | Ronald        | Beach          | 56.9371              | 1012692 | MISCELLANEOUS      | Seat Belts                                     | False                      |
| 29 | 3914       | Dena          | Carli          | 33.6288              | 1012693 | MISCELLANEOUS      | SecondarySpecial Employment                    | False                      |
| 30 | 9019       | Derek         | Fowler         | 69.2513              | 1012694 | MISCELLANEOUS      | Seat Belts                                     | False                      |
| 31 | 13939      | Harold        | Junior         | 51.4165              | 1012695 | MISCELLANEOUS      | Seat Belts                                     | False                      |
| 32 | 18213      | Anthony       | Mc Gowan       | 63.1462              | 1012696 | MISCELLANEOUS      | Seat Belts                                     | False                      |

**Fig 7 - Looking at allegations and the drastic changes in meaning**

One can see "MISCELLANEOUS" showing up often under **old\_category** and "Miscellaneous" showing up under **new\_allegation\_name**. There are still quite a few interesting ones, such as "PUSHPULLGRAB" being changed to "Fail To Obtain A Complaint Register Number." Clearly these are unrelated and the severity of these is different.



Next we thought it would be interesting to try and see which allegations were having their categories changed to something “less severe.” We decided to use *sentiment* as a proxy for severity: if the sentiment is more positive, the category is less severe (and vice versa).

### Sentiment (most allegations will probably be negative 🙌)

- Excessive Force:  $-.25$
- Unnecessary Physical Contact On Duty No Injury:  $-.2$
- Use Of Profanity:  $0$
- Seat Belts:  $0$
- We will compare how the sentiment (proxy for severity) changed from the old allegation categories to the new allegation categories

One can see “Excessive Force” has negative sentiment (more negative than Unnecessary Physical Contact which makes sense). Use of Profanity and Seat Belts are considered neutral. We then calculated the absolute difference between the old category and the new category, getting an idea of how the sentiment (severity) changed.

|     | officer_id | officer_fname | officer_lname | complaint_percentile | crid    | old_category                          | new_allegation_name               |
|-----|------------|---------------|---------------|----------------------|---------|---------------------------------------|-----------------------------------|
| 334 | 5323       | Ronald        | Cooper        | 74.7927              | 1014535 | CONDUCT UNBECOMING                    | Sexual Orientation                |
| 335 | 18907      | Tomi          | Methipara     | 73.2184              | 1014535 | CONDUCT UNBECOMING                    | Sexual Orientation                |
| 336 | 5323       | Ronald        | Cooper        | 74.7927              | 1014535 | CONDUCT UNBECOMING                    | Sexual Orientation                |
| 337 | 18907      | Tomi          | Methipara     | 73.2184              | 1014535 | CONDUCT UNBECOMING                    | Sexual Orientation                |
| 702 | 28925      | Joseph        | Tripoli       | 84.5898              | 1017193 | RACIAL ETHNIC ETC                     | Sexual Orientation                |
| 703 | 8592       | Frank         | Fish          | 52.0271              | 1017193 | RACIAL ETHNIC ETC                     | Sexual Orientation                |
| 762 | 24506      | Stan          | Rogers        | 74.7851              | 1017952 | SEXUAL ORIENTATION                    | Excessive Force On Duty No Injury |
| 763 | 16660      | Lawrence      | Lowrey        | 71.4570              | 1017952 | SEXUAL ORIENTATION                    | Excessive Force On Duty No Injury |
| 764 | 13796      | Nancy         | Jones         | 53.9163              | 1017952 | SEXUAL ORIENTATION                    | Excessive Force On Duty No Injury |
| 765 | 23351      | Brandon       | Rambert       | 74.5214              | 1017952 | SEXUAL ORIENTATION                    | Excessive Force On Duty No Injury |
| 798 | 27725      | Mark          | Struke        | 82.2552              | 1018181 | INADEQUATE FAILURE TO PROVIDE SERVICE | Properly Direct Subordinate       |
| 799 | 13909      | Ted           | Jozefczak     | 95.2046              | 1018181 | INADEQUATE FAILURE TO PROVIDE SERVICE | Properly Direct Subordinate       |
| 800 | 17134      | Melissa       | Malm          | 78.4953              | 1018181 | INADEQUATE FAILURE TO PROVIDE SERVICE | Properly Direct Subordinate       |
| 818 | 13637      | Robert        | Johnson       | 95.8695              | 1018246 | RACIAL ETHNIC ETC                     | Sexual Orientation                |
| 819 | 32383      | Armando       | Ugarte        | 99.6981              | 1018246 | RACIAL ETHNIC ETC                     | Sexual Orientation                |

It looks like many of the major differences have to do with Sexual Orientation. In this analysis, we did not find particular allegations that were grossly different. Given enough time, further analysis could be done where one uses a different sentiment analysis model, or uses the

[Severity Analysis](#) done by the Fearless Hamsters. Using the **Severity Analysis**, one could truly compare whether allegation categories were changed from something severe to something less severe (like "Seat Belts" 😞)

## Conclusion

We learned a lot in this project. Not only did we learn new tools and analytical methods, but we learned a lot about the Chicago Police and the CPDP project. A large percentage of allegations have their category strings changed (in our estimation, about  $\frac{1}{3}$ ). Most of the changes are benign, but there is a subset of allegations that change to a category that is completely different. There are others that changed to something much less severe. Why is this? Are some of these clerical errors? Is this a technique the Department uses to mask the severity of certain allegations?

We noticed that "Miscellaneous" is used as a catch-all term. There are some very serious allegations that are being defined as "Miscellaneous." The term is vague and non-descriptive and also masks the severity of certain allegations. Another common cover is "Seat Belts." This category is used in CPDB often when the previous allegation category is completely unrelated.

In closing, we believe we have just scratched the surface and we hope that our work is a baseline for further investigation. Hopefully those with more knowledge about the project can read our analysis and either make confirmations, or come up with new questions that can be answered using our data and analysis.

## Future Work

In each checkpoint, we mention ways to expand on the questions we asked for the analysis. Overall, the most promising future analysis would be integrating the [Severity Analysis](#) with our **mapping** data. As opposed to doing sentiment analysis, we could truly see if allegations were changed to a category that is less severe. This would allow the Invisible Institute to have an even smaller subset of allegations where the categories changed, ones in which the severity of the allegation was altered.