

MARCH 2020

CITY OF CHICAGO OFFICE OF INSPECTOR GENERAL

ADVISORY CONCERNING THE CITY OF CHICAGO DATA PORTAL





JOSEPH M. FERGUSON
INSPECTOR GENERAL

CITY OF CHICAGO
OFFICE OF INSPECTOR GENERAL
740 NORTH SEDGWICK STREET, SUITE 200
CHICAGO, ILLINOIS 60654
TELEPHONE: (773) 478-7799
FAX: (773) 478-3949

VIA ELECTRONIC MAIL

Nicholas Lucius
Chief Data Officer
Office of the Mayor
121 North LaSalle Street, 4th Floor
Chicago, Illinois 60602

David Reynolds
Commissioner
Department of Assets, Information, and Services
30 North LaSalle Street, Suite 300
Chicago, Illinois 60602

Dear Chief Data Officer Lucius and Commissioner Reynolds:

An OIG inquiry has determined that the City does not comply with the requirements of Executive Order No. 2012-2 (E.O. 2012-2 or “the Order”), which governs the administration of the Chicago Data Portal. Additionally, in the course of several OIG projects, we have observed multiple issues that impact the Portal’s usability. As the City reengages with its open data strategy via the efforts of the Department of Assets, Information, and Services (AIS) and the movement of senior technology officers to the Mayor’s Office, we encourage it to focus on both meeting existing obligations and reviewing open data processes to improve accessibility, transparency, and accountability.

I. BACKGROUND

The Mayor’s Office formalized the City of Chicago Data Portal under E.O 2012-2, which committed the City to “an unprecedented level of transparency, honesty and accountability” in its treatment of public data.¹ The Data Portal allows the public to “freely download data” for analysis.² As of January 2, 2020, the Portal housed 488 datasets covering a wide range of topics, including building permit and inspection

¹ City of Chicago, “Executive Order 2012-2 (City of Chicago Data Portal),” December 12, 2012, accessed January 2, 2020, https://www.chicago.gov/city/en/narr/foia/open_data_executiveorder.html.

² City of Chicago, “Chicago Data Portal,” accessed January 2, 2020, <https://data.cityofchicago.org/>.

information, crime incidents and locations, City lobbyist registries, employee salaries, and Freedom of Information Act request logs.³

As part of Mayor Lightfoot's transition into office, the *Report of the Transition Committees*, published in May 2019, identified "Transparency Through Open Data" as a means of readjusting the City's culture to one of "radical transparency."⁴ The transition committee suggested that the new administration,

- introduce an ordinance to codify Open Data E.O. 2012-2; and
- direct the City to make public transparency the default position, limiting use of exemptions to rare, narrow circumstances, even where a legal basis for withholding data might exist.

In January 2020, the Departments of Innovation and Technology (DoIT) and Fleet and Facility Management merged to become AIS, and the City's chief data officer (CDO) and chief information officer (CIO) positions moved into the Mayor's Office to spearhead a "renewed open data strategy."⁵

E.O. 2012-2 requires City agencies to perform certain tasks to achieve better public data transparency. Some of these tasks enable communication between agencies to develop and implement a comprehensive open data strategy, while others support communication to the public about City open data initiatives.

AIS, as the successor to DoIT, is responsible for the continuing administration of the Data Portal pursuant to E.O. 2012-2. As of November 15, 2019, however, only one staff member was assigned to support the Portal. Moreover, that person had additional responsibilities unrelated to Portal administration. The Order requires the CDO to "coordinate implementation, compliance, and expansion of the City's open data policy," and to chair an Open Data Advisory Group comprised of open data coordinators from each City agency.⁶ Departments are responsible for identifying appropriate datasets to post on the Portal, taking into consideration whether making

³ For the purposes of this report, "datasets" are defined as online data presented in tabular format. Other data is available and presented in the form of maps, charts, or other data "lenses."

⁴ Office of the Mayor, "Report of the Transition Committees," p. 21, published May 17, 2019, accessed December 18, 2019, https://www.chicago.gov/city/en/depts/mayor/supp_info/transition-report.html.

⁵ Office of the Mayor, "Mayor Lightfoot Announces Proposal to Merge Departments of Innovation and Technology And Fleet and Facility Management in 2020," October 9, 2019, accessed December 18, 2019, https://www.chicago.gov/city/en/depts/mayor/press_room/press_releases/2019/october/ProposalDoIT2FM_Merger.html. As of January 15, 2020, the CDO position has been filled but the CIO position remains vacant.

⁶ The Order defines "City agency" as "any city department, office, administrative unit, commission, board, advisory committee or other division of the government of the City within the official jurisdiction of the executive branch."

a particular dataset publicly available increases accountability, efficiency, responsiveness, service delivery, or public knowledge of the agency and its operations.

II. THE CITY IS NOT COMPLYING WITH THE OPEN DATA EXECUTIVE ORDER.

The City has not met some of its core requirements to ensure data transparency and a unified open data strategy. In particular,

- the Open Data Advisory Group has not met since 2014 and, at that time, only included representatives from 18 of 29 City agencies;⁷
- the City has not published an Open Data Compliance Report since 2014; and
- the CDO position was vacant for 18 months, from June 2018 until January 2020.

Without a CDO or Open Data Advisory Group, the City has had limited means to implement and expand its open data strategy, of which the Data Portal is a central component. An advisory group including representatives from all agencies would enable the development of a cohesive, Citywide open data strategy. Without representatives from all agencies, the City lacks a complete picture of each agency's data and needs. This impairs the effectiveness of any policies the City may create. Finally, the continuing lack of an annual Open Data Compliance Report prevents the public from learning about the City's open data efforts.

Prior to the creation of AIS, DoIT executive management stated that it began Citywide engagement efforts in summer 2019 and was attempting to convene a group of active data users across the agencies. DoIT stated it planned to leverage these users to reach out to unrepresented agencies, identify potential open data coordinators, and reconvene the Open Data Advisory Group.

III. OIG HAS FOUND NUMEROUS PROBLEMS WITH THE COMPLETENESS, ACCURACY, AND USABILITY OF THE DATA PORTAL.

As an active user of the City Data Portal, OIG has found problems with the completeness, accuracy, and usability of datasets on the Portal. These problems

⁷ The 29 agencies represent those in existence in 2014 and exclude OIG and the Chicago Board of Election Commissioners. The 2014 Advisory Group did not include representatives from the Departments of Human Services, Cultural Affairs and Special Events, Public Library, Police, Fire, Buildings, Animal Care and Control, Aviation, nor the Police Board, Independent Police Review Authority, or License Appeal Commission.

prevent users of the Data Portal from accessing relevant City information and reaching accurate conclusions about City initiatives and resources. The creation of AIS and centralization of the CDO and CIO positions presents a unique opportunity to address issues with the Portal and strengthen aspects the City's pre-existing open data initiatives.

A. THE CITY OMITTS SOME RECORDS WITHOUT NOTING THEY ARE MISSING OR PROVIDING A RATIONALE FOR THEIR OMISSION.

Over the course of its work, OIG has determined that several datasets exclude relevant fields or records. However, DoIT and City agencies could not always explain why data was excluded. For example,

- A 2019 audit of the Chicago Department of Public Health's (CDPH) environmental permitting and enforcement activities found that the Environmental Complaints and Environmental Inspections datasets on the Portal were missing records.⁸ CDPH discovered that a source data query⁹ had caused the missing inspection records, but the Department could not identify a reason for the missing complaint records.
- A 2018 audit found that the Department of Buildings (DOB) Building Violations dataset excluded certain inspection records.¹⁰ Neither DoIT nor DOB were able to provide a rationale for excluding those types from the source query. In a 2019 follow-up, DOB acknowledged "significant problems" with both the Building Permits and Building Violations datasets on the Portal.¹¹ DOB management acknowledged these were due to issues with the "filtering rules and reporting logic" underlying those datasets, but stated it could not resolve these issues until it was able to address source technology obstacles.

Per Executive Order 2012-2, AIS and City agencies are jointly responsible for determining what information to post on the Portal, based on the principle that all data should be public unless a reason exists to limit its availability. The decentralized

⁸ Chicago Department of Public Health Air Pollution Enforcement Audit (#17-0525), published September 16, 2019, <https://igchicago.org/2019/09/16/chicago-department-of-public-health-air-pollution-enforcement-audit/>.

⁹ A "query" is the underlying computer code designed by DoIT that pulls data from internal agency databases to populate datasets on the Portal.

¹⁰ Department of Buildings Complaint-Based Inspections Audit (#16-0301), published April 10, 2018, <https://igchicago.org/2018/04/10/dob-complaint-based-inspections-audit/>.

¹¹ Department of Buildings Complaint-Based Inspections Follow-Up Inquiry (#19-0589), published November 19, 2019, <https://igchicago.org/2019/11/19/department-of-buildings-complaint-based-inspections-follow-up-inquiry/>.

nature of the Portal, with AIS collecting and posting data from other agencies, requires clear communication and knowledge sharing. When public datasets are missing fields or records, they are less useful and less accessible, and can result in inaccurate conclusions to public inquiries.

Although AIS must collaborate with agencies to determine what data to make available on the Portal, prior to the merger, DoIT stated that it did not have a standardized process to help agencies make these determinations, nor did it have a standardized system for documenting query designs to prevent unintended data omissions. Instead, DOIT memorialized query design decisions in emails to a central account as a method of keeping the information available in the event of staffing changes.

The Open Data Executive Order empowers the CIO with the “authority to promulgate reasonable rules and regulations to implement the requirements of this Order.” However, no such rules and regulations have been promulgated, and the CIO position is currently vacant. Prior to the departmental merger, DoIT stated it had taken initial steps to develop and improve the City’s open data strategy. For example, the Department had contracted with a vendor to “perform a full-scale IT assessment” in 2020 that would include creating “a [Citywide] inventory of technology systems and their associated datasets.” DoIT further stated that, where possible, it had configured datasets to update automatically.¹²

B. THE CITY DOES NOT CONSISTENTLY COMMUNICATE CHANGES THAT MAY IMPACT DATASET USABILITY.

Although the City uses various methods to notify the public of changes to the structure, queries, or general configuration of datasets, these notifications are not consistently clear about which datasets have been changed.

The City publishes descriptions of dataset changes via two primary methods—its Developers website and the Chicago Digital Hub.¹³ The Developers website, which communicates technical information and changes to datasets, is designed primarily for software developers. In contrast, DoIT management described the Digital Hub (last updated in December 2017) as a place to provide more general information, like

¹² DoIT management stated it has already configured 103 datasets with automatic updates. Management noted that there are many datasets on the Portal which are static and do not need to be updated (such as budget ordinance data), and many which cannot automatically update due to their format (for example, data stored in Excel spreadsheets).

¹³ See <http://dev.cityofchicago.org/blog/> and <https://digital.cityofchicago.org/>. DoIT executive management stated that it uses its judgment to determine whether revisions are too minor to be publicized, but “if a dataset revision could be interpreted as calling into question our commitment to transparency, we try to communicate it no matter how minor or transient.”

original design decisions behind certain datasets. DoIT stated that, as the Portal has evolved, most updates are posted to the Developers website; the Digital Hub is primarily a historical record of prior communications.

Recent blog posts on the Developers website highlight changes to datasets and link directly to them. However, the datasets on the Portal do not consistently link back to the relevant blog posts explaining the changes. As of December 2019, for example, the City-Owned Land Inventory dataset does not link to the blog post notifying users that certain entries were removed after August 2019.¹⁴ In contrast, the description of the Building Permits dataset on the Portal does link to the blog post explaining significant revisions.¹⁵ In another case, a November 2017 Digital Hub post explains the initial design decisions and differences behind three Civilian Office of Police Accountability datasets.¹⁶ However, none of these datasets on the Portal include links to this useful information on the Digital Hub.

Without a consistent method of communicating dataset design decisions and modifications, data users may continue using a dataset without realizing that the underlying structure has changed. This increases the risk that users will reach incorrect conclusions, thereby undermining the City's efforts to improve transparency. The inconsistency of notification and the use of separate websites for communication may also impede users' attempts to understand public data. Prior to the departmental merger, DoIT management stated that it was working toward streamlining the Digital Hub and the Developers website, which may include phasing out one or both sites.

C. MULTIPLE CITY WEBSITES CONTAIN INCONSISTENT SETS OF THE SAME DATA.

Over the course of its work, OIG has analyzed the same or similar data posted on multiple City websites and discovered inconsistencies in that data. Although providing access to data in more than one place is not prohibited, the City's use of multiple public platforms to convey what purports to be the same data undermines transparency by making it harder for users to identify an authoritative or appropriate data source. For example,

¹⁴ Chicago Data Portal, "City-Owned Land Inventory," accessed January 2, 2020, <https://data.cityofchicago.org/Community-Economic-Development/City-Owned-Land-Inventory/aksk-kvfp>. The Developer website's blog post is available at <http://dev.cityofchicago.org/open%20data/data%20portal/2019/08/23/city-owned-land-revision.html>.

¹⁵ Chicago Data Portal, "Building Permits," accessed January 2, 2020, <https://data.cityofchicago.org/Buildings/Building-Permits/ydr8-5enu>.

¹⁶ Chicago Digital Hub, "COPA Adds More Transparency with Police Complaint Data," November 1, 2017, accessed January 2, 2020, <https://digital.cityofchicago.org/index.php/copa-adds-more-transparency-with-police-complaint-data/>.

- Building permit and violation data is available on at least three City webpages.¹⁷ During a 2018 audit, OIG found the Building Violations dataset on the Portal and separate DOB Warehouse contained conflicting information, did not provide a property's full violation history, were not user-friendly, and exacerbated confusion regarding building violation information. Although DOB explained that these platforms were designed to meet different needs, the pages themselves did not describe their intended uses.¹⁸
- Lobbyist registries are available on both the Data Portal and the Board of Ethics Electronic Lobbyist Filing (ELF) website. During a 2016 audit, OIG found that filing dates differed between the Data Portal and ELF.¹⁹ In 2016, DoIT published an article on the Digital Hub detailing its choice to present certain information on each site.²⁰ However, while the ELF site currently links directly to the Digital Hub, the individual datasets on the Portal do not.
- Chicago Department of Transportation service request data historically appeared in two locations. In a 2014 audit, OIG determined that one (the City's since-discontinued Service Delivery Metrics webpage) was inaccurate, while the same information on the Portal was accurate but incomplete.²¹ OIG learned that DoIT had decided to only include the most frequent service requests on the Portal, but had not publicly notified users of this decision.
- Tax Increment Financing data is split across at least five City web pages, only some of which have been recently updated.
- Different historical versions of the Chicago Police Department's (CPD) "party to violence" risk models appear on the Data Portal and CPD's website. The static datasets are based on different models and therefore contain different data; however, neither of them notes these differences or explains why the data was made available on separate platforms. CPD has since discontinued use of both models.²²

¹⁷ The Building Permits and Building Violations datasets appear on the Portal, while DOB maintains a separate website (previously known as the "Warehouse") on their City webpage.

¹⁸ Department of Buildings Complaint-Based Inspections Audit (#16-0301), published April 10, 2018, <https://igchicago.org/2018/04/10/dob-complaint-based-inspections-audit/>.

¹⁹ Board of Ethics Lobbyist Registration Audit (#14-0328), published March 17, 2016, <https://igchicago.org/2016/03/17/audit-of-board-of-ethics-lobbyist-registration/>.

²⁰ Department of Innovation and Technology, "Improved Lobbyist Data," February 2, 2016, accessed January 2, 2020, <https://digital.cityofchicago.org/index.php/improved-lobbyist-data/>.

²¹ Chicago Department of Transportation Service Request Performance Reporting Audit (#13-0006), published January 4, 2014, <https://igchicago.org/2014/01/04/departement-of-transportation-service-request-performance-reporting-audit/>.

²² The Data Portal contains a static version of CPD's Strategic Subjects List, while CPD's website contains a more recent version known as the Crime and Victimization Risk Model. However, CPD decommissioned this model on November 1, 2019.

The use of multiple web platforms for City data may be warranted and appropriate. However, the City's open data policy will be more effective if each platform has a clear purpose, and if the CIO, CDO, and AIS collaborate with all City agencies to ensure a cohesive Citywide approach to open data. Clear public documentation also enables users to better understand and use the Data Portal and other public data resources.

IV. OIG SUGGESTIONS

The CIO and CDO have been tasked with developing a "vision for the use of information technology in City services and infrastructure" to help the administration "spearhead a renewed open data strategy." OIG recommends that, in those efforts, the CIO and CDO work with AIS to comply with the current executive order and strengthen the City's existing open data processes.

To comply with the current executive order, the City should,

- ensure that all City agencies designate an open data coordinator;
- resume meetings of the Open Data Advisory Group, this time including representatives from all City agencies; and
- resume publication of its required annual Open Data Compliance Report to provide transparency and accountability in its open data efforts.

The City should also standardize and streamline its open data initiatives. Specifically, the City should:

- **Standardize and formalize the dataset planning process.** This may include developing a template with standardized questions for AIS and agencies to consider when identifying new data to be made available on the Portal.
- **Standardize documentation for query design decisions** in order to ensure this knowledge is retained for current and future stakeholders.
- **Streamline methods for communicating dataset changes.** For example, AIS should link datasets on the Portal directly to blog posts and/or other historical sources (such as the Digital Hub) that explain the design of that dataset and any subsequent changes.
- **Review existing City data websites and identify opportunities to consolidate them, where possible.** When there is a need to publish similar data in multiple locations, consider updating the dataset descriptions to clearly communicate the differences between the sources and how they should be used.

- **Conduct a staffing analysis.** With only a single staff member assigned to the Data Portal, the City should consider the workload of this employee and determine if more staff are necessary to meet the City's open data goals.
- **Continue to automate data portal processes, when possible.** Such automation should continue to include technological notifications of failed processes, incomplete data transfers, and other situations that cause data quality issues.

While these are not the only open data processes that would benefit from standardization across the City, they provide an opportunity for the City to introduce more consistency in its open data initiatives.

OIG invites the City to respond in writing before March 17, 2020. Any such response will be made public together with this OIG Advisory.

Respectfully,



Joseph M. Ferguson
Inspector General
City of Chicago

cc: Carleton Nolan, Chief Technology Officer, AIS

MISSION

The City of Chicago Office of Inspector General (OIG) is an independent, nonpartisan oversight agency whose mission is to promote economy, efficiency, effectiveness, and integrity in the administration of programs and operations of City government. OIG achieves this mission through,

- administrative and criminal investigations by its Investigations Section;
- performance audits of City programs and operations by its Audit and Program Review Section;
- inspections, evaluations and reviews of City police and police accountability programs, operations, and policies by its Public Safety Section; and
- compliance audit and monitoring of City hiring and employment activities by its Hiring Oversight Unit.

From these activities, OIG issues reports of findings and disciplinary and other recommendations,

- to assure that City officials, employees, and vendors are held accountable for violations of laws and policies;
- to improve the efficiency and cost-effectiveness of government operations; and
- to prevent, detect, identify, expose, and eliminate waste, inefficiency, misconduct, fraud, corruption, and abuse of public authority and resources.

AUTHORITY

OIG's authority to produce reports of its findings and recommendations is established in the City of Chicago Municipal Code §§ 2-56-030(d), -035(c), -110, -230, and 240.

Cover image courtesy of Department of Assets, Information, and Services.

PROJECT TEAM

NICK MCFADDEN, PERFORMANCE ANALYST
CLAIRE KRUSCHKE, SENIOR PERFORMANCE ANALYST
DARWYN JONES, CHIEF PERFORMANCE ANALYST
LISE VALENTINE, DEPUTY INSPECTOR GENERAL

PUBLIC INQUIRIES:

NATALIE A. KURIATA: (773) 478-8417 | NKURIATA@IGCHICAGO.ORG

TO SUGGEST WAYS TO IMPROVE CITY GOVERNMENT, VISIT:
IGCHICAGO.ORG/CONTACT-US/HELP-IMPROVE-CITY-GOVERNMENT

TO REPORT FRAUD, WASTE, AND ABUSE IN CITY PROGRAMS:
CALL OIG'S TOLL-FREE TIP LINE (866) 448-4754 / TTY: (773) 478-2066

OR VISIT OUR WEBSITE:

IGCHICAGO.ORG/CONTACT-US/REPORT-FRAUD-WASTE-ABUSE/