HIERARCHICAL DATABASE MODEL FOR STREET IMPROVEMENT PROJECT TRACKING

Street Improvement Projects

Street Improvement Projects (SIPs):

- SIPs are a Significant part of NYC DOT's mission to provide safe, efficient and environmentally responsible movement of people and goods in the City of New York
- Reengineer city streets to improve **safety, comfort, accessibility, and mobility**
- NYC DOT implements more than 100 SIPs per year designed by 8 different units

Traffic Calming, Transit Prioritization

Traffic Calming, Bike Infrastructure Development

Traffic Calming, Bike Infrastructure, & Plaza

Street Closure, Pedestrian Plaza Implementation

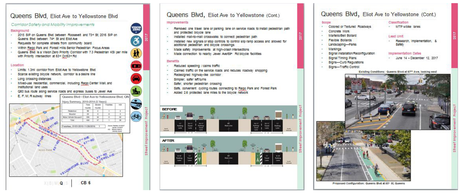
Historic Tracking

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
Inconsistent Treatment Description

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 and project details.

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

Database Model

Database Design:

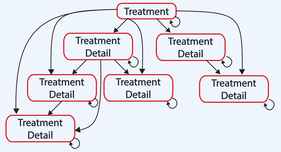
System requirements:

- Record all project and treatment details
- Allow for diversity of treatment details (different treatments have differing levels of details)
- Standardize treatment attributes and related geographic information
- Facilitate modifications to treatment options
- Data entry must be possible by the Project Manager (no 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 (treatment detail and relationships)
- Allows for detail and flexibility to store treatments with differing levels of detail in one schema

Closure Hierarchy Model

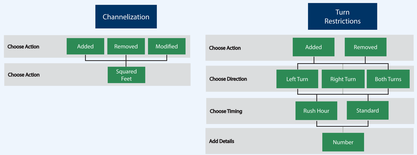


Implementation:

Projects are a collection of treatments:

- Project and treatment data are stored in separate, but related tables
- Project attributes describe the study (e.g. unit, implementation dates)
- Do not require Closure model

Treatment Hierarchy Structures

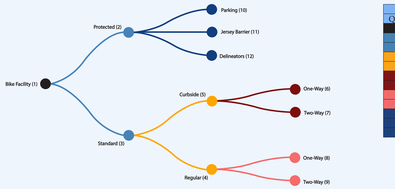


Treatments are a standardized collection of decisions:

- Each treatment must be abstracted into a series of decisions
- The decision trees do not need to have 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 decision trees are converted into the closure tables

- Each possible decision is represented in the Questions table
- The relationships between 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



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	Jersey Barrier
12	Delineators

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

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

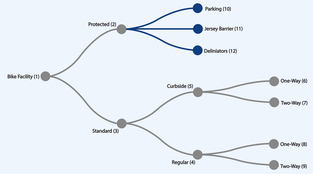
Exploration in the database:

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 for finding all **broad-stroke treatments** (e.g. All protected bike paths, regardless of details).

```
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;
```

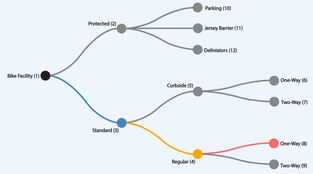
qid integer	question text
1	Protected
2	Parking
3	Jersey Barrier
4	Delineators



The treatment attributes are limited to the final decision node the user passed through. Recreating the full decision path of a treatment detail necessitates querying **up the tree**. This is used for finding all **specifics of a treatment** (e.g. tracing breadcrumbs to see full treatment).

```
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

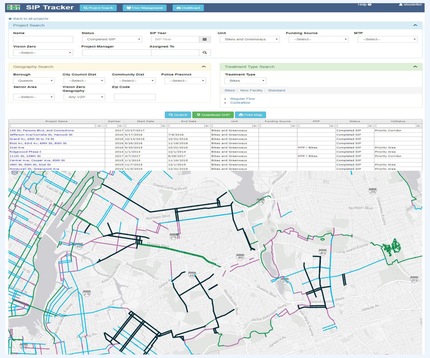


Exploration in the application:

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

The application provides a number of **reference layers** to give context to a manager's projects:

- 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.

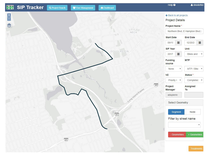
The results of the search in both the table and the map link to the projects' full detail pages.

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.

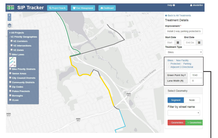
Step 1: Project data entry process:

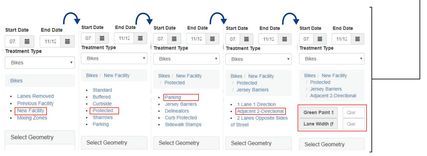
1. Enter project level attributes
2. Option to delegate the project to another user (share ownership)
3. Select project geography



Step 2: 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



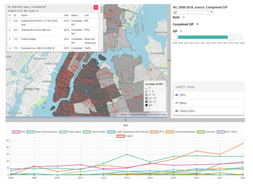


Applications

Planning

The application includes a dashboard to help unit directors explore their group's work. NYC DOT now uses the **dashboard as a tool to plan future work** programs at a high level (e.g. identifying gaps or over-representations in program's work portfolio).

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



Analysis

NYC DOT has begun to analyze the **equity** of the SIP program as a whole to ensure all New Yorkers are 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 **safety impacts** of the SIP program as a whole. Every SIP project is analyzed individually to ensure they are improving safety, but a holistic analysis has historically not been practical.

Credits

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

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