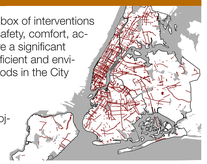
# STREET IMPROVEMENT PROJECT TRACKING APPLICATION: USING A HIERARCHICAL DATABASE MODEL

## Street Improvement Projects

Street Improvement Projects (SIP) utilize a broad toolbox of interventions to reengineer city streets with the goal of improving safety, comfort, accessibility, and mobility for all roadway users. They are a significant part of NYC DOT's mission to provide for the safe, efficient and environmentally responsible movement of people and goods in the City of New York.



NYC DOT implements more than 100 distinct SIP projects per year, carried out by 8 different units.

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Develop-

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementa-

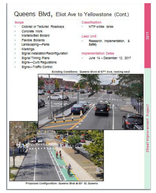
## Historic Tracking

Historically, SIP projects were tracked for the program in a master tracking sheet and at the project level in a summary document.

| Excel tracking sheet | PDF project description |
|---|---|
| Maintained centrally by program manager | Maintained by project manager |
| Project name | Project background |
| One-line project description | High level description |
| Implementation dates | High level treatment descriptions |
| Project manager name and unit | Project images and map |

Data Problems:
- Not Query-able
- Limited to High-Profile Treatments
- Inconsistant Treatment Discription

Geographic Problems:
- No Spatial Data
- Imprecise Project Extents
- No Treatment Extents



This system made it very challenging and labor intensive to answer questions about the program as a whole or even specific projects. Finding all of the projects that contained a specific treatment (for example) became an exercise in archeology and institutional memory. Programmatic analysis was nearly impossible.

NYC DOT decided to create a new system to track the SIPs that utilized a distributed data entry process.

## Database Model

### Database Design:

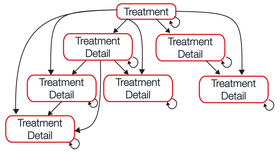
SIP project tracking system requirements:

- Record all project and treatment details
- Account for different project types
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate the adoption of new treatments and modifications to existing treatment options
- Data entry must be possible by the Project Manager and not require specialized skills or software
- Data must be query-able and accessible to analysts

Closure database model:

- Hierarchical tree structure representation of the data
- Stores relationships between each node to each of its descendants (including itself)
- Data is stored in two tables (data and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model




### Implementation:

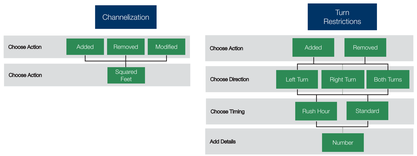
Projects are a collection of treatments:

- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study and simple standardized
- Do not require Closure model

Treatments are a standardized collection of decisions:

- Each potential treatment must be abstracted into a series of decisions.
- The decision trees that represent the various treatments do not need tohave a uniform breadth or depth.
- Treatment trees define what is possible to record for any treatment
- Treatment trees may be changed to reflect new choices

Treatment Hierarchy Structures




Treatment decision trees are converted into the closure tables

- Each possible decision is represented in the Questions table.
- The relationships between the every decision in the treatment trees are added to the Question Path table.
- Each node and each questions path is unique.
- A 3rd attribute table is added to allow for freeform data fields (such as square feet), these fields must always be the last node in the tree.
- Treatment trees define the possible data can be entered for a project.

| Questions | |
|---|---|
| Question ID | Question |
| 1 | Bike Facility |
| 2 | Protected |
| 3 | Standard |
| 4 | Regular |
| 5 | Curbside |
| 6 | One-Way |
| 7 | Two-Way |
| 8 | One-Way |
| 9 | Two-Way |
| 10 | Parking |
| 11 | Jersy Barrier |
| 12 | Delinlator |

| Question Path | | |
|---|---|---|
| Ancestor | Descendant | Depth |
| 1 | 1 | 0 |
| 1 | 2 | 1 |
| 1 | 3 | 1 |
| 1 | 4 | 2 |
| 1 | 5 | 2 |
| 1 | 6 | 3 |
| 1 | 7 | 3 |
| 1 | 8 | 3 |
| 1 | 9 | 3 |
| 1 | 10 | 2 |
| 1 | 11 | 2 |
| 1 | 12 | 2 |
| 2 | 2 | 0 |
| 2 | 10 | 1 |
| 2 | 11 | 1 |
| 2 | 12 | 1 |
| 3 | 3 | 0 |
| 3 | 4 | 1 |
| 3 | 5 | 1 |
| 3 | 6 | 2 |
| 3 | 7 | 2 |
| 3 | 8 | 2 |
| 3 | 9 | 2 |
| 4 | 4 | 0 |
| 4 | 8 | 1 |
| 4 | 9 | 1 |
| 8 | 8 | 0 |
| 9 | 9 | 0 |

### Recording Treatments:

Project data is recorded in two tables: project attribute table and project treatments table.

- The last decision point that the user passed though is what is recorded in the project treatments table.
- All treatments, regardless of complexity are stored in a single row in the project treatments table.
- The data stored in the project treatments table requires joining to the decision and decision path tables to be understood.

## Data Exploration

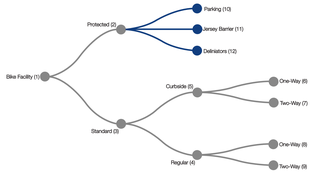
### Exploration in the database:

This data structure is simple to work with in the database. It provides a lot of flexibility for querying at different levels of granularity or recreating the full treatment details from the last decision node.

The data structure allows for ad hock groupings within the hierarchy. Queries down the tree return all of the choices below a given node. This is used to find all treatments that include at least a particular decision. For example: Finding all protected bike paths, regardless of the specific details of the subsequent decisions.

```
SELECT q.*
FROM sip_questions q
JOIN (
   SELECT descendant d, depth
   FROM sip_questions_path
   -- All treatments with a protected bike facility
   WHERE ancestor = 2
) AS p ON q.qid = p.d
ORDER BY p.depth;
```

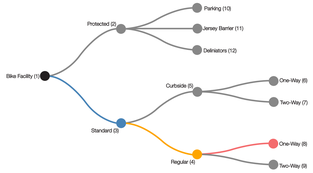
| qid integer | question text |
|---|---|
| 2 | Protected |
| 10 | Parking |
| 11 | Jersey Barrier |
| 12 | Delinators |




The treatment attributes that are recorded in association with a project are limited to the final decision node the user passed through. The ability to interpret the treatment relies on being able to recreate the full set of choices the user made. Recreating the full decision path of a treatment detail necessitates querying up the tree. For example: One-way (QID 8) is only a meaningful treatment description once the full path back to the root is reviled as a One-way standard non-curbside bike facility.

```
SELECT q.*
FROM sip_questions q
JOIN (
   SELECT ancestor a, depth
   FROM sip_questions_path
   -- All treatments with a one-way
      standard bike facility
   WHERE descendant = 8
) AS p ON q.qid = p.a
ORDER BY p.depth desc;
```

| qid integer | question text |
|---|---|
| 1 | Bike Facility |
| 3 | Standard |
| 4 | Regular |
| 8 | One-way |

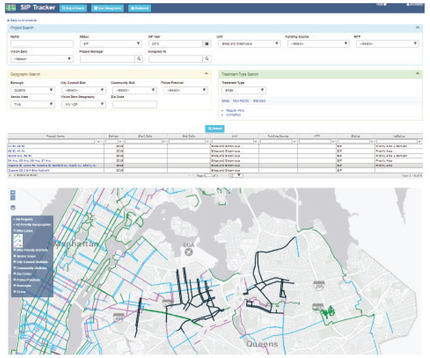



### Exploration in the application:

The application provides planners, program managers, analysts and outreach/communications staff with the ability to quickly and easily explore the data.

Project managers need to know what other projects are being developed and/or implemented near their project. To facilitate this, the application provides a number of reference layers that can be toggled in the map provide an overlay of:

- All projects (complete and under development)
- Political and administrative districts
- Priority geographies
- Existing bicycle network



For a more targeted search, the application provides a query tool that allows for searching for projects using three categories of criteria

- Project
- Geography
- Treatment

The treatment search utilizes the treatment decision trees to allow user defined specificity.

The results of the search in both the table and the map link to the projects' full detail page.

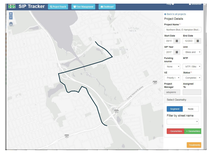
## Data Entry

The primary purpose of the application is to facilitate the collection and storage of the project and treatment attributes. Data entry is intended to be done by the project manager.

The data entry process is straightforward and consists of two levels, project and treatment.

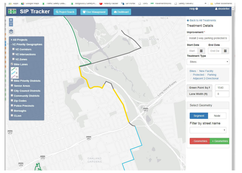
Project data entry process:

1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography



Treatment data entry process:

1. Enter treatment level attributes
2. Enter treatment type details via decision tree
3. Select treatment geography
4. Repeat for all other treatments associated with the project

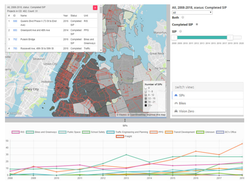


## Applications

### Planning

The application includes a dashboard to help unit directors to explore their group's work. NYC DOT now uses the dashboard as a tool to plan future work programs at a high level

The dashboard provides a powerful tool to ensure that communities were not over nor underrepresented in the SIP program's work portfolio.



### Analysis

NYC DOT has begun to analyze the equitability of the SIP program as a whole to ensure all New Yorkers were being served. This analysis explores the geographic distribution of projects throughout the city across a number of different demographic metrics.

NYC DOT has begun to analyze the full SIP work program's efficacy from a safety perspective. Every SIP project is, analyzed individually to ensure they are improving safety, but a holistic analysis has not historically been practical.

## Credits

Project Team:

Jason Fitzsimmons
Arthur Getman
Zhi Keng He
Seth Hostetter
Chandrima Pal
Sidney Salnave
Tanuja Ramesh Vanjarapu

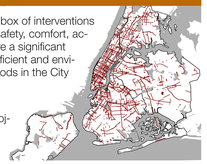
Technologies:

PostgreSQL / PostGIS
C# .Net
Geoserver v2.8
OpenLayers 3
ESRI map services

# STREET IMPROVEMENT PROJECT TRACKING APPLICATION: USING A HIERARCHICAL DATABASE MODEL

## Street Improvement Projects

Street Improvement Projects (SIP) utilize a broad toolbox of interventions to reengineer city streets with the goal of improving safety, comfort, accessibility, and mobility for all roadway users. They are a significant part of NYC DOT's mission to provide for the safe, efficient and environmentally responsible movement of people and goods in the City of New York.



NYC DOT implements more than 100 distinct SIP projects per year, carried out by 8 different units.

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Develop-

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementa-

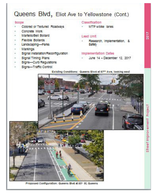
## Historic Tracking

Historically, SIP projects were tracked for the program in a master tracking sheet and at the project level in a summary document.

| Excel tracking sheet | PDF project description |
|---|---|
| Maintained centrally by program manager | Maintained by project manager |
| Project name | Project background |
| One-line project description | High level description |
| Implementation dates | High level treatment descriptions |
| Project manager name and unit | Project images and map |

Data Problems:
- Not Query-able
- Limited to High-Profile Treatments
- Inconsistant Treatment Discription

Geographic Problems:
- No Spatial Data
- Imprecise Project Extents
- No Treatment Extents



This system made it very challenging and labor intensive to answer questions about the program as a whole or even specific projects. Finding all of the projects that contained a specific treatment (for example) became an exercise in archeology and institutional memory. Programmatic analysis was nearly impossible.

NYC DOT decided to create a new system to track the SIPs that utilized a distributed data entry process.

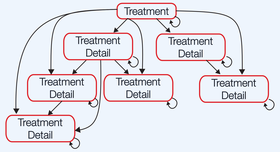
## Database Model

### Database Design:

SIP project tracking system requirements:
- Record all project and treatment details
- Account for different project types
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate the adoption of new treatments and modifications to existing treatment options
- Data entry must be possible by the Project Manager and not require specialized skills or software
- Data must be query-able and accessible to analysts

Closure database model:
- Hierarchical tree structure representation of the data
- Stores relationships between each node to each of its descendants (including itself)
- Data is stored in two tables (data and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model




### Implementation:

Projects are a collection of treatments:
- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study and simple standardized
- Do not require Closure model

Treatments are a standardized collection of decisions:
- Each potential treatment must be abstracted into a series of decisions.
- The decision trees that represent the various treatments do not need tohave a uniform breadth or depth.
- Treatment trees define what is possible to record for any treatment
- Treatment trees may be changed to reflect new choices

Treatment Hierarchy Structures

Treatment decision trees are converted into the closure tables
- Each possible decision is represented in the Questions table.
- The relationships between the every decision in the treatment trees are added to the Question Path table.
- Each node and each questions path is unique.
- A 3rd attribute table is added to allow for freeform data fields (such as square feet), these fields must always be the last node in the tree.
- Treatment trees define the possible data can be entered for a project.

| Questions | |
|---|---|
| Question ID | Question |
| 1 | Bike Facility |
| 2 | Protected |
| 3 | Standard |
| 4 | Regular |
| 5 | Curbside |
| 6 | One-Way |
| 7 | Two-Way |
| 8 | One-Way |
| 9 | Two-Way |
| 10 | Parking |
| 11 | Jersy Barrier |
| 12 | Delinlator |

| Question Path | | |
|---|---|---|
| Ancestor | Descendant | Depth |
| 1 | 1 | 0 |
| 1 | 2 | 1 |
| 1 | 3 | 1 |
| 1 | 4 | 2 |
| 1 | 5 | 2 |
| 1 | 6 | 3 |
| 1 | 7 | 3 |
| 1 | 8 | 3 |
| 1 | 9 | 3 |
| 1 | 10 | 2 |
| 1 | 11 | 2 |
| 1 | 12 | 2 |
| 2 | 2 | 0 |
| 2 | 10 | 1 |
| 2 | 11 | 1 |
| 2 | 12 | 1 |
| 3 | 3 | 0 |
| 3 | 4 | 1 |
| 3 | 5 | 1 |
| 3 | 6 | 2 |
| 3 | 7 | 2 |
| 3 | 8 | 2 |
| 3 | 9 | 2 |
| 4 | 4 | 0 |
| 4 | 8 | 1 |
| 4 | 9 | 1 |
| 8 | 8 | 0 |
| 9 | 9 | 0 |

### Recording Treatments:

Project data is recorded in two tables: project attribute table and project treatments table.
- The last decision point that the user passed though is what is recorded in the project treatments table.
- All treatments, regardless of complexity are stored in a single row in the project treatments table.
- The data stored in the project treatments table requires joining to the decision and decision path tables to be understood.

USER INPUT →

| Project | Treatment |
|---|---|
| 1 | 7 |
| 1 | 10 |
| 2 | 5 |

→ Two-Way + Curbside + Standard + Bike Facility
→ Parking + Protected + Bike Facility
→ Curbside + Standard + Bike Facility

## Data Exploration

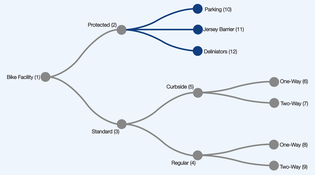
### Exploration in the database:

This data structure is simple to work with in the database. It provides a lot of flexibility for querying at different levels of granularity or recreating the full treatment details from the last decision node.

The data structure allows for ad hock groupings within the hierarchy. Queries down the tree return all of the choices below a given node. This is used to find all treatments that include at least a particular decision. For example: Finding all protected bike paths, regardless of the specific details of the subsequent decisions.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT descendant d, depth
    FROM sip_questions_path
    -- All treatments with a protected bike facility
    WHERE ancestor = 2
) AS p ON q.qid = p.d
ORDER BY p.depth;
```

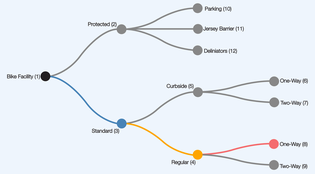
| qid integer | question text |
|---|---|
| 1 | Protected |
| 2 | Parking |
| 3 | Jersey Barrier |
| 4 | Delinators |




The treatment attributes that are recorded in association with a project are limited to the final decision node the user passed through. The ability to interpret the treatment relies on being able to recreate the full set of choices the user made. Recreating the full decision path of a treatment detail necessitates querying up the tree. For example: One-way (QID 8) is only a meaningful treatment description once the full path back to the root is revileed as a One-way standard non-curbside bike facility.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT ancestor a, depth
    FROM sip_questions_path
    -- All treatments with a one-way
       standard bike facility
    WHERE descendant = 8
  ) AS p ON q.qid = p.a
ORDER BY p.depth desc;
```

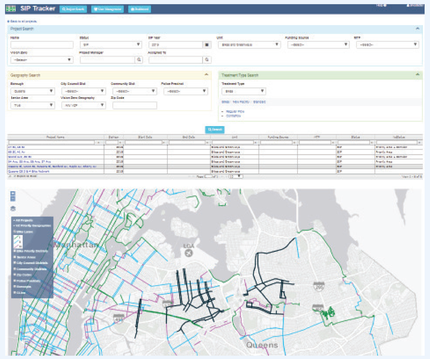
| qid integer | question text |
|---|---|
| 1 | Bike Facility |
| 2 | Standard |
| 3 | Regular |
| 4 | One-way |




### Exploration in the application:

The application provides planners, program managers, analysts and outreach/communications staff with the ability to quickly and easily explore the data.

Project managers need to know what other projects are being developed and/or implemented near their project. To facilitate this, the application provides a number of reference layers that can be toggled in the map provide an overlay of:
- All projects (complete and under development)
- Political and administrative districts
- Priority geographies
- Existing bicycle network



For a more targeted search, the application provides a query tool that allows for searching for projects using three categories of criteria
- Project
- Geography
- Treatment

The treatment search utilizes the treatment decision trees to allow user defined specificity.

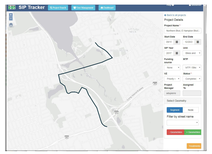
The results of the search in both the table and the map link to the projects' full detail page.

## Data Entry

The primary purpose of the application is to facilitate the collection and storage of the project and treatment attributes. Data entry is intended to be done by the project manager.

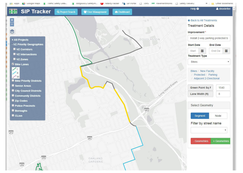
The data entry process is straightforward and consists of two levels, project and treatment.

Project data entry process:
1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography



Treatment data entry process:

1. Enter treatment level attributes
2. Enter treatment type details via decision tree
3. Select treatment geography
4. Repeat for all other treatments associated with the project

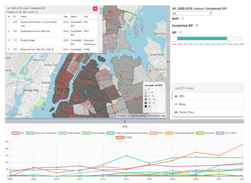


## Applications

### Planning

The application includes a dashboard to help unit directors to explore their group's work. NYC DOT now uses the dashboard as a tool to plan future work programs at a high level

The dashboard provides a powerful tool to ensure that communities were not over nor underrepresented in the SIP program's work portfolio.



### Analysis

NYC DOT has begun to analyze the equitability of the SIP program as a whole to ensure all New Yorkers were being served. This analysis explores the geographic distribution of projects throughout the city across a number of different demographic metrics.

NYC DOT has begun to analyze the full SIP work program's efficacy from a safety perspective. Every SIP project is, analyzed individually to ensure they are improving safety, but a holistic analysis has not historically been practical.

## Credits

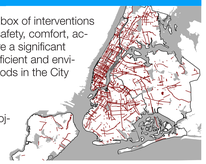
Project Team:
Jason Fitzsimmons
Arthur Getman
Zhi Keng He
Seth Hostetter
Chandrima Pal
Sidney Salnave
Tanuja Ramesh Vanjarapu

Technologies:
PostgreSQL / PostGIS
C# .Net
Geoserver v2.8
OpenLayers 3
ESRI map services

# STREET IMPROVEMENT PROJECT TRACKING APPLICATION: USING A HIERARCHICAL DATABASE MODEL

## Street Improvement Projects

Street Improvement Projects (SIP) utilize a broad toolbox of interventions to reengineer city streets with the goal of improving safety, comfort, accessibility, and mobility for all roadway users. They are a significant part of NYC DOT's mission to provide for the safe, efficient and environmentally responsible movement of people and goods in the City of New York.



NYC DOT implements more than 100 distinct SIP projects per year, carried out by 8 different units.

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Develop-

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementa-

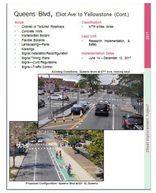
## Historic Tracking

Historically, SIP projects were tracked for the program in a master tracking sheet and at the project level in a summary document.

| Excel tracking sheet | PDF project description |
|---|---|
| Maintained centrally by program manager | Maintained by project manager |
| Project name | Project background |
| One-line project description | High level description |
| Implementation dates | High level treatment descriptions |
| Project manager name and unit | Project images and map |

Data Problems:
- Not Query-able
- Limited to High-Profile Treatments
- Inconsistant Treatment Discription

Geographic Problems:
- No Spatial Data
- Imprecise Project Extents
- No Treatment Extents




This system made it very challenging and labor intensive to answer questions about the program as a whole or even specific projects. Finding all of the projects that contained a specific treatment (for example) became an exercise in archeology and institutional memory. Programmatic analysis was nearly impossible.

NYC DOT decided to create a new system to track the SIPs that utilized a distributed data entry process.

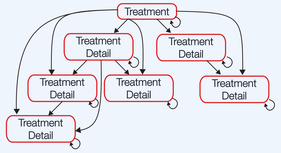
## Database Model

### Database Design:

SIP project tracking system requirements:
- Record all project and treatment details
- Account for different project types
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate the adoption of new treatments and modifications to existing treatment options
- Data entry must be possible by the Project Manager and not require specialized skills or software
- Data must be query-able and accessible to analysts

Closure database model:
- Hierarchical tree structure representation of the data
- Stores relationships between each node to each of its descendants (including itself)
- Data is stored in two tables (data and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model



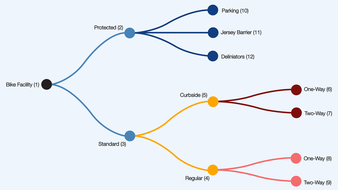

### Implementation:

Projects are a collection of treatments:
- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study and simple standardized
- Do not require Closure model

Treatments are a standardized collection of decisions:
- Each potential treatment must be abstracted into a series of decisions.
- The decision trees that represent the various treatments do not need tohave a uniform breadth or depth.
- Treatment trees define what is possible to record for any treatment
- Treatment trees may be changed to reflect new choices

Treatment Hierarchy Structures

Treatment decision trees are converted into the closure tables
- Each possible decision is represented in the Questions table.
- The relationships between the every decision in the treatment trees are added to the Question Path table.
- Each node and each questions path is unique.
- A 3rd attribute table is added to allow for freeform data fields (such as square feet), these fields must always be the last node in the tree.
- Treatment trees define the possible data can be entered for a project.




| Questions | |
|---|---|
| Question ID | Question |
| 1 | Bike Facility |
| 2 | Protected |
| 3 | Standard |
| 4 | Regular |
| 5 | Curbside |
| 6 | One-Way |
| 7 | Two-Way |
| 8 | One-Way |
| 9 | Two-Way |
| 10 | Parking |
| 11 | Jersy Barrier |
| 12 | Delinlator |

| Question Path | | |
|---|---|---|
| Ancestor | Descendant | Depth |
| 1 | 1 | 0 |
| 1 | 2 | 1 |
| 1 | 3 | 1 |
| 1 | 4 | 2 |
| 1 | 5 | 2 |
| 1 | 6 | 3 |
| 1 | 7 | 3 |
| 1 | 8 | 3 |
| 1 | 9 | 3 |
| 1 | 10 | 2 |
| 1 | 11 | 2 |
| 1 | 12 | 2 |
| 2 | 2 | 0 |
| 2 | 10 | 1 |
| 2 | 11 | 1 |
| 2 | 12 | 1 |
| 3 | 3 | 0 |
| 3 | 4 | 1 |
| 3 | 5 | 1 |
| 3 | 6 | 2 |
| 3 | 7 | 2 |
| 3 | 8 | 2 |
| 3 | 9 | 2 |
| 4 | 4 | 0 |
| 4 | 8 | 1 |
| 4 | 9 | 1 |
| 8 | 8 | 0 |
| 9 | 9 | 0 |

### Recording Treatments:

Project data is recorded in two tables: project attribute table and project treatments table.
- The last decision point that the user passed though is what is recorded in the project treatments table.
- All treatments, regardless of complexity are stored in a single row in the project treatments table.
- The data stored in the project treatments table requires joining to the decision and decision path tables to be understood.

USER INPUT →

| Project | Treatment | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | = | Two-Way | + | Curbside | + | Standard | + | Bike Facility |
| 1 | 10 | = | Parking | + | Protected | + | Bike Facility | | |
| 2 | 5 | = | Curbside | + | Standard | + | Bike Facility | | |

## Data Exploration

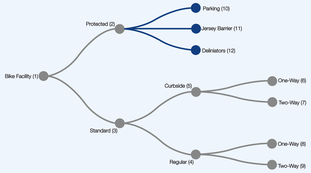
### Exploration in the database:

This data structure is simple to work with in the database. It provides a lot of flexibility for querying at different levels of granularity or recreating the full treatment details from the last decision node.

The data structure allows for ad hock groupings within the hierarchy. Queries down the tree return all of the choices below a given node. This is used to find all treatments that include at least a particular decision. For example: Finding all protected bike paths, regardless of the specific details of the subsequent decisions.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT descendant d, depth
    FROM sip_questions_path
    -- All treatments with a protected bike facility
    WHERE ancestor = 2
) AS p ON q.qid = p.d
ORDER BY p.depth;
```

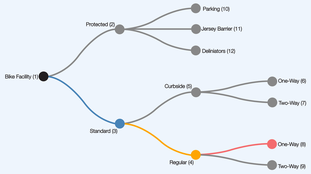
| qid integer | question text |
|---|---|
| 1 | 2 Protected |
| 2 | 10 Parking |
| 3 | 11 Jersey Barrier |
| 4 | 12 Delinators |




The treatment attributes that are recorded in association with a project are limited to the final decision node the user passed through. The ability to interpret the treatment relies on being able to recreate the full set of choices the user made. Recreating the full decision path of a treatment detail necessitates querying up the tree. For example: One-way (QID 8) is only a meaningful treatment description once the full path back to the root is reviled as a One-way standard non-curbside bike facility.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT ancestor a, depth
    FROM sip_questions_path
    -- All treatments with a one-way
       standard bike facility
    WHERE descendant = 8
  ) AS p ON q.qid = p.a
ORDER BY p.depth desc;
```

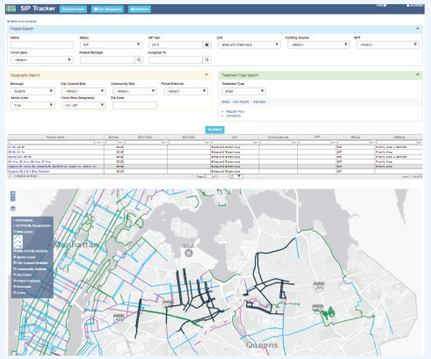
| qid integer | question text |
|---|---|
| 1 | 1 Bike Facility |
| 2 | 3 Standard |
| 3 | 4 Regular |
| 4 | 8 One-way |




### Exploration in the application:

The application provides planners, program managers, analysts and outreach/communications staff with the ability to quickly and easily explore the data.

Project managers need to know what other projects are being developed and/or implemented near their project. To facilitate this, the application provides a number of reference layers that can be toggled in the map provide an overlay of:
- All projects (complete and under development)
- Political and administrative districts
- Priority geographies
- Existing bicycle network




For a more targeted search, the application provides a query tool that allows for searching for projects using three categories of criteria
- Project
- Geography
- Treatment

The treatment search utilizes the treatment decision trees to allow user defined specificity.

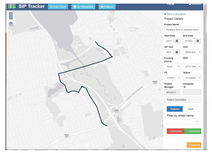
The results of the search in both the table and the map link to the projects' full detail page.

## Data Entry

The primary purpose of the application is to facilitate the collection and storage of the project and treatment attributes. Data entry is intended to be done by the project manager.

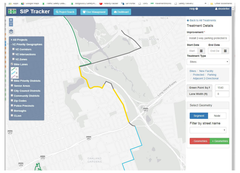
The data entry process is straightforward and consists of two levels, project and treatment.

Project data entry process:
1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography




Treatment data entry process:

1. Enter treatment level attributes
2. Enter treatment type details via decision tree
3. Select treatment geography
4. Repeat for all other treatments associated with the project

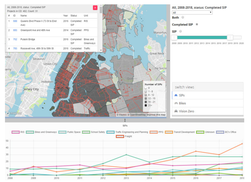



## Applications

### Planning

The application includes a dashboard to help unit directors to explore their group's work. NYC DOT now uses the dashboard as a tool to plan future work programs at a high level

The dashboard provides a powerful tool to ensure that communities were not over nor underrepresented in the SIP program's work portfolio.



### Analysis

NYC DOT has begun to analyze the equitability of the SIP program as a whole to ensure all New Yorkers were being served. This analysis explores the geographic distribution of projects throughout the city across a number of different demographic metrics.

NYC DOT has begun to analyze the full SIP work program's efficacy from a safety perspective. Every SIP project is, analyzed individually to ensure they are improving safety, but a holistic analysis has not historically been practical.

## Credits

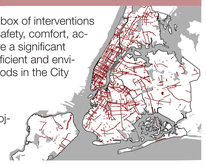
Project Team:
- Jason Fitzsimmons
- Arthur Getman
- Zhi Keng He
- Seth Hostetter
- Chandrima Pal
- Sidney Salnave
- Tanuja Ramesh Vanjarapu

Technologies:
- PostgreSQL / PostGIS
- C# .Net
- Geoserver v2.8
- OpenLayers 3
- ESRI map services

# STREET IMPROVEMENT PROJECT TRACKING APPLICATION: USING A HIERARCHICAL DATABASE MODEL

## Street Improvement Projects

Street Improvement Projects (SIP) utilize a broad toolbox of interventions to reengineer city streets with the goal of improving safety, comfort, accessibility, and mobility for all roadway users. They are a significant part of NYC DOT's mission to provide for the safe, efficient and environmentally responsible movement of people and goods in the City of New York.



NYC DOT implements more than 100 distinct SIP projects per year, carried out by 8 different units.

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Develop-

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementa-

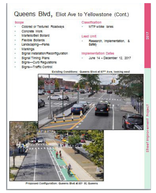
## Historic Tracking

Historically, SIP projects were tracked for the program in a master tracking sheet and at the project level in a summary document.

| Excel tracking sheet | PDF project description |
|---|---|
| Maintained centrally by program manager | Maintained by project manager |
| Project name | Project background |
| One-line project description | High level description |
| Implementation dates | High level treatment descriptions |
| Project manager name and unit | Project images and map |

Data Problems:
- Not Query-able
- Limited to High-Profile Treatments
- Inconsistant Treatment Discription

Geographic Problems:
- No Spatial Data
- Imprecise Project Extents
- No Treatment Extents



This system made it very challenging and labor intensive to answer questions about the program as a whole or even specific projects. Finding all of the projects that contained a specific treatment (for example) became an exercise in archeology and institutional memory. Programmatic analysis was nearly impossible.

NYC DOT decided to create a new system to track the SIPs that utilized a distributed data entry process.

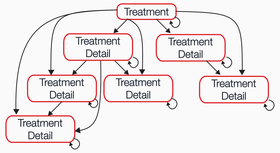
## Database Model

### Database Design:

SIP project tracking system requirements:
- Record all project and treatment details
- Account for different project types
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate the adoption of new treatments and modifications to existing treatment options
- Data entry must be possible by the Project Manager and not require specialized skills or software
- Data must be query-able and accessible to analysts

Closure database model:
- Hierarchical tree structure representation of the data
- Stores relationships between each node to each of its descendants (including itself)
- Data is stored in two tables (data and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model



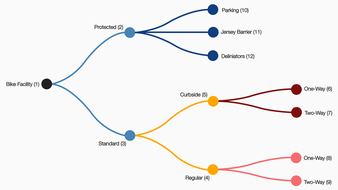

### Implementation:

Projects are a collection of treatments:
- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study and simple standardized
- Do not require Closure model

Treatments are a standardized collection of decisions:
- Each potential treatment must be abstracted into a series of decisions.
- The decision trees that represent the various treatments do not need tohave a uniform breadth or depth.
- Treatment trees define what is possible to record for any treatment
- Treatment trees may be changed to reflect new choices

Treatment Hierarchy Structures

Treatment decision trees are converted into the closure tables
- Each possible decision is represented in the Questions table.
- The relationships between the every decision in the treatment trees are added to the Question Path table.
- Each node and each questions path is unique.
- A 3rd attribute table is added to allow for freeform data fields (such as square feet), these fields must always be the last node in the tree.
- Treatment trees define the possible data can be entered for a project.




| Questions | |
|---|---|
| Question ID | Question |
| 1 | Bike Facility |
| 2 | Protected |
| 3 | Standard |
| 4 | Regular |
| 5 | Curbside |
| 6 | One-Way |
| 7 | Two-Way |
| 8 | One-Way |
| 9 | Two-Way |
| 10 | Parking |
| 11 | Jersy Barrier |
| 12 | Delinlator |

| Question Path | | |
|---|---|---|
| Ancestor | Descendant | Depth |
| 1 | 1 | 0 |
| 1 | 2 | 1 |
| 1 | 3 | 1 |
| 1 | 4 | 2 |
| 1 | 5 | 2 |
| 1 | 6 | 3 |
| 1 | 7 | 3 |
| 1 | 8 | 3 |
| 1 | 9 | 3 |
| 1 | 10 | 2 |
| 1 | 11 | 2 |
| 1 | 12 | 2 |
| 2 | 2 | 0 |
| 2 | 10 | 1 |
| 2 | 11 | 1 |
| 2 | 12 | 1 |
| 3 | 3 | 0 |
| 3 | 4 | 1 |
| 3 | 5 | 1 |
| 3 | 6 | 2 |
| 3 | 7 | 2 |
| 3 | 8 | 2 |
| 3 | 9 | 2 |
| 4 | 4 | 0 |
| 4 | 8 | 1 |
| 4 | 9 | 1 |
| 8 | 8 | 0 |
| 9 | 9 | 0 |

### Recording Treatments:

Project data is recorded in two tables: project attribute table and project treatments table.
- The last decision point that the user passed though is what is recorded in the project treatments table.
- All treatments, regardless of complexity are stored in a single row in the project treatments table.
- The data stored in the project treatments table requires joining to the decision and decision path tables to be understood.

USER INPUT →

| Project | Treatment |
|---|---|
| 1 | 7 |
| 1 | 10 |
| 2 | 5 |

→ Two-Way + Curbside + Standard + Bike Facility
→ Parking + Protected + Bike Facility
→ Curbside + Standard + Bike Facility

## Data Exploration

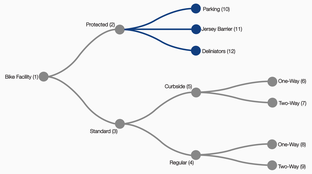
### Exploration in the database:

This data structure is simple to work with in the database. It provides a lot of flexibility for querying at different levels of granularity or recreating the full treatment details from the last decision node.

The data structure allows for ad hock groupings within the hierarchy. Queries down the tree return all of the choices below a given node. This is used to find all treatments that include at least a particular decision. For example: Finding all protected bike paths, regardless of the specific details of the subsequent decisions.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT descendant d, depth
    FROM sip_questions_path
    -- All treatments with a protected bike facility
    WHERE ancestor = 2
) AS p ON q.qid = p.d
ORDER BY p.depth;
```

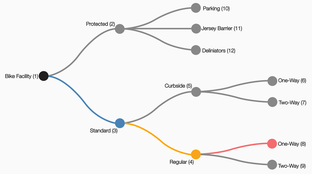
| qid integer | question text |
|---|---|
| 1 | 2 Protected |
| 2 | 10 Parking |
| 3 | 11 Jersey Barrier |
| 4 | 12 Delinators |




The treatment attributes that are recorded in association with a project are limited to the final decision node the user passed through. The ability to interpret the treatment relies on being able to recreate the full set of choices the user made. Recreating the full decision path of a treatment detail necessitates querying up the tree. For example: One-way (QID 8) is only a meaningful treatment description once the full path back to the root is revilled as a One-way standard non-curbside bike facility.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT ancestor a, depth
    FROM sip_questions_path
    -- All treatments with a one-way
       standard bike facility
    WHERE descendant = 8
  ) AS p ON q.qid = p.a
ORDER BY p.depth desc;
```

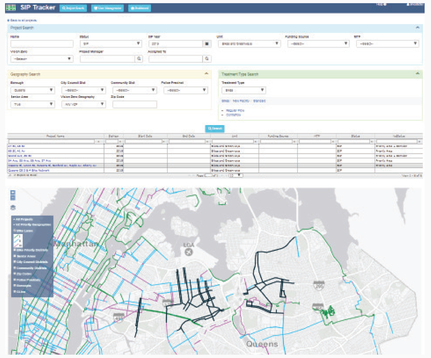
| qid integer | question text |
|---|---|
| 1 | 1 Bike Facility |
| 2 | 3 Standard |
| 3 | 4 Regular |
| 4 | 8 One-way |




### Exploration in the application:

The application provides planners, program managers, analysts and outreach/communications staff with the ability to quickly and easily explore the data.

Project managers need to know what other projects are being developed and/or implemented near their project. To facilitate this, the application provides a number of reference layers that can be toggled in the map provide an overlay of:
- All projects (complete and under development)
- Political and administrative districts
- Priority geographies
- Existing bicycle network



For a more targeted search, the application provides a query tool that allows for searching for projects using three categories of criteria
- Project
- Geography
- Treatment

The treatment search utilizes the treatment decision trees to allow user defined specificity.

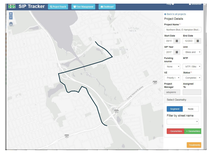
The results of the search in both the table and the map link to the projects' full detail page.

## Data Entry

The primary purpose of the application is to facilitate the collection and storage of the project and treatment attributes. Data entry is intended to be done by the project manager.

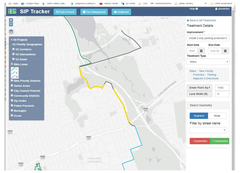
The data entry process is straightforward and consists of two levels, project and treatment.

Project data entry process:
1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography



Treatment data entry process:

1. Enter treatment level attributes
2. Enter treatment type details via decision tree
3. Select treatment geography
4. Repeat for all other treatments associated with the project

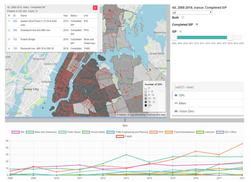


## Applications

### Planning

The application includes a dashboard to help unit directors to explore their group's work. NYC DOT now uses the dashboard as a tool to plan future work programs at a high level

The dashboard provides a powerful tool to ensure that communities were not over nor underrepresented in the SIP program's work portfolio.



### Analysis

NYC DOT has begun to analyze the equitability of the SIP program as a whole to ensure all New Yorkers were being served. This analysis explores the geographic distribution of projects throughout the city across a number of different demographic metrics.

NYC DOT has begun to analyze the full SIP work program's efficacy from a safety perspective. Every SIP project is, analyzed individually to ensure they are improving safety, but a holistic analysis has not historically been practical.

## Credits

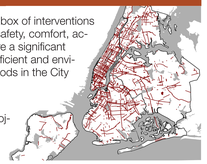
Project Team:
Jason Fitzsimmons
Arthur Getman
Zhi Keng He
Seth Hostetter
Chandrima Pal
Sidney Salnave
Tanuja Ramesh Vanjarapu

Technologies:
PostgreSQL / PostGIS
C# .Net
Geoserver v2.8
OpenLayers 3
ESRI map services

# STREET IMPROVEMENT PROJECT TRACKING APPLICATION: USING A HIERARCHICAL DATABASE MODEL

## Street Improvement Projects

Street Improvement Projects (SIP) utilize a broad toolbox of interventions to reengineer city streets with the goal of improving safety, comfort, accessibility, and mobility for all roadway users. They are a significant part of NYC DOT's mission to provide for the safe, efficient and environmentally responsible movement of people and goods in the City of New York.



NYC DOT implements more than 100 distinct SIP projects per year, carried out by 8 different units.

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Develop-

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementa-

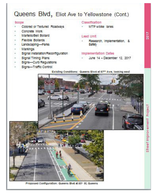
## Historic Tracking

Historically, SIP projects were tracked for the program in a master tracking sheet and at the project level in a summary document.

| Excel tracking sheet | PDF project description |
|---|---|
| Maintained centrally by program manager | Maintained by project manager |
| Project name | Project background |
| One-line project description | High level description |
| Implementation dates | High level treatment descriptions |
| Project manager name and unit | Project images and map |

Data Problems:
- Not Query-able
- Limited to High-Profile Treatments
- Inconsistant Treatment Discription

Geographic Problems:
- No Spatial Data
- Imprecise Project Extents
- No Treatment Extents



This system made it very challenging and labor intensive to answer questions about the program as a whole or even specific projects. Finding all of the projects that contained a specific treatment (for example) became an exercise in archeology and institutional memory. Programmatic analysis was nearly impossible.

NYC DOT decided to create a new system to track the SIPs that utilized a distributed data entry process.

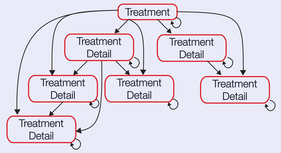
## Database Model

### Database Design:

SIP project tracking system requirements:
- Record all project and treatment details
- Account for different project types
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate the adoption of new treatments and modifications to existing treatment options
- Data entry must be possible by the Project Manager and not require specialized skills or software
- Data must be query-able and accessible to analysts

Closure database model:
- Hierarchical tree structure representation of the data
- Stores relationships between each node to each of its descendants (including itself)
- Data is stored in two tables (data and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model

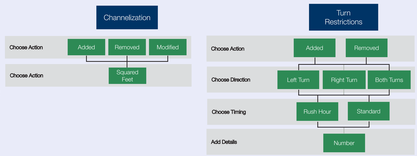



### Implementation:

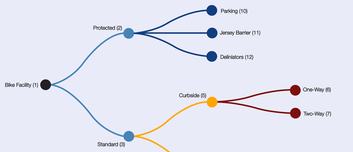
Projects are a collection of treatments:
- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study and simple standardized
- Do not require Closure model

Treatments are a standardized collection of decisions:
- Each potential treatment must be abstracted into a series of decisions.
- The decision trees that represent the various treatments do not need tohave a uniform breadth or depth.
- Treatment trees define what is possible to record for any treatment
- Treatment trees may be changed to reflect new choices

Treatment Hierarchy Structures




Treatment decision trees are converted into the closure tables
- Each possible decision is represented in the Questions table.
- The relationships between the every decision in the treatment trees are added to the Question Path table.
- Each node and each questions path is unique.
- A 3rd attribute table is added to allow for freeform data fields (such as square feet), these fields must always be the last node in the tree.
- Treatment trees define the possible data can be entered for a project.




| Questions | |
|---|---|
| Question ID | Question |
| 1 | Bike Facility |
| 2 | Protected |
| 3 | Standard |
| 4 | Regular |
| 5 | Curbside |
| 6 | One-Way |
| 7 | Two-Way |
| 8 | One-Way |
| 9 | Two-Way |
| 10 | Parking |
| 11 | Jersy Barrier |
| 12 | Delinlator |

| Question Path | | |
|---|---|---|
| Ancestor | Descendant | Depth |
| 1 | 1 | 0 |
| 1 | 2 | 1 |
| 1 | 3 | 1 |
| 1 | 4 | 2 |
| 1 | 5 | 2 |
| 1 | 6 | 3 |
| 1 | 7 | 3 |
| 1 | 8 | 3 |
| 1 | 9 | 3 |
| 1 | 10 | 2 |
| 1 | 11 | 2 |
| 1 | 12 | 2 |
| 2 | 2 | 0 |
| 2 | 10 | 1 |
| 2 | 11 | 1 |
| 2 | 12 | 1 |
| 3 | 3 | 0 |
| 3 | 4 | 1 |
| 3 | 5 | 1 |
| 3 | 6 | 2 |
| 3 | 7 | 2 |
| 3 | 8 | 2 |
| 3 | 9 | 2 |
| 4 | 4 | 0 |
| 4 | 8 | 1 |
| 4 | 9 | 1 |
| 8 | 8 | 0 |
| 9 | 9 | 0 |

### Recording Treatments:

Project data is recorded in two tables: project attribute table and project treatments table.
- The last decision point that the user passed though is what is recorded in the project treatments table.
- All treatments, regardless of complexity are stored in a single row in the project treatments table.
- The data stored in the project treatments table requires joining to the decision and decision path tables to be understood.

## Data Exploration

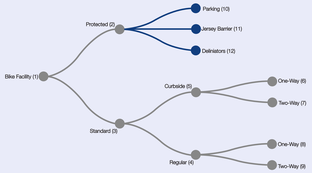
### Exploration in the database:

This data structure is simple to work with in the database. It provides a lot of flexibility for querying at different levels of granularity or recreating the full treatment details from the last decision node.

The data structure allows for ad hock groupings within the hierarchy. Queries down the tree return all of the choices below a given node. This is used to find all treatments that include at least a particular decision. For example: Finding all protected bike paths, regardless of the specific details of the subsequent decisions.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT descendant d, depth
    FROM sip_questions_path
    -- All treatments with a protected bike facility
    WHERE ancestor = 2
) AS p ON q.qid = p.d
ORDER BY p.depth;
```

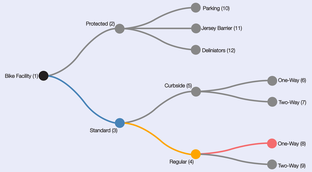
| qid integer | question text |
|---|---|
| 1 | 2 Protected |
| 2 | 10 Parking |
| 3 | 11 Jersey Barrier |
| 4 | 12 Delinators |




The treatment attributes that are recorded in association with a project are limited to the final decision node the user passed through. The ability to interpret the treatment relies on being able to recreate the full set of choices the user made. Recreating the full decision path of a treatment detail necessitates querying up the tree. For example: One-way (QID 8) is only a meaningful treatment description once the full path back to the root is revileed as a One-way standard non-curbside bike facility.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT ancestor a, depth
    FROM sip_questions_path
    -- All treatments with a one-way
       standard bike facility
    WHERE descendant = 8
  ) AS p ON q.qid = p.a
ORDER BY p.depth desc;
```

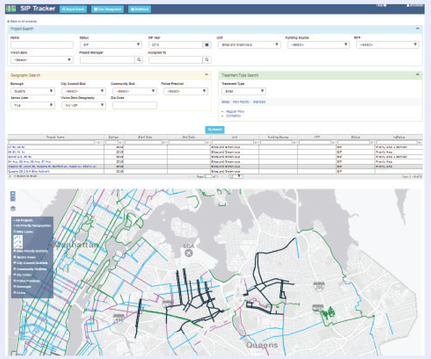
| qid integer | question text |
|---|---|
| 1 | 1 Bike Facility |
| 2 | 3 Standard |
| 3 | 4 Regular |
| 4 | 8 One-way |




### Exploration in the application:

The application provides planners, program managers, analysts and outreach/communications staff with the ability to quickly and easily explore the data.

Project managers need to know what other projects are being developed and/or implemented near their project. To facilitate this, the application provides a number of reference layers that can be toggled in the map provide an overlay of:
- All projects (complete and under development)
- Political and administrative districts
- Priority geographies
- Existing bicycle network




For a more targeted search, the application provides a query tool that allows for searching for projects using three categories of criteria
- Project
- Geography
- Treatment

The treatment search utilizes the treatment decision trees to allow user defined specificity.

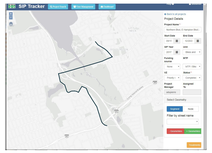
The results of the search in both the table and the map link to the projects' full detail page.

## Data Entry

The primary purpose of the application is to facilitate the collection and storage of the project and treatment attributes. Data entry is intended to be done by the project manager.

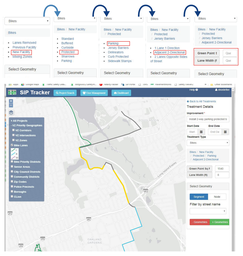
The data entry process is straightforward and consists of two levels, project and treatment.

Project data entry process:
1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography



Treatment data entry process:

1. Enter treatment level attributes
2. Enter treatment type details via decision tree
3. Select treatment geography
4. Repeat for all other treatments associated with the project

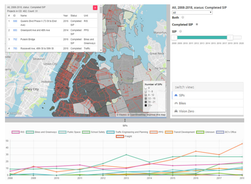



## Applications

### Planning

The application includes a dashboard to help unit directors to explore their group's work. NYC DOT now uses the dashboard as a tool to plan future work programs at a high level

The dashboard provides a powerful tool to ensure that communities were not over nor underrepresented in the SIP program's work portfolio.



### Analysis

NYC DOT has begun to analyze the equitability of the SIP program as a whole to ensure all New Yorkers were being served. This analysis explores the geographic distribution of projects throughout the city across a number of different demographic metrics.

NYC DOT has begun to analyze the full SIP work program's efficacy from a safety perspective. Every SIP project is, analyzed individually to ensure they are improving safety, but a holistic analysis has not historically been practical.

## Credits

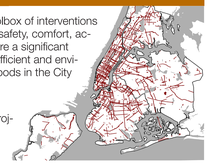

Project Team:
Jason Fitzsimmons
Arthur Getman
Zhi Keng He
Seth Hostetter
Chandrima Pal
Sidney Salnave
Tanuja Ramesh Vanjarapu

Technologies:
PostgreSQL / PostGIS
C# .Net
Geoserver v2.8
OpenLayers 3
ESRI map services

# STREET IMPROVEMENT PROJECT TRACKING APPLICATION: USING A HIERARCHICAL DATABASE MODEL

## Street Improvement Projects

Street Improvement Projects (SIP) utilize a broad toolbox of interventions to reengineer city streets with the goal of improving safety, comfort, accessibility, and mobility for all roadway users. They are a significant part of NYC DOT's mission to provide for the safe, efficient and environmentally responsible movement of people and goods in the City of New York.



NYC DOT implements more than 100 distinct SIP projects per year, carried out by 8 different units.

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Develop-

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementa-

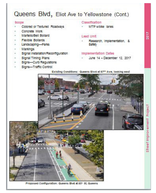
## Historic Tracking

Historically, SIP projects were tracked for the program in a master tracking sheet and at the project level in a summary document.

| Excel tracking sheet | PDF project description |
|---|---|
| Maintained centrally by program manager | Maintained by project manager |
| Project name | Project background |
| One-line project description | High level description |
| Implementation dates | High level treatment descriptions |
| Project manager name and unit | Project images and map |

Data Problems:
- Not Query-able
- Limited to High-Profile Treatments
- Inconsistant Treatment Discription

Geographic Problems:
- No Spatial Data
- Imprecise Project Extents
- No Treatment Extents



This system made it very challenging and labor intensive to answer questions about the program as a whole or even specific projects. Finding all of the projects that contained a specific treatment (for example) became an exercise in archeology and institutional memory. Programmatic analysis was nearly impossible.

NYC DOT decided to create a new system to track the SIPs that utilized a distributed data entry process.

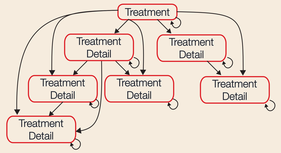
## Database Model

### Database Design:

SIP project tracking system requirements:
- Record all project and treatment details
- Account for different project types
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate the adoption of new treatments and modifications to existing treatment options
- Data entry must be possible by the Project Manager and not require specialized skills or software
- Data must be query-able and accessible to analysts

Closure database model:
- Hierarchical tree structure representation of the data
- Stores relationships between each node to each of its descendants (including itself)
- Data is stored in two tables (data and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model



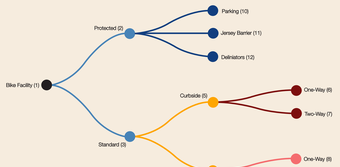

### Implementation:

Projects are a collection of treatments:
- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study and simple standardized
- Do not require Closure model

Treatments are a standardized collection of decisions:
- Each potential treatment must be abstracted into a series of decisions.
- The decision trees that represent the various treatments do not need tohave a uniform breadth or depth.
- Treatment trees define what is possible to record for any treatment
- Treatment trees may be changed to reflect new choices

Treatment Hierarchy Structures

Treatment decision trees are converted into the closure tables
- Each possible decision is represented in the Questions table.
- The relationships between the every decision in the treatment trees are added to the Question Path table.
- Each node and each questions path is unique.
- A 3rd attribute table is added to allow for freeform data fields (such as square feet), these fields must always be the last node in the tree.
- Treatment trees define the possible data can be entered for a project.




| Questions | |
|---|---|
| Question ID | Question |
| 1 | Bike Facility |
| 2 | Protected |
| 3 | Standard |
| 4 | Regular |
| 5 | Curbside |
| 6 | One-Way |
| 7 | Two-Way |
| 8 | One-Way |
| 9 | Two-Way |
| 10 | Parking |
| 11 | Jersy Barrier |
| 12 | Delinlator |

| Question Path | | |
|---|---|---|
| Ancestor | Descendant | Depth |
| 1 | 1 | 0 |
| 1 | 2 | 1 |
| 1 | 3 | 1 |
| 1 | 4 | 2 |
| 1 | 5 | 2 |
| 1 | 6 | 3 |
| 1 | 7 | 3 |
| 1 | 8 | 3 |
| 1 | 9 | 3 |
| 1 | 10 | 2 |
| 1 | 11 | 2 |
| 1 | 12 | 2 |
| 2 | 2 | 0 |
| 2 | 10 | 1 |
| 2 | 11 | 1 |
| 2 | 12 | 1 |
| 3 | 3 | 0 |
| 3 | 4 | 1 |
| 3 | 5 | 1 |
| 3 | 6 | 2 |
| 3 | 7 | 2 |
| 3 | 8 | 2 |
| 3 | 9 | 2 |
| 4 | 4 | 0 |
| 4 | 8 | 1 |
| 4 | 9 | 1 |
| 8 | 8 | 0 |
| 9 | 9 | 0 |

### Recording Treatments:

Project data is recorded in two tables: project attribute table and project treatments table.
- The last decision point that the user passed though is what is recorded in the project treatments table.
- All treatments, regardless of complexity are stored in a single row in the project treatments table.
- The data stored in the project treatments table requires joining to the decision and decision path tables to be understood.

USER INPUT →

| Project | Treatment |
|---|---|
| 1 | 7 |
| 1 | 10 |
| 2 | 5 |

→ Two-Way + Curbside + Standard + Bike Facility
→ Parking + Protected + Bike Facility
→ Curbside + Standard + Bike Facility

## Data Exploration

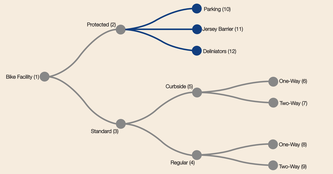
### Exploration in the database:

This data structure is simple to work with in the database. It provides a lot of flexibility for querying at different levels of granularity or recreating the full treatment details from the last decision node.

The data structure allows for ad hock groupings within the hierarchy. Queries down the tree return all of the choices below a given node. This is used to find all treatments that include at least a particular decision. For example: Finding all protected bike paths, regardless of the specific details of the subsequent decisions.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT descendant d, depth
    FROM sip_questions_path
    -- All treatments with a protected bike facility
    WHERE ancestor = 2
) AS p ON q.qid = p.d
ORDER BY p.depth;
```

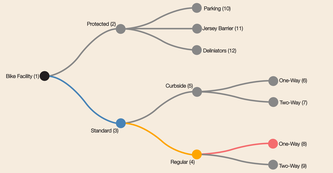
| qid integer | question text |
|---|---|
| 1 | 2 Protected |
| 2 | 10 Parking |
| 3 | 11 Jersey Barrier |
| 4 | 12 Delinators |




The treatment attributes that are recorded in association with a project are limited to the final decision node the user passed through. The ability to interpret the treatment relies on being able to recreate the full set of choices the user made. Recreating the full decision path of a treatment detail necessitates querying up the tree. For example: One-way (QID 8) is only a meaningful treatment description once the full path back to the root is revileed as a One-way standard non-curbside bike facility.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT ancestor a, depth
    FROM sip_questions_path
    -- All treatments with a one-way
       standard bike facility
    WHERE descendant = 8
  ) AS p ON q.qid = p.a
ORDER BY p.depth desc;
```

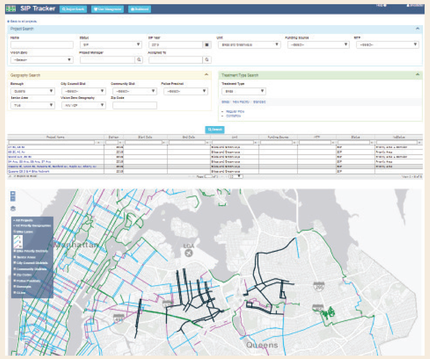
| qid integer | question text |
|---|---|
| 1 | 1 Bike Facility |
| 2 | 3 Standard |
| 3 | 4 Regular |
| 4 | 8 One-way |




### Exploration in the application:

The application provides planners, program managers, analysts and outreach/communications staff with the ability to quickly and easily explore the data.

Project managers need to know what other projects are being developed and/or implemented near their project. To facilitate this, the application provides a number of reference layers that can be toggled in the map provide an overlay of:
- All projects (complete and under development)
- Political and administrative districts
- Priority geographies
- Existing bicycle network



For a more targeted search, the application provides a query tool that allows for searching for projects using three categories of criteria
- Project
- Geography
- Treatment

The treatment search utilizes the treatment decision trees to allow user defined specificity.

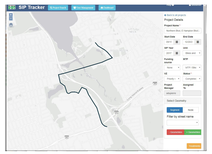
The results of the search in both the table and the map link to the projects' full detail page.

## Data Entry

The primary purpose of the application is to facilitate the collection and storage of the project and treatment attributes. Data entry is intended to be done by the project manager.

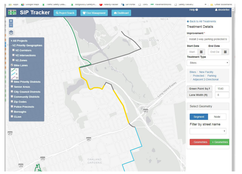
The data entry process is straightforward and consists of two levels, project and treatment.

Project data entry process:
1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography



Treatment data entry process:

1. Enter treatment level attributes
2. Enter treatment type details via decision tree
3. Select treatment geography
4. Repeat for all other treatments associated with the project

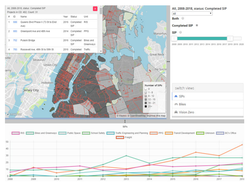


## Applications

### Planning

The application includes a dashboard to help unit directors to explore their group's work. NYC DOT now uses the dashboard as a tool to plan future work programs at a high level

The dashboard provides a powerful tool to ensure that communities were not over nor underrepresented in the SIP program's work portfolio.



### Analysis

NYC DOT has begun to analyze the equitability of the SIP program as a whole to ensure all New Yorkers were being served. This analysis explores the geographic distribution of projects throughout the city across a number of different demographic metrics.

NYC DOT has begun to analyze the full SIP work program's efficacy from a safety perspective. Every SIP project is, analyzed individually to ensure they are improving safety, but a holistic analysis has not historically been practical.

## Credits

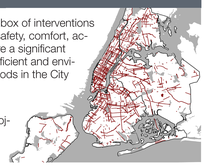

Project Team:
Jason Fitzsimmons
Arthur Getman
Zhi Keng He
Seth Hostetter
Chandrima Pal
Sidney Salnave
Tanuja Ramesh Vanjarapu


Technologies:
PostgreSQL / PostGIS
C# .Net
Geoserver v2.8
OpenLayers 3
ESRI map services

# STREET IMPROVEMENT PROJECT TRACKING APPLICATION: USING A HIERARCHICAL DATABASE MODEL

## Street Improvement Projects

Street Improvement Projects (SIP) utilize a broad toolbox of interventions to reengineer city streets with the goal of improving safety, comfort, accessibility, and mobility for all roadway users. They are a significant part of NYC DOT's mission to provide for the safe, efficient and environmentally responsible movement of people and goods in the City of New York.



NYC DOT implements more than 100 distinct SIP projects per year, carried out by 8 different units.

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Develop-

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementa-

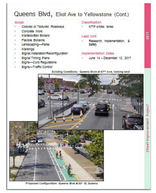
## Historic Tracking

Historically, SIP projects were tracked for the program in a master tracking sheet and at the project level in a summary document.

| Excel tracking sheet | PDF project description |
|---|---|
| Maintained centrally by program manager | Maintained by project manager |
| Project name | Project background |
| One-line project description | High level description |
| Implementation dates | High level treatment descriptions |
| Project manager name and unit | Project images and map |

Data Problems:
- Not Query-able
- Limited to High-Profile Treatments
- Inconsistant Treatment Discription

Geographic Problems:
- No Spatial Data
- Imprecise Project Extents
- No Treatment Extents



This system made it very challenging and labor intensive to answer questions about the program as a whole or even specific projects. Finding all of the projects that contained a specific treatment (for example) became an exercise in archeology and institutional memory. Programmatic analysis was nearly impossible.

NYC DOT decided to create a new system to track the SIPs that utilized a distributed data entry process.

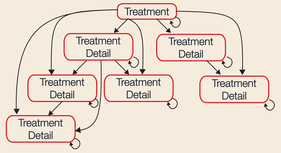
## Database Model

### Database Design:

SIP project tracking system requirements:
- Record all project and treatment details
- Account for different project types
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate the adoption of new treatments and modifications to existing treatment options
- Data entry must be possible by the Project Manager and not require specialized skills or software
- Data must be query-able and accessible to analysts

Closure database model:
- Hierarchical tree structure representation of the data
- Stores relationships between each node to each of its descendants (including itself)
- Data is stored in two tables (data and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model

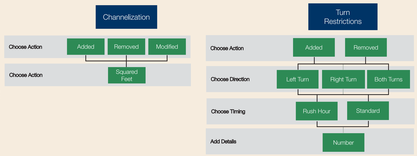



### Implementation:

Projects are a collection of treatments:
- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study and simple standardized
- Do not require Closure model

Treatments are a standardized collection of decisions:
- Each potential treatment must be abstracted into a series of decisions.
- The decision trees that represent the various treatments do not need tohave a uniform breadth or depth.
- Treatment trees define what is possible to record for any treatment
- Treatment trees may be changed to reflect new choices

Treatment Hierarchy Structures




Treatment decision trees are converted into the closure tables
- Each possible decision is represented in the Questions table.
- The relationships between the every decision in the treatment trees are added to the Question Path table.
- Each node and each questions path is unique.
- A 3rd attribute table is added to allow for freeform data fields (such as square feet), these fields must always be the last node in the tree.
- Treatment trees define the possible data can be entered for a project.

| Questions | |
|---|---|
| Question ID | Question |
| 1 | Bike Facility |
| 2 | Protected |
| 3 | Standard |
| 4 | Regular |
| 5 | Curbside |
| 6 | One-Way |
| 7 | Two-Way |
| 8 | One-Way |
| 9 | Two-Way |
| 10 | Parking |
| 11 | Jersy Barrier |
| 12 | Delinlator |

| Question Path | | |
|---|---|---|
| Ancestor | Descendant | Depth |
| 1 | 1 | 0 |
| 1 | 2 | 1 |
| 1 | 3 | 1 |
| 1 | 4 | 2 |
| 1 | 5 | 2 |
| 1 | 6 | 3 |
| 1 | 7 | 3 |
| 1 | 8 | 3 |
| 1 | 9 | 3 |
| 1 | 10 | 2 |
| 1 | 11 | 2 |
| 1 | 12 | 2 |
| 2 | 2 | 0 |
| 2 | 10 | 1 |
| 2 | 11 | 1 |
| 2 | 12 | 1 |
| 3 | 3 | 0 |
| 3 | 4 | 1 |
| 3 | 5 | 1 |
| 3 | 6 | 2 |
| 3 | 7 | 2 |
| 3 | 8 | 2 |
| 3 | 9 | 2 |
| 4 | 4 | 0 |
| 4 | 8 | 1 |
| 4 | 9 | 1 |
| 8 | 8 | 0 |
| 9 | 9 | 0 |

### Recording Treatments:

Project data is recorded in two tables: project attribute table and project treatments table.
- The last decision point that the user passed though is what is recorded in the project treatments table.
- All treatments, regardless of complexity are stored in a single row in the project treatments table.
- The data stored in the project treatments table requires joining to the decision and decision path tables to be understood.

USER INPUT →

| Project | Treatment |
|---|---|
| 1 | 7 |
| 1 | 10 |
| 2 | 5 |

→ Two-Way + Curbside + Standard + Bike Facility
→ Parking + Protected + Bike Facility
→ Curbside + Standard + Bike Facility

## Data Exploration

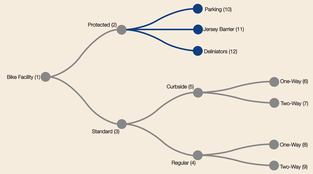
### Exploration in the database:

This data structure is simple to work with in the database. It provides a lot of flexibility for querying at different levels of granularity or recreating the full treatment details from the last decision node.

The data structure allows for ad hock groupings within the hierarchy. Queries down the tree return all of the choices below a given node. This is used to find all treatments that include at least a particular decision. For example: Finding all protected bike paths, regardless of the specific details of the subsequent decisions.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT descendant d, depth
    FROM sip_questions_path
    -- All treatments with a protected bike facility
    WHERE ancestor = 2
) AS p ON q.qid = p.d
ORDER BY p.depth;
```

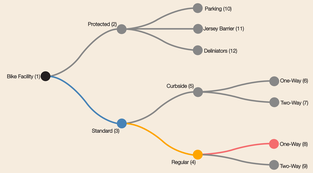
| qid integer | question text |
|---|---|
| 2 | Protected |
| 10 | Parking |
| 11 | Jersey Barrier |
| 12 | Delinators |




The treatment attributes that are recorded in association with a project are limited to the final decision node the user passed through. The ability to interpret the treatment relies on being able to recreate the full set of choices the user made. Recreating the full decision path of a treatment detail necessitates querying up the tree. For example: One-way (QID 8) is only a meaningful treatment description once the full path back to the root is revilled as a One-way standard non-curbside bike facility.

```
SELECT q.*
FROM sip_questions q
JOIN (
    SELECT ancestor a, depth
    FROM sip_questions_path
    -- All treatments with a one-way
       standard bike facility
    WHERE descendant = 8
  ) AS p ON q.qid = p.a
ORDER BY p.depth desc;
```

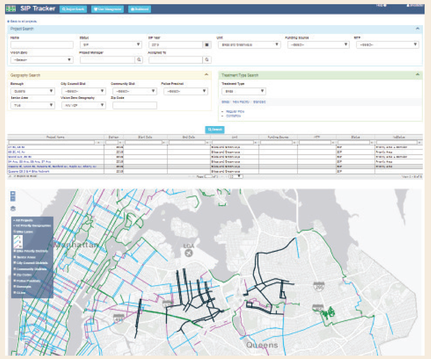
| qid integer | question text |
|---|---|
| 1 | Bike Facility |
| 3 | Standard |
| 4 | Regular |
| 8 | One-way |




### Exploration in the application:

The application provides planners, program managers, analysts and outreach/communications staff with the ability to quickly and easily explore the data.

Project managers need to know what other projects are being developed and/or implemented near their project. To facilitate this, the application provides a number of reference layers that can be toggled in the map provide an overlay of:
- All projects (complete and under development)
- Political and administrative districts
- Priority geographies
- Existing bicycle network



For a more targeted search, the application provides a query tool that allows for searching for projects using three categories of criteria
- Project
- Geography
- Treatment

The treatment search utilizes the treatment decision trees to allow user defined specificity.

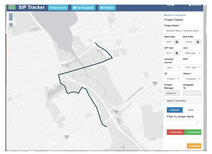
The results of the search in both the table and the map link to the projects' full detail page.

## Data Entry

The primary purpose of the application is to facilitate the collection and storage of the project and treatment attributes. Data entry is intended to be done by the project manager.

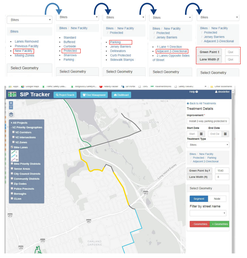
The data entry process is straightforward and consists of two levels, project and treatment.

Project data entry process:
1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography



Treatment data entry process:

1. Enter treatment level attributes
2. Enter treatment type details via decision tree
3. Select treatment geography
4. Repeat for all other treatments associated with the project

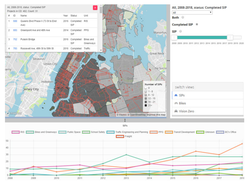


## Applications

### Planning

The application includes a dashboard to help unit directors to explore their group's work. NYC DOT now uses the dashboard as a tool to plan future work programs at a high level

The dashboard provides a powerful tool to ensure that communities were not over nor underrepresented in the SIP program's work portfolio.



### Analysis

NYC DOT has begun to analyze the equitability of the SIP program as a whole to ensure all New Yorkers were being served. This analysis explores the geographic distribution of projects throughout the city across a number of different demographic metrics.

NYC DOT has begun to analyze the full SIP work program's efficacy from a safety perspective. Every SIP project is, analyzed individually to ensure they are improving safety, but a holistic analysis has not historically been practical.

## Credits

Project Team:

Jason Fitzsimmons
Arthur Getman
Zhi Keng He
Seth Hostetter
Chandrima Pal
Sidney Salnave
Tanuja Ramesh Vanjarapu

Technologies:
PostgreSQL / PostGIS
C# .Net
Geoserver v2.8
OpenLayers 3
ESRI map services