

December 1, 2016

Dear Mayor de Blasio:

Pursuant to Local Law 8 of 2016, you tasked the Mayor’s Office of Data Analytics (MODA) with conducting a series of examinations and verifications on mayoral agencies to ensure their compliance with Local Law 11 of 2012, also known as the Open Data Law. As New York City’s Chief Analytics Officer, Chief Open Platform Officer, and Director of MODA, I hereby submit my office’s findings on the Department of Sanitation (DSNY), Department of Correction (DOC), and Department of Housing Preservation and Development (HPD). This report also includes recommendations on improving the City’s overall compliance with the Open Data Law in line with our vision of *Open Data for All*.<sup>1</sup>

Please find enclosed:

- Examination and Verification 2016 Findings Report
  - Process and Complications
  - Summary of Results
  - Recommendations for Better Citywide Compliance
  - Appendix A: Local Law 8 of 2016
  - Appendix B: MODA Examination and Verification Plan
  - Appendix C: Department of Investigation Letter of Approval
  - Appendix D: Executive Certification Letter template
- Examination and Verification 2016 Results Workbook
  - Department of Sanitation
  - Department of Correction
  - Department of Housing Preservation and Development

As specified in Local Law 8 of 2016, these findings will be made publicly available on nyc.gov within ten days.

Our ongoing partnership with the Open Data Coordinators and other Open Data stewards at the participating agencies is helping MODA build a better examination and verification process for subsequent years. We wish to thank them for their cooperation and will continue to collaborate on identifying and publishing valuable datasets to the Open Data Portal.

Respectfully Submitted,



Dr. Amen Ra Mashariki  
Chief Analytics Officer and Chief Open Platform Officer

cc: James Vacca, Chair, City Council Committee on Technology  
Kathryn Garcia, Commissioner, Department of Sanitation  
Joseph Ponte, Commissioner, Department of Correction  
Vicki Been, Commissioner, Department of Housing Preservation & Development  
Mark Peters, Commissioner, Department of Investigation

---

<sup>1</sup> Mayor’s Office of Data Analytics, *Open Data for All* (July 2015):  
<http://www1.nyc.gov/assets/home/downloads/pdf/reports/2015/NYC-Open-Data-Plan-2015.pdf>

## EXAMINATION AND VERIFICATION 2016 FINDINGS REPORT

### *Process and Complications*

In December 2015, the New York City Council voted in favor of Intro No. 916-A, a law requiring an agency designated by the Mayor to conduct examinations and verifications of the compliance of certain mayoral agencies with the Open Data Law. The purpose of this law is to improve citywide compliance by creating a more systematic way to locate datasets that may have inadvertently or purposefully been excluded in agencies' self-reported Open Data compliance plans.

In January 2016, Mayor de Blasio signed Int. No. 916-A and it became Local Law 8 of 2016 (Appendix A). He designated the Mayor's Office of Data Analytics (MODA) to conduct this process. MODA then prepared an Examination and Verification plan (Appendix B), which was approved by the Commissioner of the Department of Investigation (Appendix C).

The Department of Sanitation (DSNY), Department of Correction (DOC), and Department of Housing Preservation and Development (HPD) were the three agencies named for the first round of the Examination and Verification process.

MODA's plan required these agencies to assemble the following items:

- Dataset questionnaire
- Executive Certification Letter (Appendix D)
- Public nominations

The dataset questionnaire familiarized MODA with each agency's current Open Data footprint, routine information reporting requirements, data management systems, and organizational structure. In the certification letter, an executive at the agency attested to the accuracy and completeness of the information provided.

From October 28 to November 14, 2016, the Open Data team invited the public to suggest additional datasets for consideration. While users always have the option to nominate datasets for publication on the Open Data Portal, this window for public feedback specifically invited the public to participate in the Examination and Verification process.

Based on this information, MODA expected to assemble a comprehensive inventory of eligible datasets within the surveyed agencies that MODA analysts could determine to be "public" or "private." The "public" datasets that had not previously been disclosed in the agencies' compliance plans would be named in this report and subjected to future publication on the Open Data Portal.

This proved more complicated than the plan intended. Creating a list of eligible datasets is not a cut and dried process, and raised questions regarding the definitions of "public," "data," and "dataset." The way data is represented reflects a series of decisions about the collection, organization, and depiction of digital information; distilling a stable "dataset" within this series is often a complex matter of discretion involving an array of actors.

Consider a few examples:

- The **Department of Sanitation (DSNY)** publishes monthly garbage collection statistics in PDFs on its website. These *reports* are formatted as tables of statistics, which are surfaced from underlying data. Is the “dataset” in question the tables of statistics in the reports, or the unstructured information that is aggregated into reported metrics?
- The **Department of Correction (DOC)** maintains the “Inmate Information System,” a jail management technology that is rife with Personally Identifiable Information. This data is aggregated and reported as an indicator in the Mayor’s Management Report (MMR). If a properly de-identified dataset might be contrived, but does not yet exist, is it eligible for publication on Open Data?
- The **Department of Housing Preservation and Development (HPD)** conducts building inspections, some of which result in violations issued. The agency maintains an “inspections file” and a “violations file.” Violations data is published on the Open Data Portal; inspections data is not. When a data source represents similar or redundant information to data already on the Portal, should it be published?

Datasets are like waves: it is not always clear where one ends and the other begins.

These questions warrant further consideration and clarification. We outline these challenges and proposals for steps forward in the “Recommendations for Better Citywide Compliance” section of this report.

### *Summary of Results*

Itemized results from the agency surveys can be found in the attached “Examination and Verification 2016 Results Workbook.” This information is meant to give users a snapshot of the technical environment of the agency and a better understanding of how data from an agency’s data system becomes a usable dataset on the Open Data Portal.

We encourage members of the public to make use of this information and the dataset nomination process – which guarantees a formal review and timely response by Local Law 109 of 2015 – to help us push the City closer to fulfilling the intention of the Open Data Law.

In summary, we found:

- All three agencies are in good standing with the Open Data Law.
- Several data systems *may* contain data that can be published on the Open Data Portal, but warrant further review.
- Three dataset nominations received during the public feedback window were referred to HPD. One was determined to be a dataset managed by the Department of Finance, one was determined to not be an existing dataset, and one is under further review.
- None of the agencies listed any new datasets determined to be public through the FOIL review process.

### Results Snapshot

	DSNY	DOC	HPD
<b>Total Open Data Datasets</b>	15	12	9
Currently on Open Data Portal (ODP)	14	9	7
Planned for future release	1	3	2
<b>Datasets currently on ODP - Automations</b>			
Automated	14	9	6
Non-automated	0	0	1
<b>Datasets currently on ODP - Update Frequency</b>			
Biannually	0	0	1
Quarterly	0	1	0
Monthly	7	7	6
Weekly	1	0	0
Daily	0	1	0
As needed	6	0	0
<b>Data associated with MMR indicators<sup>2</sup></b>			
On ODP or future release	1	2	5
Public but not on ODP or future release	14	3	3
Private	0	7	3
<b>Public requests for data</b>			
Already available on ODP	6	0	2
Potentially new public data	2	0	1
Non-public data	0	0	0
Not agency data	2	0	3

<sup>2</sup> Agencies were required to report names of datasets associated with each of their MMR indicators. The number of *clearly public* datasets is totaled in “On ODP or future release.” The number of *clearly private* datasets is totaled in “Private.” The datasets totaled in “Public but not on ODP or future release” are less definitive. Many do not refer to a specific dataset, but a database or application that requires further investigation as to whether a dataset eligible for publication on Open Data could be produced. For more detail, see the Examination and Verification 2016 Results Workbook.

*Recommendations for Better Citywide Compliance*

Up until this year, the Open Data Law required City agencies to self-submit compliance plans that laid out a timeline for publishing “public datasets.” After the package of amendments to the Open Data Law was passed during the last year, the agency compliance plan became supplemented by the following legal mandates:

- **Public requests:** Local Law 109 of 2015 guarantees timely and thorough responses to all public requests for new datasets on the Open Data Portal.
- **Timely updates:** Local Law 110 of 2015 requires all data published on agency websites to be included and kept up-to-date on the Open Data Portal.
- **FOIL responses including data:** Local Law 7 of 2016 requires agencies to review Freedom of Information Law (FOIL) requests containing data to determine whether they contained *new* public datasets that could be published on the Open Data Portal.
- **Examinations and Verifications:** Local Law 8 of 2016 requires MODA to examine three mayoral agencies each year to verify that all public datasets have been disclosed.

These statutory measures collaborate to form a framework for locating data that are, by Local Law 11 of 2012, eligible for publication on the Open Data Portal by the end of 2018. Over and above the specific statutory mandates for Open Data compliance, it is incumbent on MODA, City agencies, and other citywide actors to fill in other identified gaps in the Open Data program.

To that end, the Examination and Verification law requires that the office or agency conducting the examinations and verifications “make recommendations to improve the disclosure and inclusion of all public data sets required to be on the single web portal.” We outline a series of specific recommendations below.

1. **Agencies should make their technical ecosystems more accommodating to Open Data by:**
  - i. Using automations, rather than manual uploads, to update datasets currently on the Open Data Portal.
  - ii. Writing Open Data requirements into procurements of new data systems and analytics technologies.
  - iii. Allocating more resources to Open Data personnel, especially Open Data Coordinators.
2. **The Open Data team should empower Open Data Coordinators by:**
  - i. Surveying Open Data Coordinators to better understand their roles, priorities, and communication preferences.
  - ii. Producing documents clarifying the roles and responsibilities of Open Data Coordinators, including guidelines on complying with legal mandates.
3. **MODA should improve the Examination and Verification plan for future years by:**
  - i. Consulting with the Department of Investigation on potential improvements.
  - ii. Creating clear guidelines and definitions of “data” and “dataset.”
  - iii. Creating clear guidelines on determining whether a dataset is “public” or “private.”

Opening data is not a one-off obligation: it is an ongoing, virtuous cycle that creates operational efficiency and communication across silos of government. As Open Data becomes the norm for city data, it makes agencies more aware of the data they have and the data they produce – and spurs better upkeep and disclosure of information. As Open Data shifts into an established, routine process across NYC government, we will continue to look for ways to try out bold new ideas, as well as opportunities to make incremental adjustments to existing processes.

**APPENDIX A: Local Law 8 of 2016**

*Passed by New York City Council on December 16, 2015 and approved by the Mayor on January 5, 2016.*

Introduced by Council Members Vacca, Koo, Constantinides, Greenfield, Kallos, Mealy and Vallone.

**A LOCAL LAW****In relation to an open data law agency compliance examination.**

*Be it enacted by the Council as follows:*

Section 1. Open data law agency compliance examination. a. An office or agency designated by the mayor shall conduct a series of examinations and verifications, as described in subdivision c, and make recommendations to improve the disclosure and inclusion of all public data sets required to be on the single web portal pursuant to section 23-502 of the administrative code of the city of New York.

b. Within 60 days of the effective date of this local law, an office or agency designated by the mayor shall present to the commissioner of investigation a plan for conducting the examinations and verifications described in subdivision c. The commissioner of investigation shall review such plan to ensure that it conforms with either a generally accepted auditing process or a process that the department of investigation would itself use in such an examination. The commissioner of investigation shall report to both the mayor and the council when a plan has been approved. The office or agency designated by the mayor may amend the plan with the approval of the commissioner of investigation.

c. Not later than December 1, 2016, and each December 1 thereafter for the next two years, the office or agency designated by the mayor shall conduct an examination and verification of the compliance with the requirements of subdivision a of section 23-502 of the administrative code of the city of New York, of no less than three mayoral agencies and submit the findings of such examination and verification to the mayor, the council and the examined mayoral agencies. Such findings shall include a list of all public data sets that such mayoral agencies did not make available on the single web portal in accordance with subdivision a of section 23-502 of the administrative code of the city of New York or disclose in the agency compliance plan required by section 23-506 of the administrative code of the city of New York as of the date of the findings, as well as a description of any deviations in the examination and verification process from the plan approved pursuant to subdivision b. For the findings due December 1, 2016, the mayoral agencies examined shall at a minimum consist of the department of sanitation, the department of correction and the department of housing preservation and development. For the findings due December 1, 2017, the mayoral agencies examined shall at a minimum consist of the department of buildings, the department of environmental protection and the fire department. For the findings due December 1, 2018, the mayoral agencies examined shall at a minimum consist of the business integrity commission, the department of transportation and the department of small business services. In preparing such findings, the office or agency designated by the mayor shall accept suggestions from the public as to possible public data sets within mayoral agencies that have not yet been disclosed.

d. Not later than December 1, 2019, the office or agency designated by the mayor shall submit a written report to the mayor and the council describing the city's compliance with the requirements of subdivision a of section 23-502 of the administrative code of the city of New York, including a complete list of public data sets discovered by the office or agency designated by the mayor that were not previously made available on the single web portal or disclosed in the agency compliance plan and recommendations to improve the disclosure and inclusion of all public data sets required to be on the single web portal. In preparing this report, the office or agency designated by the mayor shall also accept suggestions from the public as to possible public data sets within mayoral agencies that have not yet been disclosed.

e. The report and findings required by this local law shall be posted on the city's website no later than ten days after being submitted.

§ 2. This local law takes effect immediately.

**APPENDIX B: MODA Examination and Verification Plan**

In order to fulfill the requirements of Intro 916-2015, the Mayor’s Office of Data Analytics (MODA) proposes a three-part Examination and Verification Plan.

For each agency named in Intro 916-2015, the examination and verification will consist of the following:

- I. Agency questionnaire
- II. Public feedback
- III. Agency affirmation / certification

As described in Intro 916-2016, for the report due December 1, 2016, the agencies examined will consist of the Department of Sanitation, the Department of Correction, and the Department of Housing Preservation and Development.

**AGENCY QUESTIONNAIRE**

The questionnaire is designed to acquaint the office conducting the examination and verification with basic information about the processes, people, and technical infrastructure associated with data at the agency.

As stated in Local Law 11 of 2012: For purposes of prioritizing public data sets, agencies shall consider whether information embodied in the public data set:

- (1) can be used to increase agency accountability and responsiveness;
- (2) improves public knowledge of the agency and its operations;
- (3) furthers the mission of the agency;
- (4) creates economic opportunity; or
- (5) responds to a need or demand identified by public consultation.

The questionnaire draws upon the priorities above, other Open Data legislation with complementary reporting requirements, and the annual guidance given to Open Data coordinators before the compliance plan update. Sample questions are available in Appendix A.

**PUBLIC FEEDBACK**

For a pre-announced amount of time, the public will be able to nominate any additional datasets for the scheduled agencies. This nomination ability exists currently, but will be publicized on the Open Data portal. An example of the current “Suggest a Dataset” feature on the portal is included in Appendix B.

**AGENCY AFFIRMATION**

The agency affirmation is designed to provide an additional layer of agency certification that they are compliant with Local Law 11 of 2012.

The affirmation letter is modeled after the original letters sent to commissioners and agency heads before the first release of the Open Data compliance plan in 2013.

**APPENDIX C: Department of Investigation Letter of Approval**



The City of New York  
Department of Investigation

MARK G. PETERS  
COMMISSIONER

80 MAIDEN LANE  
NEW YORK, NY 10038  
212-825-5900

TO: Bill de Blasio, Mayor, City of New York  
Melissa Mark-Viverito, Speaker, New York City Council

cc: Mindy Tarlow, Director, Mayor's Office of Operations

FROM: Mark G. Peters, Commissioner, NYC Department of Investigation

RE: Local Law 8 of 2016 (Council Int. No. 916-A of 2015)  
In relation to an open data law agency compliance examination

DATE: November 23, 2016

A handwritten signature in blue ink, appearing to read "MGP", is written over the "FROM:" line of the memorandum.

---

The purpose of this Memorandum is to advise the Mayor and the City Council that, pursuant to Subdivision (b) of Local Law 8 of 2016, I have reviewed the examination and verification plan presented by the Mayor's Office of Data Analytics and approve the use of such plan.

**APPENDIX D: Executive Certification Letter template**

As the Chief Executive Officer of \_\_\_\_\_, or their designee, I do hereby certify that, to the best of my knowledge, information, and reasonable belief, the attached inventory, submitted pursuant to said agency’s obligations under Local Law 8 of 2016, represents complete and accurate information for each survey question.

Furthermore, I affirm that this inventory includes a summary description of all public data sets under the control of said agency as of the date of the findings, including a list of all public data sets such that said agency did not make available on the single web portal in accordance with Local Law 11 of 2012.

If such public data set or sets cannot be made available on the Open Data portal on or before December 31, 2018, this plan states the reasons why such set or sets cannot be made available, and, to the extent practicable, the date by which the said agency believes that it will be available on the single web portal.

---

Signature

---

Printed Name

---

Title

---

Date