TECHNICAL DOCUMENTATION

Opportunities for Integrated Water Management

Jamaica Bay, New York City

Interactive Spatial Atlas — Web Map Documentation

Prepared by
Chanaporn (New) Tohsuwanwanich
Master of Urban Planning · NYU Wagner School of Public Service
Commissioned by Town & Gown NYC · NYC Department of Design and Construction (DDC)
Version 1.0 · May 2026

https://ctohsuwanwanich.github.io/ddc_webmap/

Executive Summary

New York City's Jamaica Bay watershed is the site of dozens of concurrent capital infrastructure projects managed by four distinct municipal agencies: the Department of Environmental Protection (DEP), the Department of Transportation (DOT), the Department of Design and Construction (DDC), and the Department of Parks and Recreation (DPR). Although each agency operates according to its own planning mandate and procurement cycle, many of these projects overlap spatially and temporally, occurring on the same streets, blocks, or parcels without formal coordination.

This project develops an integrated, web-based spatial atlas that compiles, cleans, and visualizes these capital investments across the Jamaica Bay Long-Term Control Plan study area. By drawing together datasets from multiple agency sources into a unified geospatial framework, the atlas surfaces patterns of infrastructure co-occurrence that are otherwise obscured by the fragmented nature of city agency data systems.

The resulting interactive web map enables planners, researchers, and agency staff to:

- Visualize capital projects across agencies and infrastructure types in a single shared spatial view
- Identify clusters of overlapping investment at the tax-block level through a choropleth concentration layer
- Explore temporal patterns in project delivery using completion-year filtering
- Navigate a curated tour of eleven inter-agency cluster sites where coordination opportunities are strongest

The tool is intended to support planning conversations around coordinated infrastructure delivery, helping agencies identify where joint implementation could reduce construction disruption, lower redundant costs, and generate compounded environmental benefit across water quality, flood resilience, and open space objectives.

1. Project Overview

1.1 Problem Framing

Infrastructure planning in New York City is structurally distributed across agencies with overlapping yet formally separate mandates. Streets and transportation corridors fall under DOT jurisdiction; sewer, stormwater, and water quality systems are managed by DEP; parks, greenways, and shoreline environments by DPR; and large capital construction projects across agencies are often delivered by DDC. While this division of responsibility allows for specialization, it creates conditions for significant inefficiency when multiple agencies are working in close spatial and temporal proximity.

Common consequences of this fragmentation include duplicated construction staging and street disruption, missed opportunities for integrated green infrastructure design, uncoordinated scheduling that extends overall project timelines, and an inability to identify where combined investment could yield compounded environmental returns. Jamaica Bay represents a particularly acute case of these dynamics. The watershed encompasses aging combined sewer infrastructure, flood-prone transportation corridors, ecologically sensitive tidal shoreline environments, and intensive multi-agency capital investment — all concentrated within a relatively compact geographic area spanning Brooklyn and Queens.

1.2 Project Summary

Parameter	Detail
Study Area	Jamaica Bay watershed (Brooklyn and Queens boroughs)
Agencies Covered	DEP, DOT, DDC, DPR
Primary Deliverable	Interactive web map with spatial analysis layer
GIS Platform	ArcGIS Pro (spatial processing and analysis)
Web Platform	Mapbox GL JS (interactive visualization)
Audience	City agency staff, urban planners, and researchers
Version	1.0 — May 2026

1.3 Interactive Map — Overview

The web map is accessible at the URL below and serves as the primary deliverable of this project. It is designed for use in desktop browsers and provides two entry modes: a free-exploration mode for browsing all capital projects, and a guided cluster tour that walks through eleven pre-identified sites of inter-agency overlap.

https://ctohsuwanwanich.github.io/ddc_webmap/

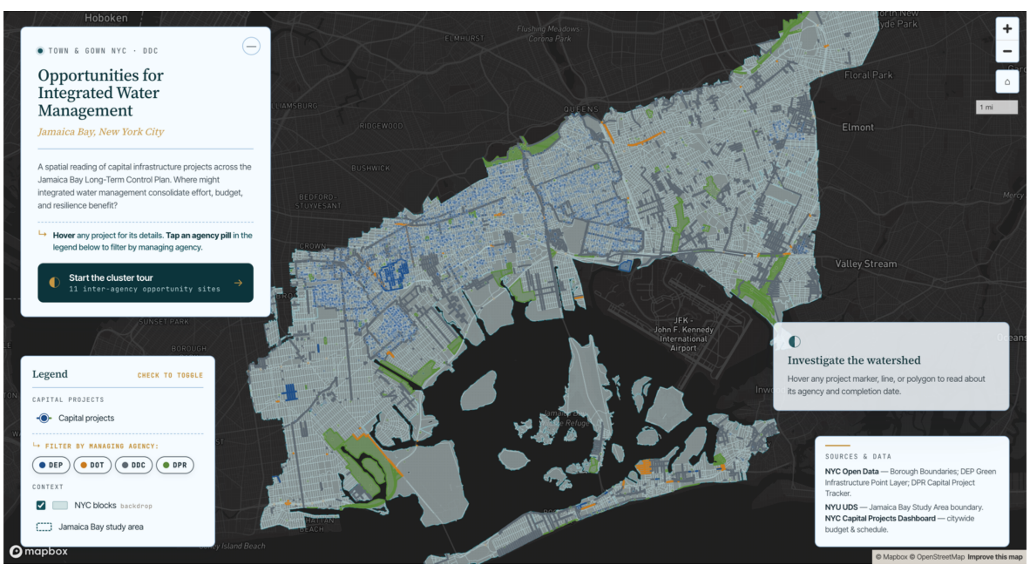

Figure 1 — The main map interface showing project markers colored by agency, the top navigation bar, and the legend panel

2. Data Sources

2.1 Key Datasets

The project integrates seven primary datasets sourced from NYC Open Data, agency-internal systems, and research institutions. Each dataset contributes distinct spatial or attribute information to the final atlas.

Dataset	Source	Type	Contribution
Borough Boundaries	NYC Open Data	Polygon	Administrative context and clipping boundary for map display
Jamaica Bay Watershed Boundary	NYU Urban Data Studies	Polygon	Defines the study area extent; used as the primary clip boundary for all analyses
NYC Capital Projects Dashboard	NYC OMB	Tabular	Citywide capital project budget and schedule data with fiscal year attributes
DDC Capital Project Data	DDC (internal)	Tabular / Polygon	Agency-specific project records with detailed attributes; primary source for DDC projects
DPR Capital Project Tracker	NYC DPR	Point / Polygon	Parks department capital projects including park improvements and shoreline work
DEP Green Infrastructure Points	NYC Open Data / DEP	Point	Locations of DEP-managed green infrastructure installations across the watershed
TIGER/Line Shapefiles — NY Tax Blocks	US Census Bureau	Polygon	Tax block boundaries used as the unit of analysis for project concentration mapping

2.2 Data Limitations

Several constraints shaped the methodological approach and should be understood when interpreting the outputs of this project.

Incomplete Cross-Dataset Integration

Source datasets vary substantially in their completeness. Some contain spatial geometry but lack descriptive attributes; others carry detailed project-level metadata but no geometric component. This asymmetry required a multi-step integration process and means that some project records in the final atlas carry fewer attributes than others.

Duplicate Records Across Sources

Multiple source datasets contain overlapping records for the same physical projects — particularly where both agency-internal data and the citywide OMB dashboard reference the same capital investments. A systematic deduplication process was applied (described in Section 3.2), but some ambiguity may persist where project identifiers do not match consistently across sources.

Temporal Inconsistency

Fiscal year and completion date information are not reported consistently across all datasets. Some projects carry estimated completion years; others carry fiscal year windows; still others lack temporal data entirely. The time-slider feature in the web map reflects only those projects for which reliable temporal data was available.

CAUTION	Users should treat completion dates and fiscal year information as indicative rather than definitive. Agency contacts should be consulted to verify current project status before using this data to inform procurement or coordination decisions.

3. Methodology

The methodology comprises three sequential components: (1) capital project mapping and data integration, (2) cross-source deduplication, and (3) clustering analysis to quantify spatial project concentration at the tax-block level. All spatial processing was conducted in ArcGIS Pro; final outputs were exported to GeoJSON format for rendering in the Mapbox GL JS web environment.

3.1 Capital Project Mapping and Data Preparation

The first stage establishes the study area boundary, assembles source datasets, and integrates spatial and tabular capital project data into a unified feature class.

Step 1 — Define the Study Area

Borough boundary shapefiles were loaded into ArcGIS Pro to provide geographic context. The Jamaica Bay watershed boundary was obtained from NYU Urban Data Studies via ArcGIS Online and added using the Add Data from Path function. Because the source dataset contains geometries extending beyond the study area, the Copy Features tool was used to extract only the watershed polygon. The resulting clean polygon feature class (jmb_boundary) served as the clipping boundary for all subsequent spatial operations.

Step 2 — Review and Assess Capital Project Data

Three distinct capital project datasets were evaluated and found to have complementary rather than redundant content:

Dataset Type	Strengths	Limitations
Fiscal Year Dataset	Complete fiscal year coverage across projects	No spatial geometry; limited project-level attributes
Project Detail Dataset	Comprehensive attributes (agency, name, budget)	No fiscal year data; no spatial geometry
Polygon Dataset	Contains spatial geometries for capital projects	Incomplete attribute fields; no fiscal year tagging

This complementary structure required a multi-stage integration strategy rather than simple dataset merging. No single source provided a complete combination of spatial geometry, project attributes, and temporal data.

Step 3 — Clip and Filter Spatial Data

All spatial capital project datasets — comprising point, line, and polygon feature classes — were clipped to the Jamaica Bay watershed boundary using the ArcGIS Clip tool. A subset of records was then selected by managing agency using Select by Attributes with the expression:

```
magency IN ('DOT', 'DEP', 'DDC', 'DPR')
```

This filter ensured that only projects under the jurisdiction of the four target agencies were carried forward into analysis.

Step 4 — Data Cleaning and Integration

To consolidate the fiscal year and project detail datasets into a single authoritative record per project, the following steps were taken:

1. Project records were filtered to retain the lowest project number per fiscal year, establishing a single canonical entry per project per fiscal period.
2. A left join was performed with the fiscal year dataset as the base table and the project detail dataset joined to it.
3. The resulting unified dataset carries both fiscal year information and full project-level attributes for all matched records.

DATA	Unmatched records — where fiscal year data existed but no corresponding project detail record was found — were retained in the dataset with null attribute fields, ensuring no project is silently dropped from the spatial layer.

3.2 Cross-Source Deduplication

Because multiple source datasets contain overlapping records — notably the OMB Capital Projects Dashboard, the DCP Capital Project Tracker, and the DPR Capital Project Tracker — a systematic spatial deduplication process was applied before analysis. The process followed three steps.

Step 1 — Identify Spatially Overlapping Records

The Select by Location tool (Intersect criterion) was used to identify features in one dataset that spatially overlapped with features in another. This was applied across each pairing of the key datasets to build a complete picture of duplication.

Step 2 — Preserve Duplicate Records Before Deletion

Prior to removal, all identified duplicate features were exported into separate archive layers using Copy Features. This step maintains a traceable record of removed data and allows for future auditing or methodological review.

Step 3 — Remove Duplicates by Dataset Pair

Deduplication was applied across four dataset pairings, with the following results:

Dataset A	Dataset B	Method	Duplicates Found	Outcome
Capital_PARKS_jmb (389 records)	jmb_capitalproj_selection (238 records)	Intersect	151	151 duplicate records removed from Capital_PARKS_jmb
Capital_PARKS_jmb	jmb_capitalproj_pts_selection	Intersect	0	No spatial overlap detected; no records removed

Dataset A	Dataset B	Method	Duplicates Found	Outcome
IRF Project Joined Clip (DDC, 98 records)	jmb_capitalproj_selection	Completely Within	33	33 duplicate records removed from DDC dataset
IRF Project Joined Clip (DDC)	jmb_capitalproj_pts_selection	Intersect	0	No overlap detected; no records removed

NOTE	The 'Completely Within' criterion applied to DDC data is more conservative than Intersect, requiring that a feature be fully contained within another feature's boundary to be flagged as a duplicate. This approach reduces the risk of false-positive matches for features that only partially overlap.

3.3 Clustering Analysis

The clustering analysis quantifies the spatial concentration of capital projects across the watershed by aggregating all project geometries to a common unit of analysis: NYC tax blocks. This produces a block-level measure of infrastructure intensity that underlies the choropleth concentration layer in the web map.

Step 1 — Convert Polygon Projects to Points

Polygon-based capital projects were converted to representative point features using the Create Spatial Sampling Locations tool, enabling them to be aggregated alongside point and line features in subsequent steps. Sampling parameters were configured as follows:

Parameter	Value	Rationale
Sampling method	Systematic	Produces an evenly distributed grid of points within each polygon, avoiding spatial bias
Bin shape	Square	Consistent, uniform coverage across varied polygon geometries
Bin size	400 sq meters	Fine enough to capture within-block variation while remaining computationally tractable
Minimum distance between sample points	0 meters	Allows adjacent bins; density of output reflects polygon area proportionally

Step 2 — Standardize Point Geometry

The sampled polygon output may contain multipoint features — single records representing multiple coordinate pairs. These were converted to individual single-point features using the Multipart to Singlepart tool. This ensures that each record in the final point dataset represents exactly one project occurrence, preventing double-counting in block-level aggregation.

Step 3 — Aggregate Projects by Tax Block

The Summarize Within tool was applied to compute the number of capital project occurrences within each NYC tax block in the study area:

- Input polygon layer: nyblock_jmb (NYC tax blocks clipped to the Jamaica Bay boundary)
- Summarized feature layer: all capital infrastructure points combined from previous steps
- Output variable: proj_count — the total count of capital project occurrences per block

The resulting proj_count field serves as the foundation for the choropleth concentration layer rendered in the web map. Blocks are classified into five concentration categories using the natural breaks (Jenks) method, which identifies natural groupings in the data distribution:

Category	Description
Very Low	Isolated individual project; minimal or no spatial overlap with other agency work
Low	Small number of projects; limited multi-agency presence
Moderate	Meaningful concentration; at least two agencies represented in the block
High	Significant overlap; multiple concurrent projects across agencies
Very High	Highest concentration; strongest candidates for inter-agency coordination

Even though the clustering result was produced, it was not included into the final web map rendering since it provided a duplicate information as the capital infrastructure markers. This process, instead, confirms the spatial clustering of the capital infrastructure which should be emphasized as the core of this project.

4. Symbology and Visualization

4.1 Map Layer Structure

The web map is composed of four data layers, each representing a distinct type of capital project geometry or analytical output. Layers are rendered in a stacking order designed to keep contextual information visible beneath project-specific data.

Layer	Geometry Type	Description
Capital Project Points	Point	Individual project locations represented as dots; colored by managing agency
Capital Project Lines	Line	Linear infrastructure such as road corridors and utility routes; colored by agency
Capital Project Polygons	Polygon	Area-based projects including parks improvements and stormwater management zones; semi-transparent agency color fill
Jamaica Bay Boundary	Polygon	Outline of the study area; provides geographic orientation and clipping reference

4.2 Agency Color Classification

Each of the four managing agencies is assigned a distinct color applied consistently across all point, line, and polygon layers, as well as the interactive filter pills in the map legend.

Agency	Color	Infrastructure Domain
DEP	Blue	Water quality, stormwater management, and green infrastructure
DOT	Orange / Amber	Streets, transportation corridors, and public rights-of-way
DDC	Gray / Silver	Capital construction delivery across agency clients
DPR	Green	Parks, open space, shoreline, and greenway improvements

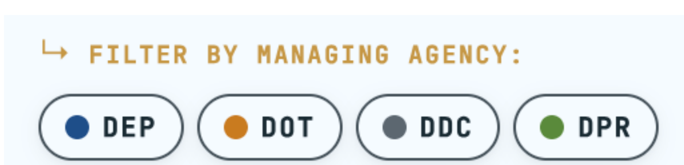

Figure 2 — Agency filter pills in the legend panel. Each pill toggles the corresponding agency's layers on or off across all geometry types.

4.3 Interactive Features

The map provides four categories of interactive functionality:

Feature	Description
Hover tooltips	Moving the cursor over any project marker, line, or polygon displays a popup showing the managing agency, project name, and estimated completion date
Agency filter pills	Tapping any of the four agency pills in the legend toggles that agency's projects on or off across all layers; multiple agencies can be active simultaneously
Cluster tour	A guided walkthrough of eleven pre-identified inter-agency overlap sites; each site includes a narrative description of the coordination opportunity
Map navigation	Standard pan, zoom, and reset controls; the map also auto-pans and zooms to center on each cluster tour site as the user navigates

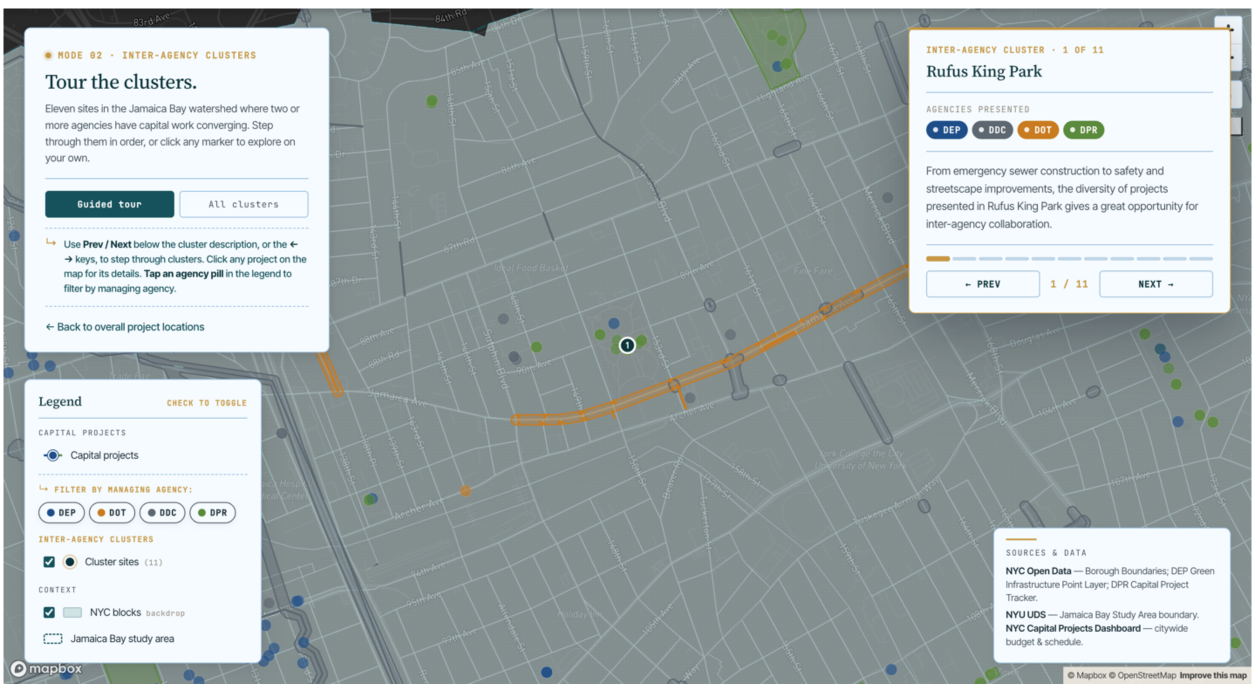

Figure 3 — The cluster tour side panel displaying site narrative and Prev / Next navigation controls

Figure 4 — A hover tooltip showing project-level details for a DEP green infrastructure marker

5. Technical Implementation

5.1 Technology Stack

Component	Technology	Role
Spatial Processing	ArcGIS Pro	Data integration, clipping, deduplication, clustering analysis, and GeoJSON export
Web Mapping Library	Mapbox GL JS	Interactive tile-based map rendering, layer management, and event handling
Frontend	HTML5, CSS3, JavaScript (ES6+)	Page structure, visual styling, and application logic
Data Format	GeoJSON	Lightweight vector format for serving spatial data to the browser
Code Repository	GitHub	Version control and static site hosting via GitHub Pages

5.2 Project File Structure

The codebase follows a modular structure that separates concerns across dedicated files:

File / Directory	Purpose
index.html	Main application entry point; defines the page layout, map container, legend panel, and cluster tour panel
style.css	Visual styling for all interface elements including the legend, tooltips, tour panel, and responsive layout
map.js	Core map logic: layer loading, filter state management, cluster tour navigation, and popup event handlers
popup-content.js	Templates for hover tooltip content, assembled dynamically from GeoJSON feature properties
data/	Directory containing all GeoJSON datasets exported from ArcGIS Pro

The repository is publicly accessible at: https://github.com/ctohsuwanwanich/ddc_webmap

5.3 Data Pipeline Summary

The end-to-end data pipeline moves from raw agency datasets through GIS processing to browser-ready GeoJSON files:

1	Source datasets collected from NYC Open Data, DDC internal systems, DPR, DEP, and NYU Urban Data Studies
2	Data loaded into ArcGIS Pro; study area boundary extracted and all datasets clipped to Jamaica Bay extent
3	Agency filter applied (DOT, DEP, DDC, DPR); fiscal year and project detail datasets joined
4	Cross-source deduplication applied across four dataset pairings; 184 duplicate records removed total
5	Polygon projects converted to systematic sample points; single-part conversion applied
6	Summarize Within analysis run on combined point dataset against NYC tax blocks; proj_count field generated
7	All layers exported from ArcGIS Pro as GeoJSON; placed in the data/ directory of the web repository
8	Mapbox GL JS renders layers in the browser; interactive features connected via JavaScript event handlers

6. Limitations

The following limitations should be considered when using this atlas for planning or research purposes:

Uneven Data Quality Across Agencies

Source datasets vary in completeness and standardization. DEP and DPR datasets are relatively well-structured with consistent spatial geometry and attribute coverage. DDC internal data required more extensive processing to integrate. DOT data relied heavily on the citywide OMB dashboard, which aggregates projects at a coarser resolution than agency-specific systems.

Inconsistent Completion Date Reporting

Temporal data, including estimated completion years and fiscal year designations, is not uniformly reported across all datasets.

Analysis Limited to Tax Block Scale

The clustering analysis aggregates projects to the NYC tax block level. This is appropriate for identifying neighborhoods or corridors of concentrated activity, but cannot resolve intra-block variation or differentiate between projects that occupy the same block but do not physically intersect. Street-level or parcel-level resolution would require a more granular unit of analysis.

Polygon Features Simplified to Points

Large-area projects (parks, stormwater zones) are represented by systematically sampled points for the purposes of block-level aggregation for the clustering analysis. While this preserves spatial distribution across a polygon's footprint, it does not fully capture the project's extent or edge geometry in the concentration analysis.

Static Dataset

The current version of the atlas is based on a static snapshot of data collected in early 2026. Capital project portfolios change continuously as projects are funded, initiated, and completed. The data should be treated as a point-in-time representation rather than a live planning record.

Partial Duplication

Due to the limitations of the intersect tool, the current efforts from the deduplication process could not completely eliminate all duplicates across datasets from different agencies. A further methodology exploration is needed to offer a better steps to deduplicate layers with more concise manners to avoid confusion and overrepresentation in the visualization.

7. Next Steps and Future Development

This project establishes a foundation that can be substantially extended in subsequent phases. Recommended directions for future development include:

Live Data Integration

Connecting the map to live API feeds from NYC Open Data and agency dashboards would enable the atlas to refresh automatically as project records are updated, eliminating the need for manual data exports and maintaining currency with agency portfolios.

Budget-Based Analysis

Incorporating capital budget data from the NYC OMB Capital Projects Dashboard would allow the atlas to visualize not only where projects overlap but also the combined financial scale of co-located investments, providing a stronger basis for making the case for coordination.

Cross-Agency Coordination Metrics

Developing quantitative metrics for coordination potential, incorporating factors such as project timeline overlap, proximity, shared infrastructure type, and combined budget, would make the tool more directly actionable for agency planners and procurement staff.

Dataset Validation with Agencies

A structured validation process with DEP, DOT, DDC, and DPR project managers would improve data accuracy, resolve ambiguities in cross-source matching, and build agency trust in the tool as a planning resource.

Expanded Study Area

The methodology developed for Jamaica Bay could be extended to other watershed areas or planning districts across New York City, enabling comparison of inter-agency coordination patterns across different urban contexts.

8. Conclusion

This project demonstrates how spatial data integration can surface patterns of infrastructure investment that remain invisible within the fragmented systems of individual agencies. By bringing together capital project data from DEP, DOT, DDC, and DPR into a single geospatial framework, the Jamaica Bay Integrated Water Management Atlas enables a kind of spatial legibility that has not previously existed for this watershed.

The eleven inter-agency cluster sites identified through the analysis represent tangible opportunities for more coordinated infrastructure delivery. In each case, multiple agencies are investing public resources in the same physical areas, often within the same fiscal periods. Whether those investments are designed and executed in coordination or in isolation is a planning decision, and this tool is intended to inform that decision with evidence.

More broadly, the project illustrates a replicable methodology: assemble fragmented agency data, establish a common spatial framework, apply systematic deduplication and aggregation, and surface the results through an accessible web interface. The same approach could be applied across New York City or other municipalities where infrastructure planning responsibilities are distributed across multiple agencies operating in shared urban environments.

References

- NYC Open Data — Borough Boundaries. https://data.cityofnewyork.us/City-Government/Borough-Boundaries/gthc-hcne/about_data
- NYC Department of Environmental Protection — Green Infrastructure Dataset. https://data.cityofnewyork.us/Environment/DEP-Green-Infrastructure-Point-Layer-/df32-vzax/about_data
- NYC Department of Parks and Recreation — Capital Project Tracker. https://data.cityofnewyork.us/Recreation/Capital-Project-Tracker/qiwj-i2jk/about_data
- NYC Office of Management and Budget — Capital Projects Dashboard. https://app.powerbigov.us/view?r=eyJrIjoiMTkwYWMyNGEtMDNiZC00OTY4LTk4YjEtYzI0MzhlOTA3MzIiIiwidCI6IjM1YzgyODE2LTZjNTYtNDQzYi1iYWY2LTgzMTIxNjNjYWRjMSJ9
- NYU Urban Data Studies — Jamaica Bay Watershed Boundary. ArcGIS Online. https://nyuds.maps.arcgis.com/home/item.html?id=934450c68a534a67b05878eed18d0235#overview
- US Census Bureau — TIGER/Line Shapefiles: New York Blocks.

Appendix — Web Map User Guide

The following pages provide a concise reference guide to navigating the interactive web map. For the complete standalone user guide document, refer to the standalone user guide document for this project.

A.1 Accessing the Map

The map is available at: https://ctohsuwanwanich.github.io/ddc_webmap/

No login or software installation is required. The map runs in any modern desktop web browser (Chrome, Firefox, Safari, Edge). An active internet connection is required to load the basemap tiles.

A.2 Entry Points

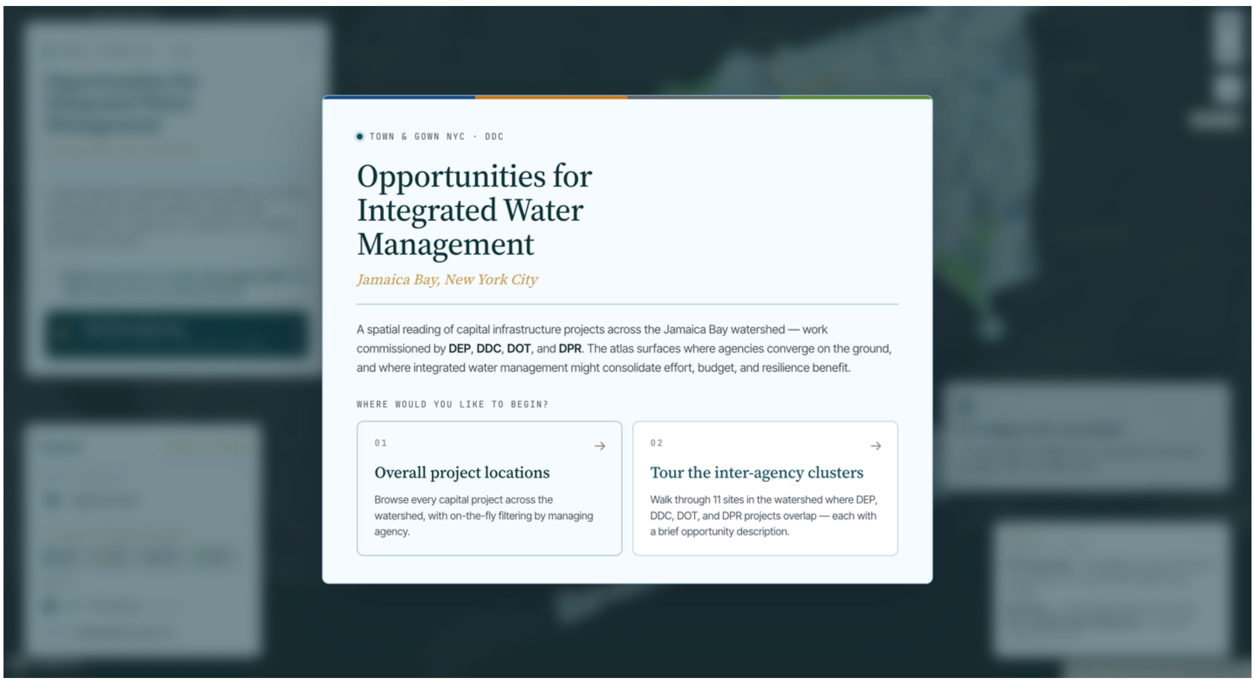

Figure A1 — The landing screen presents two entry options: free project exploration or the guided cluster tour

01	Overall project locations — Enters the map in free-exploration mode. All agency layers are active and can be filtered using the legend pills.
02	Tour the inter-agency clusters — Enters the guided tour of 11 inter-agency overlap sites, each with a narrative description.

A.3 Quick Reference

I want to…	How
See all capital projects	Open the map — all four agency filter pills are active by default
Focus on one agency's projects	Click the other three agency pills to deactivate them
Read details about a project	Hover the cursor over (or tap) any marker, line, or polygon
Start the cluster tour	Click '◐ Start the cluster tour' in the top navigation bar
Move between cluster sites	Use the ← Prev and Next → buttons in the tour side panel
Close the cluster tour	Click the × button in the tour panel to return to free exploration
Toggle context layers	Use the checkboxes under 'Context' in the legend panel
Return to the landing screen	Refresh the page in your browser