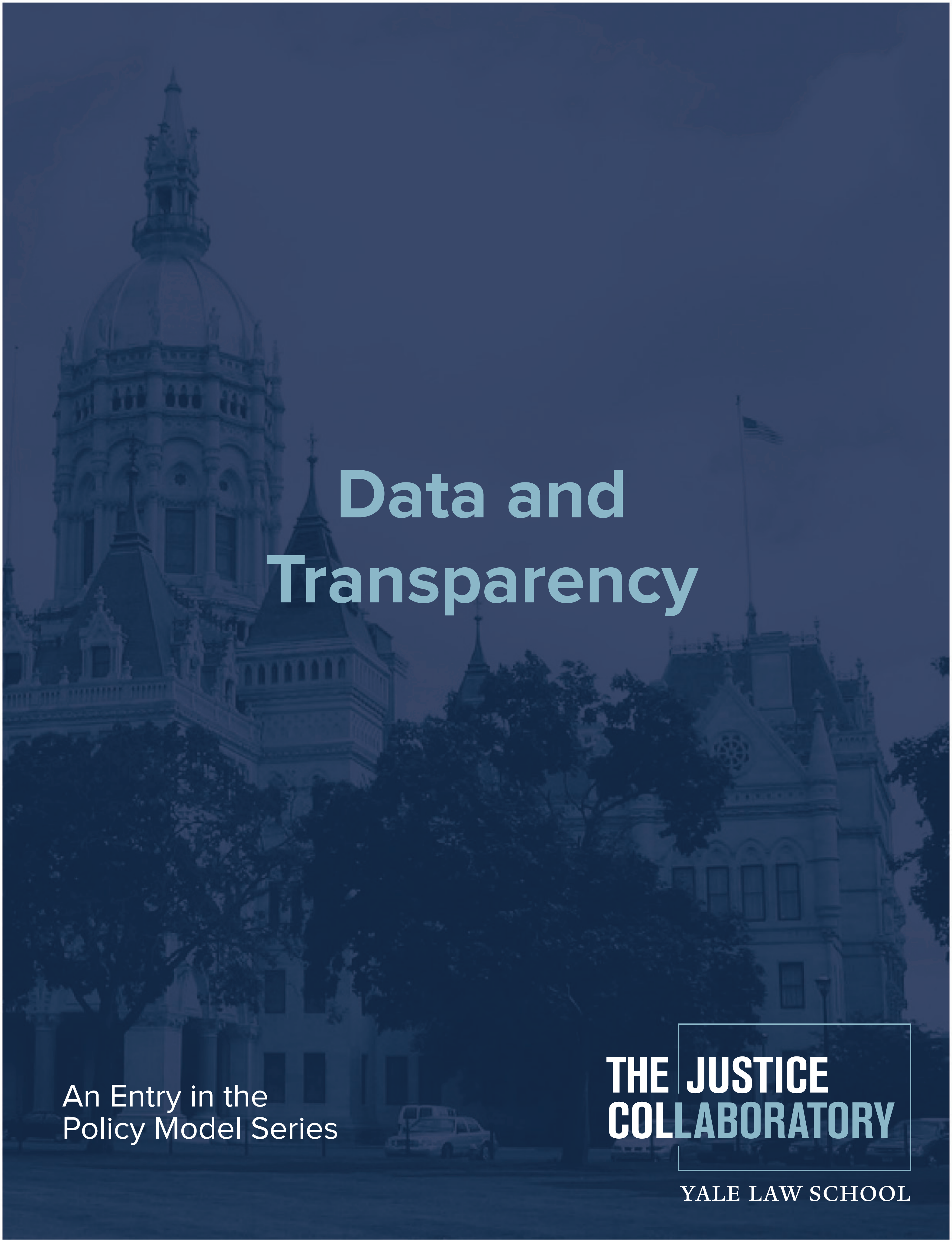

Data and Transparency

An Entry in the
Policy Model Series

THE JUSTICE COLLABORATORY

YALE LAW SCHOOL

About the Justice Collaboratory

The Justice Collaboratory at Yale Law School is a group of nationally recognized academics, researchers, and social scientists who have joined together to build a more just, effective, and democratic criminal legal system by advancing public policies that are scientifically proven to build strong and safe communities where all citizens can thrive.

About the Policy Model Series

The Justice Collaboratory's Policy Model Series offers concrete proposals to those striving to achieve a community-centered justice system, one focused on promoting vibrancy over mere criminal control. Entries in the series are concise, plainly worded, and reflect the latest thinking by leading experts. Our models are intended to serve as templates for state and local laws, though their substance may also be incorporated into agency policies, regulations, and guidance.

For more information about the Justice Collaboratory and its work, please visit us at

justicehappenshere.yale.edu

twitter.com/JCollaboratory

linkedin.com/company/yale-justice-collaboratory

Preface

Municipal police departments, particularly those serving metropolitan areas, are sophisticated bureaucracies that generate a wealth of data and information. However, despite the notable efforts of some departments to promote transparency and access to the data they maintain, most policing data remains inaccessible and, as a result, underutilized. Luckily, the increased adoption of digital record management systems by police departments has drastically improved their ability to not only gather rich repositories of policing data, but also to compile and release that data into useful datasets that hold the potential to yield important public safety insights.

Open data and transparency are key

The benefits of releasing open data are myriad. As noted by the **National Policing Institute**, open police data "[allows] members of the public, community groups, and law enforcement agencies to independently and collaboratively analyze the data to identify potential problems, improve understanding of the challenges faced by law enforcement and their responses, craft solutions, and improve their communities." Through the release of open data, police departments would simultaneously improve public transparency and permit the crowdsourcing of departmental data, avoiding the substantial cost of performing such analysis only in-house.

Our model legislation offers an economical and implementable pathway to promoting policing data transparency

The model's goals are threefold:

- to compel more police departments to make their data openly available,
- to ensure that police data is readily usable by independent researchers, and
- to do so economically.

By requiring only the disclosure of digital data actually collected by departments, the model avoids imposing new, costly mandates while organically requiring additional disclosures as a department's digital recordkeeping practices mature.

This model adds to the efforts of other organizations who have previously offered models and guidance on policing data by specifically targeting the issues of implementation and affordability, two persistent impediments to achieving greater transparency. Whereas organizations like the **NYU Policing Project** and the **Center for Policing Equity** offer invaluable roadmaps for improving data collection practices and transparency in critical areas of interest in policing, the Justice Collaboratory's model aims to trigger a broad release of policing data as currently collected by police departments and as already permitted by their current resources. By leveraging existing departmental capabilities to export digital data into open-source formats—a feature of virtually all digital record management systems—this approach can achieve immediate transparency at minimal cost. States can support this effort by funding increased adoption of digital recordkeeping, the cost of which typically becomes more economical each year.

The interests of both privacy and transparency are balanced

The model is structured to maximize data access while also protecting vital interests in personal privacy and departmental operations. To balance transparency and effective law enforcement, the model excludes from disclosure any personally identifiable information relating to specific departmental employees or persons associated with departmental matters, like witnesses, victims, and suspects. It also excludes from disclosure any information related to an open investigation that, if disclosed, could undermine the investigation.

To facilitate research, the model requires the inclusion of unique identifiers in any released data set to permit separate data sets to be linked to one another, allowing users to track trends relating to specific locations, dates, times, and individual—yet anonymized—officers. Finally, to ensure that current transparency laws are not inadvertently superseded, the model includes a provision expressly stating that it does not alter the scope or applicability of any other law requiring the disclosure of data or records maintained by police departments. The goal here is to expand on current transparency efforts, not to supplant them.

Compliance is required and will be monitored

The model's mandate to comply with its disclosure requirements is clear from the outset. However, to ease the task of implementation, the manner of compliance is largely left to police departments themselves, provided they successfully satisfy the law's requirements concerning the content, composition, and timing of released data. Accordingly, departments can decide how they compile data sets, including exportation from their current record management systems or integration with separate data reporting systems, and how they structure their data sets, including whether to group their disclosures by timespan, subject matter, location, operational unit, or some combination of these or other factors. Once initial compliance practices are implemented, continued compliance with the law should get progressively easier, with the potential for partial or full automation of the disclosure process.

To facilitate the process of monitoring compliance across all covered police departments statewide, the model requires each department to submit to the state attorney general a link to the page on the department's website from which its disclosed datasets may be accessed, including the date of the most recent disclosure. This information must then be published on the attorney general's website, with an indication as to whether the department is in compliance with the timeliness and recency requirements for disclosed datasets. The aim is to provide a statewide snapshot of compliance with the model's disclosure requirements, making it easy to identify those departments that are honoring their transparency obligations and those that are not.

The Model

Section I Duty to Disclose

Every police department must publicly disclose its data as set forth in this law.

Section II Definition, Scope, and Composition of Data

a.) Definition. The term "data" refers to any information collected and maintained by a police department in the ordinary course of its business that is stored in a digital record management system.

b.) Scope.

(i) Data required to be disclosed under this law includes, but is not limited to, the following information:

(A) Incident reporting, including information submitted to the federal National Incident-based Reporting System, the Uniform Crime Reporting system, the Law Enforcement Management and Administrative Statistics survey, or similar systems maintained by the federal [and state] government[s];

(B) Contacts between police officers and members of the public, regardless of whether the contact resulted in arrest and regardless of how contact was initiated;

(C) Calls for service received by the department, including calls received through emergency and non-emergency systems such as 911, 311, 211, 988, or police dispatch services;

(D) Criminal and non-criminal complaints submitted to the department;

(E) Departmental personnel, including employee demographics, qualifications, shifts, deployment, departmental assignments, length of employment, rank, seniority, and salary;

(F) Finances, including funding, budget, and expenditure;

(G) Officer discipline, including information on complaints and outcomes of disciplinary proceedings;

(H) Locational and timestamp information maintained by the department including any geocodes, geotags, coordinates, or other markers indicating the location of officers, incidents, or responses at particular times; and

(I) Datasets compiled by the department for release to any third-party, where public disclosure of the dataset is not otherwise prohibited by law.

(ii) Data to be disclosed under this law does not include:

(A) Information relating to active law enforcement investigations that, if disclosed, would impede the investigation;

(B) Personally identifiable information, including names and social security numbers, relating to departmental employees or persons identified as witnesses, victims, suspects, arrestees, or any other individual associated with a departmental investigation; or

(C) Any information whose disclosure is prohibited by law.

(ii) This law does not alter the scope or applicability of any other law requiring the disclosure of data or records maintained by police departments.

c.) Composition. Data to be disclosed under this law must, prior to disclosure, be compiled as described in this subdivision.

(i) Machine-readable. Data must be structured in a non-proprietary digital format that can be processed by a computer without human intervention. The data must be capable of being processed without loss of any semantic meaning during its processing.

(ii) Disaggregation. Data must be disaggregated to the most granular level maintained by the department, as permitted by law.

(iii) Unique identifiers. Where a data field in a data set contains identifiable information whose disclosure is prohibited by law, departments must replace such information with a unique and anonymized identifier that allows for cross-comparison with other disclosed data sets. To the extent practicable, identifiers must be consistently maintained across data sets to permit longitudinal analysis.

Section III Manner of Disclosure

a.) Timing and Recency. Police departments must disclose all required data within [x] days after this law is enacted, with subsequent disclosures or updates to disclosed data sets occurring no later than every [y] days after the initial disclosure. Disclosed data sets must be current as of no more than [z] days prior to their date of disclosure.

b.) Availability. Data must be made publicly available for download on the department's website and at no additional cost to the public. Data may be made available as a single data set or through multiple data sets, provided that the data sets, either individually or collectively, satisfy the scope and composition requirements of Section II.

c.) Data correction. When a police department corrects or modifies data maintained in its digital record management system, such corrections or modifications must be included in any corresponding publicly available data set, with an indication on the department's website that the data set has been corrected or modified. Corrected or modified data sets shall be published within [x] days after its corresponding data has been corrected or modified within the department's digital record management system.

d.) Codebook. In addition to disclosing data as required under this law, a police department must publish on its website a codebook that describes the content and structure of disclosed data and that defines the meaning of each disclosed data field. Changes to departmental data coding practices shall be timely reflected in its published codebook.

Section IV Manner of Disclosure

a.) Disclosure. Every police department must, within [x] day[s] s of every disclosure it makes pursuant to this law, submit to the state attorney general a link to the page on the department's website from which the disclosed data set may be accessed. The submission must also indicate the date of the dataset's disclosure and the date as of which the disclosure is current.

b.) Publication. The attorney general must publish on its website a page containing the links submitted to it by each police department pursuant to subdivision (a) and a note indicating the data set's date of disclosure and the date as of which it is current. The attorney general must also indicate whether each department is in compliance with the timing and recency requirements of subdivision (a) of section III.