Data Management Policy

Committee on Information Technology

PURPOSE AND SCOPE

Data is a key asset in meeting the demands of a ~~21st century~~modern government. Proper data management ~~can add value to the work of the City,~~-- including~~:~~

- ~~Improvements in~~ management of systems that collect, store, transform, and share data ~~consistency and quality~~
- ~~Faster, easier access to data~~
- ~~Better controls~~-- is necessary for a government to operate efficiently and ~~security~~deliver services, especially when operations
- ~~**Data** sharing~~ and ~~interoperability between datasets~~
- ~~Integrated data across~~services span multiple departments. The benefits of proper data management include:

- ~~Data analytics~~Consistent, discoverable, interoperable, secure, and ~~more~~ trustworthy data
- Seamless data integration between departments
- Rapid deployment of new scalable services or enhancements to existing ones
- Shorter time-to-insight with access advanced data science tools
- Enhanced risk mitigation and compliance monitoring
- Reduction in cost

To deliver ~~better~~these outcomes, data must be proactively managed and maintained ~~much like our capital and financial assets.~~

throughout its lifecycle. This policy applies to all information resources operated by the City and County of San Francisco ~~and~~as well as its departments~~,~~ and commissions~~. Elected~~, elected officials, employees, consultants, city-funded nonprofits, and vendors working on behalf of the City and County of San Francisco. The above listed entities are required to comply with this policy.

POLICY STATEMENT

This policy establishes a framework for the proper management of data as an asset across the City. ~~-~~Departments must adopt this framework ~~and the requirements below to support the ongoing, proactive management of data~~, which includes:

~~COIT Policy Dates~~

~~Approved: January 17, 2019~~
~~Next Anticipated Policy Update: June 2020~~

- ~~The identification~~Ongoing modernization of data systems

1. ~~Inventory~~ and classification of data ~~in **database**~~systems, and ~~**dataset inventories**~~

~~2.~~• ~~The processes and policies for appropriately sharing **open and confidential**~~the data within them
- ~~An approach to identify~~ Assessment of interdepartmental data ~~**and interdepartmental**~~uses
- Addressing security and privacy considerations

~~3.~~• Making data ~~**standards** and actively manage those data, databases, and standards~~accessible for public use

~~Together, these represent the requirements for managing data as an asset. Increasing Citywide understanding of our data assets and the purposes they serve will allow the City to proactively plan for and meet the needs of an ever-changing city.~~

POLICY REQUIREMENTS

MODERN DATA SYSTEMS

A data system, also referred to as an Enterprise System of Record as defined in California Government Code §7922.630, is the foundation of good data management. The City should prioritize modernizing its data systems whenever feasible. While this transformation won't happen overnight, departments should seek systems with the following capabilities when procuring new tools. These requirements should be added to contracts when possible:

Data ~~**1.0 Database and Dataset Inventories**~~

- **Integration:** Systems should support secure read and write access to external sources through well-documented integrations (including but not limited to direct systems access or APIs).
- **Automation & Orchestration:** To reduce manual work performed by City staff, systems should make it easy to schedule and automate repeated tasks, such as sending notifications, updating records, or moving data between systems, or integrate with orchestration tools.
- **Repository & Version Control:** Systems should support integration with code repositories to allow for version control and collaborative development.
- **Collaboration:** Systems should promote team collaboration. Code, documentation, configurations, templates, scripts, and libraries should not be stored locally on individual machines but instead be maintained in shared, version-controlled environments.
- **Continuous Integration and Continuous Delivery (CI/CD):**Systems should support automated testing, validation, and deployment pipelines to enable rapid, reliable, and secure implementation of system enhancements, integrations, and updates.
- **Cloud-native Services:**Systems should utilize cloud-native and public cloud infrastructure to enhance scalability, interoperability, and efficiency, while reducing administrative complexity compared to legacy on-premises solutions.

Modern infrastructure lets cities choose the best tool for each part of the data lifecycle, whether for collection, storage, cleaning and transformation, orchestration, sharing, or visualization, without being locked into a single vendor or platform.

The City's standard contracts, such as section the P-64X series, should be updated to more closely meet the standards listed above.

DATA SYSTEM AND DATASET INVENTORIES

Managing data as an asset starts with knowing the range and type of data under the control of the City~~, which supports:~~

- ~~improved cyber and information security,~~
- ~~improved understanding of the criticality of data access to support the Citywide DPR3 policy~~
- ~~proactive publication of data as appropriate, and enhanced use of shared data through the identification of data critical to improved operations~~

~~San Francisco Administrative Code~~ . This process enables greater cyber- and information-security and enhances efficiency and transparency by identifying candidates for Interdepartmental sharing and Open Data. San Francisco Administrative Code Chapter 22D establishes the requirement to publish "a catalogue of the Department's data that can be made public~~, including both raw data sets and application programming interfaces (API's)""~~. The inventory is not a requirement to publish data, only to catalog all existing systems and datasets owned by a department. The following requirements build on Chapter 22D and strengthen the foundation for good data management:

~~The following requirements build on Chapter 22D and form a foundation for good data management:~~

~~Collection of Database~~

1.~~•~~ **Data Systems and Dataset Inventory**. Departments must create and maintain ~~on an annual basis a Database~~a Data System Inventory and Dataset Inventory[1] per ~~procedures and templates~~procedures and templates set by the Chief Data Officer (CDO) and with guidance from the City Attorney's Office. This process should happen no less than once a year. See public Dataset and System Inventory.

2.~~•~~ **Classification of ~~Databases~~Data Systems and Datasets**. As part of the inventory, Departments must classify ~~databases~~data systems and datasets per procedures set by the CDO and City Chief Information Security Officer (CCISO) in the Data Classification Standard.~~. Data classification is an indicator for aiding in the proper management and security of data in use, in transit and at rest. It is not meant to place any additional restrictions over and above what is required by law, or administrative policy. The classifications as submitted per the Data Classification Standard can be updated at any time if classifications don't reflect the nature of data. This can happen either through departmental review or in the course of review by the CDO or CCISO during other processes described in this policy.~~

3.~~•~~ **Publication of ~~Database~~Data System and Dataset Inventory**. The CDO must publish the ~~Database~~Data System and Dataset Inventories as datasets on the City's open data portal (or successor site) and ~~publish updated inventories no less than annually by July 1st. The public inventories are inclusive of datasets regardless of classification except where local, State or Federal laws specify otherwise.~~update them whenever departments add, remove, or modify their inventory.

- **Data Retention**. Data retention will adhere to guidance of Chapter 8 of the Administrative Code.

The ~~The Data Coordinator Guidebooks~~Data Inventory Explainer, incorporated ~~by~~for reference, ~~provide~~provides more detailed direction on the ~~database~~data system and data inventory process.

[1] ~~Required by SF Admin. Code 22D.2(c)(4)(A)~~

~~2.0 Open Data~~INTERDEPARTMENTAL DATA

City services are deeply interconnected, but the data systems that support them remain siloed. To improve service delivery and operational efficiency, we must modernize how interdepartmental data is classified, stored, cleaned, discovered, integrated, analyzed, and shared.

Increasing Interdepartmental Data Availability

All data within the City and County of San Francisco should be assessed for Interdepartmental use cases. Data often has a network effect – the more who use it, the more valuable it becomes. It is often easy to evaluate the risks of sharing data, but the risk of sharing data should not be assessed in a vacuum. It must be balanced against the value of increased access, usage and standardization.

This policy establishes a three-part check to verify that data can be shared with another City entity. Specifically, a department should not withhold data from another city department as long as three requirements are met:

1. There is a valid business reason to request data, such as supporting public services, internal operations, regulatory compliance, reporting, analysis, or decision-making
2. Regulations allow the requesting department to access the data
3. The requesting department has the systems and processes in place to comply with data regulations

As outlined in "Modern Data Systems", to facilitate the sharing of Interdepartmental data, all departmental and vendor data systems should support data sharing through direct system access, APIs or other methods that enable easier access and automation.

Interdepartmental Data Standards

To realize the full value of Interdepartmental data, a shared data standard as outlined in this section below must be applied to ensure proper usage, efficiency, and interoperability. Without these standards, the risk of incorrect data usage, inefficiency when collaborating across departments, and inconsistent reporting grows. Specifically:

- Metadata Standards: Applying COIT Metadata Standards to all Interdepartmental data, ensuring all datasets are properly documented.
- Field & Column-Specific Standards: Applying best practices outlined in DataSF's Data Standards Guide, including formatting for addresses, street segments, parcels, business ID, and other common identifiers. This list is meant to grow as more standards are established.

Custodianship and Ownership of Interdepartmental Data

Interdepartmental data stored on Department of Technology (DT) infrastructure is covered by the Data Custodian and Stewardship Policy. When data is replicated from a system of record onto DT systems, ownership remains with the originating department. All Sunshine requests must be directed to the department that owns the enterprise system of record where the data originated. In all other cases, ownership and custodianship should be defined in the data-sharing MOU (see "sharing non-public data" below).

SHARING NON-PUBLIC DATA

When managing and sharing non-public data, departments must follow best practices, listed below, to protect

privacy and ensure the security of City data, datasets, replica datasets, derived datasets, and systems. Maintaining public trust requires strong safeguards and full compliance with applicable privacy laws. Secure and lawful data sharing not only reduces the risk of breaches but also improves efficiency by limiting duplicate data collection. To minimize risk:

- **Follow Cybersecurity Best Practices.** Abide by the guidance in the City and County of San Francisco Cybersecurity Requirements, and the guidance of the Office of the Chief Information Security Officer (CCISO), and departmental CISOs.
- **Assess Privacy and Security Risks.** When sharing data, departments can utilize the Privacy Toolkit and Security Toolkit established by the CISO and CDO. These resources should be applied alongside any department- or domain-specific regulations and laws, with additional guidance from the City Attorney's Office as needed.
- **Establish Data Sharing Agreements.** Standardized Memoranda of Understanding (MOUs) for data sharing are available from the City Attorney's Office and should be used when appropriate.
- **Utilize Secure Systems.** Sensitive data should only be shared using secure, approved systems that meet the City's privacy and security requirements. Departments should avoid using email or unencrypted file transfers for sensitive information and instead rely on tools designed for secure data exchange, such as encrypted portals or file-sharing platforms vetted by the Department of Technology or the CISO.
- **Avoid Dataset Replicas.** Whenever possible, departments should provide secure access to data in their systems or a central system rather than generating static copies or replications for dispersal.

OPEN DATA

Data yields the most value when it

~~The greatest value from City data is realized when anyone~~ is free to access, use and share ~~datasets~~ consistent with safety, privacy, and security~~. "Open Data" means that the dataset is:~~

- ~~Available as a whole to all at no cost and discoverable and accessible over the internet,~~
- ~~Published to minimize time between the creation and dissemination of the data,~~
- considerations. ~~Documented,~~
- ~~Provided under terms that permit re-use, redistribution, and the mixing with other datasets, and~~
- ~~Provided in an open format that is machine-readable.~~

Making open data available means it is published on the City's ~~open data portal~~Open Data Portal (https://data.sfgov.org/ or successor site~~) consistent with this definition.~~).

~~2.1~~ Publishing Process and Prioritization ~~and Plans~~

~~Departments~~Once departments have identified potential Open Data candidates though the annual inventory process, they must ~~prioritize data for publication and develop~~:

Create or update their publishing ~~plans as follows[2]:~~

1. ~~•~~ **~~Publishing Prioritization.~~plan**. Departments must prioritize datasets for publication ~~per procedures set~~using the Prioritization Guidance created by the CDO ~~and no less than annually as part of the Data Inventory process. In general, those datasets that have the highest public value and can easily be published should be prioritized.~~

~~[2] Implements SF Admin. Code 22D.2(c)(3)~~

2. **Publishing Plans.** Departments must submit a Publishing Plan per procedures set by the CDO and no less than annually. The Publishing Plan will describe the Department's commitments with respect to publishing data on the City's open data portal during the publishing plan timeframe. Datasets will be reviewed by the CCISO for privacy considerations prior to publishing.

The Data Coordinator Guidebooks, incorporated by reference, provide detailed direction on the process for prioritizing data for publication and for developing publishing plans.

2.2 Publishing Data

Departments must publish data per procedures set by the CDO and as follows:

1. **Publishing Portal.** Departments must submit datasets for publication to the Publishing Portal. The Submission Guidelines and Publishing Guidelines, incorporated by reference, provide more detailed direction on the publishing process.
2. **Dataset Documentation (Metadata).** Departments must document data prior to publication per the City's Metadata Standard.
3. **Data Standards.** Departments should publish data if possible, practical and available, per commonly used standards for that type of data.

- **Follow the publishing process**. Departments must follow the Publishing Process established by the CDO, when sharing data on the Open Data Portal

Other Publishing Considerations

4.- **Non-Public Data.** Departments must use the Open Data Release Toolkit Open Data Release Toolkit and work with the CDO to transform non-public data so that a relevant and appropriate view of that data (e.g. limited fields, aggregate data) can be published as Open Data. Non-public data includes anything properly classified as Level 2 or higher per the Data Classification Standard.[3]

5.- **Licensing Standard.** The default license for all Open Data is the Public Domain Dedication License (PDDL).[4] While this license is the default, other licenses may be used for specific datasets and as approved by the CDO.

6.- **Recording.** The CDO must update the publishing status of datasets in the Dataset Inventory when a dataset becomes public.

7.- **Language.** Data sets, including metadata, are not required to be published in additional languages beyond those used at the source. The City may opt to implementedimplement automated language translation of data sets in the future.

3.0 Confidential Data

Data and information sharing increases our ability to improve service outcomes through more accurate evaluation of policy options, improved stewardship of taxpayer dollars, and more coordinated and personalized delivery of public services.

The City and County of San Francisco is committed to inter-agency information and data sharing as a standard practice. At the same time, it is essential to maintain public trust that confidential information is safe and secure via appropriate, strong, and effective safeguards and compliance with applicable privacy rules.

[3] Supports SF Admin. Code 22D.2(c)(4)(G)

[4] License terms available here: http://opendatacommons.org/licenses/pddl/1.0/.

1. **Minimum data sharing, protection standards and privacy.** The City Chief Information Security Officer (CCISO) and CDO must establish minimum data sharing and protection standards using a risk based approach consistent with the Citywide Data Classification Standard and the needs and requirements of interdepartmental data and information sharing.
2. **Data sharing agreements and resources.** The City Attorney's Office, and with input from departments, must provide template agreements, supporting guidance and a standard and timely review process for confidential data and information sharing agreements across departments or with external partners.
3. 1. **Departmental processes.** Departments must protect confidential data per applicable law and regulations. Departments should work to maximize the value of data and data sharing for public good purposes while appropriately managing risk. In this context, Departments should actively pursue interdepartmental data sharing for reasons such as program evaluation, research, analysis, care coordination, operations, or other public good purposes and per citywide standards and resources for data sharing and protection.

4.0 Management of Interdepartmental Data

Proper interdepartmental data management can unlock value and control costs for the City. Where data can serve multiple programs, initiatives or client groups across more than one department, the City should adopt data management practices.

Data management means that data identified as interdepartmental is actively managed to:

1. **Support integrations.** Data should be accessible in a manner that users can develop self-service integrations.
2. **Minimize redundancy and errors.** Data should be easily accessible to minimize the production of derivative data products that conflict with the authoritative source.
3. **Ensure data quality.** Data should be monitored and managed to maintain data quality appropriate to the requirements of its use.
4. **Scale with the needs of users.** Data stewards should solicit, vet, prioritize and implement changes that meet the needs of the users over time and as resources allow.
5. **Properly control access.** Access to data should be managed commensurate with risks.
6. **Be clearly documented.** To promote responsible data use, data must be documented in a tool that promotes proper maintenance and accessibility of documentation.

Done well, properly managed data should bring down the overall costs to the organization by reducing errors and duplication of effort, while also increasing effective use of interdepartmental data.

In pursuit of adopting appropriate data management practices, this policy defines:

1. **Interdepartmental data.** Criteria for interdepartmental data and related requirements.
2. **Interdepartmental data standards.** Criteria for interdepartmental data standards and related requirements.
3. **Implementation.** How the City should ensure interdepartmental data and standards are managed per this policy.

4.1 Interdepartmental Data

Definition / Scope. Interdepartmental Data is data that has substantial value when used across more than one department. Properly managed, interdepartmental data should also unlock efficiencies across

departments. Data can be characterized as interdepartmental data when it meets a three-part definition:

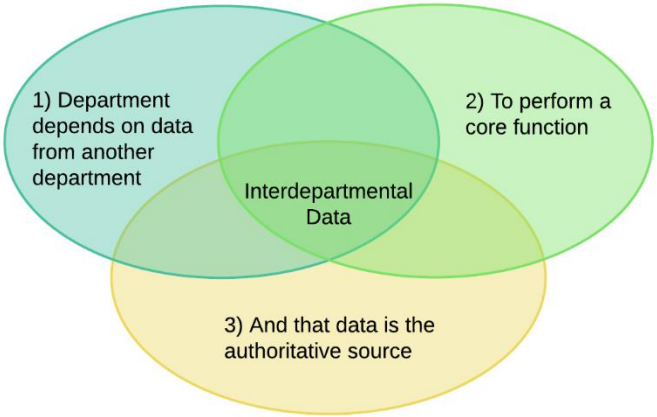
1. **Process dependencies.** At least one department is solely dependent on data from another department to conduct its work.

2. **City and County core functions.** The data supports work that is integral to the operations and management of core functions for the City and County. A core function is a service, program or activity authorized by statute, regulation, the Administrative Code or other authorizing authority. It does not include functions in support of core functions such as research or planning.

3. **Authoritative Source**. The data is the authoritative source. Authoritative sources can be codified in federal, state or local law or through administrative policy or practice.

Per this three-part definition, not all departments will have interdepartmental data under their stewardship.

4.2 Interdepartmental Data standards



Definition / Scope. Interdepartmental Data standards are those standards that are valuable when applied across more than one department and can include:

- lists of permissible values like those used in a lookup or as an identifier or code,
- formats, and
- methods for data collection, sharing and/or reporting

Standards can be applied to common attributes used across the City, including but not limited to parcel identification, demographics, department names and codes, sexual orientation and gender identity.

Standards allow easier integration of data across the City and enable consistent and comparable reporting across City agencies.

Per this criteria, not all departments will have Interdepartmental Data standards under their stewardship, but all departments are expected to use Interdepartmental Data standards to the extent reasonable.

Departments must use data standards as listed in the Data Standards Reference Handbook unless otherwise required.

4.3 Identification of and Requirements for Interdepartmental Data and Standards

Interdepartmental data should be identified, prioritized and aligned strategically over time. To that end the Committee on Information Technology and the CDO encourages and supports cross-department working groups to:

1. Identify Interdepartmental Data and Standards per this policy.

~~2. Rank and prioritize these based on the organizational value they can provide.~~

~~Based on the results on the working group, Departments managing Interdepartmental Data or Standards must work to implement the requirements in the Appendix.~~

~~5.0~~

EVOLUTION OF DATA MANAGEMENT

As technology evolves, so too must the data management policy. This section outlines specific use cases for data management, and will expand as new technologies require:

Artificial Intelligence

The use of Artificial Intelligence (AI) within the City is governed by Chapter 22J of the Administrative Code. In the context of data management, several key principles warrant particular attention:

- **Data Provenance.** The origin, sources, and any transformations of data used in the training and operation of AI and machine learning (ML) models should be thoroughly documented to ensure traceability and accountability.
- **Data Quality.** All data utilized in AI/ML systems should be evaluated for accuracy, completeness, and potential bias.
- **Algorithmic Transparency.** To the extent practicable, the City shall promote transparency in the design, inputs, and functioning of AI/ML algorithms, including the data they rely on and the decisions they generate.
- **Ethical Considerations.** AI/ML applications should be reviewed for potential ethical risks. This includes proactively addressing concerns related to bias, fairness, individual rights, and privacy.

The AI Advisory Board—convened by the Office of Emerging Technologies, in partnership with the Chief Data Officer and the City Attorney's Office—shall be responsible for developing and maintaining guidelines that promote ethical AI/ML data practices. AI Advisory Board shall also establish a framework for assessing and mitigating risks associated with the use of AI/ML technologies.

ROLES AND RESPONSIBILITIES

~~5.1~~ Chief Data Officer~~.~~

The CDO is responsible for the following:

- ~~•~~● Develop and oversee the process for creating, maintaining, and publishing the annual ~~Database~~Data System and Dataset Inventory, including data classification;
- ~~•~~● Develop processes and resources and support the publishing of open and internal data consistent with Citywide policies and standards~~;~~
- ~~•~~● Develop, with the City Attorney and input from departments, template agreements, supporting guidance, and a standard and timely review process for confidential data and information~~ -~~-sharing agreements across departments or with external partners;
- ~~•~~● Develop, with the ~~City Chief Information Security Officer,~~CISO minimum data ~~sharing~~quality and protection standards;
- ~~•~~● Define data architecture modeling standards and tools that ~~will be utilized by~~ departments~~; and~~ will use;
- ~~•~~● Engage with Data Coordinators, Stewards, and Custodians to strategically plan, and provide support and training for the publishing of data~~,~~; and
- ● ~~**5.2**~~ Lead & implement citywide data modernization via consolidation and integration into a unified data platform to achieve the objectives of the data management policy and the city's technology goals (done

in partnership with the CIO and the Department of Technology).

City Chief Information Security Officer.

The ~~CCISO~~CISO is responsible for the following:

- Develop minimum data protection standards;
- Complete and publish the annual Open Data Program Risk Assessment; and
- Advise departments on data management.

~~5.3~~

Data Coordinators.

Department Data Coordinators are responsible for the following~~, and according to Administrative Code Chapter 22D~~:

- ~~Work with the CDO to coordinate implementation of and oversee compliance with the Data Policy within their department;~~
- Coordinate the annual ~~Database~~Data system and Dataset Inventory process, including classification, in their ~~Department~~department and per CDO procedures;
- Coordinate the department prioritization of data for publication ~~and the creation of department plans and timelines for publishing data; and~~;
- Coordinate publication of data on the open data portal per procedures and standards set by the CDO.
- **~~5.4~~** Identify and record Interdepartmental data originating from their department
- Alert CDO to any changes of the coordinator or stewards within their department

Information Security Officer

Department Information Security Officers are responsible for the following:

- Apply the 3-part check to assess Interdepartmental data requests when received
- Make final decision on data access for Interdepartmental data requests

Data Stewards.

Department Data Stewards are responsible for the following:

- Work with the Department Data Coordinator to properly document data for which they are the steward according to the inventory requirements;
- Follow the requirements set out in the Data Classification standard and properly document the classification of data in the inventories; and
- Provide proper documentation of shared datasets to support the responsible use of data and metadata for all data, including unstructured documents, is accurately recorded and maintained.

~~5.5~~

Data Custodians.

Department Data Custodians are responsible for the following:

- Work with the Data Coordinator and Data Stewards to provide appropriate documentation to support the ~~database~~data system inventory;
- Work with Data Stewards and Data Coordinator to support efforts as needed to make data available through the publishing process referenced in this Policy; and
- Adequately support their department's Data Stewards, Data Coordinators, ~~Cybersecurity~~and Information Security Officer in conducting their responsibilities in this Policy and in the Data Classification Standard.

~~5.6~~

Data Users.

Data Users are responsible for the following:

- Responsibly use open data or data obtained from departments by reading or requesting documentation on data and applying analysis understanding any constraints on the data;
- Follow any constraints on use as specified in MOUs or other agreements where applicable;
- Be familiar with federal, state and local confidentiality or privacy laws pertaining to the data they collect, access, use, or maintain in conducting their work.; and
- **5.7** Adhere to data retention schedules and disposition procedures when handling data, and report any potential issues or discrepancies to the Data Steward.

Departments.

Department leadership and program management are responsible for the following:

- Assure that staff handling confidential data are sufficiently trained and aware of their duties with respect to securing and protecting private information including Data Stewards, Data Custodians and Data Users
- Make good faith effort to identifyfollow Interdepartmental data and data standards; and
- For those having Interdepartmental data or data standard candidates, work to meet expectations for management of interdepartmentalInterdepartmental data and data standards as resources allow and by priority of value to the enterprise.

5.8

City Chief Information Officer.

CIO and the Department of Technology are responsible for the following:

- Advise department leadership and program staff in technologies to support and build Interdepartmental data and data standards;
- Publish and promote standards and patterns on how to best implement the requirements for managing Interdepartmental data and data standards; and
- Help implement any Drive central procurement strategies to increase adoption of modern technology systems as described in Section 1.0; and
- Partner and support citywide data modernization, in partnership with the CDO, via consolidation and integrations for managing Interdepartmental data and into a unified data standardsplatform to achieve the objectives of the data management policy and the City's technology goals.

5.9

Controller City Services Auditor Audits.

CSA Audits is responsible for the following:

- Audit this policy for compliance as needed.

REFERENCE

- Cybersecurity Policy
- Data Classification Standard
- DPR3 Policy
- Metadata Standard

APPENDIX: Requirements

Office of Contract Administration

OCA is responsible for the following:

- Explore updating the language in the Software as a Service (SAAS) contract to include stronger language related to Modern Data Systems

METRICS FOR POLICY EFFECTIVENESS

To ensure the ongoing effectiveness of this Data Management Policy and to drive continuous improvement, the Office of the Chief Data Officer will publish the following metrics:

Inventory Metrics

- **Completion Rate:** Percentage of departments that have submitted a complete and up-to-date Data System and Dataset Inventory by the annual deadline.
    - Target: 100% compliance.

Interdepartmental Metrics

- **Participation Rate:** Percent of departments with MOU or other data sharing agreement in place with other department(s).
    - Target: 70%

Open Data~~. Departments managing Interdepartmental~~ **Metrics**

- **Publishing Rate:** Percentage of departments that are publishing data on the Open Data ~~should work~~Portal.
    - Target: 100% compliance.

EXCEPTIONS

This policy does not override any State or federal regulations such as the Health Insurance Portability and Accountability Act ("HIPAA"), Criminal Justice Information Services Policy ("CJIS"), Sunshine Ordinance, Data Retention, etc. Further, this policy does not apply to ~~accomplish the following~~data provided to the City by the State or Federal government, or by other municipal or regional jurisdictions.

DEFINITIONS

Data System

Also referred to as an Enterprise System of Record, as defined in California Government Code §7922.630

Dataset

A dataset is defined as a structured or unstructured collection of digital information that is used in the delivery of public services, enables internal operations, regulatory compliance, reporting, analysis, or decision-making. For the purposes of this inventory, a dataset must be:

- **Official and authoritative.** Recognized as the source of truth for ~~that data:~~its purpose.

1. ~~**Establish appropriate change management and user feedback practices.** Establish feedback processes appropriate to the complexity of the data. These could include one or more of the following:~~
    - ~~Create data user groups with representation~~ **Production-level**. Used in finalized, downstream applications—not temporary, intermediary, or raw data tables used only during processing or development.

This includes structured data (e.g., databases, spreadsheets, digital maps, and CAD files) and select unstructured materials (e.g., internal knowledge bases, SOPs, photos, audio, videos, PDFs, or scanned files).

Dataset Replica

A dataset replica or replications is a copy of a dataset that is functionally the same as a dataset, but does not serve as the official and authoritative source of truth. It is commonly referred to as a: Snapshot, clone, mirror, or backup.

Derived Dataset

A derived dataset is a dataset that is created by combining one or more source datasets or replica datasets to create a new asset. If a derived dataset is Official and Authoritative and Production-level, it can be treated as a new dataset with shared stewardship between the source dataset owners.

Interdepartmental Data~~**from major users of**~~

Interdepartmental Data is a dataset that meets a two-part definition:

~~a.~~• **Process dependencies**. At least one department is dependent on data from another department to ~~elicit ongoing needs~~conduct its work. This includes (but is not limited to) scenarios where it would be costly, redundant, or ~~solicit feedback when planning changes~~inefficient for the dependent to collect the data itself.

• **Authoritative Source**. The data is the authoritative source. Authoritative sources can be codified in federal, state or local law or through administrative policy or practice.

~~b. Provide opt-in communications with users of data through a listserv or similar broadcast channel.~~

~~c. Develop processes to report errors and reconcile those errors within a published target service level.~~

~~2. **Provide clear, accessible documentation.** At a minimum, document data per the Data Standard. Interdepartmental Data stewards are encouraged to supplement the minimum standard with additional information and explanations, including wikis or similar collaborative documentation.~~

~~**Provide means for accessing the data as appropriate**~~**Nonpublic Data**

Any dataset not classified as Level 1 (Public) using the COIT data classification standards.

Open Data

Open Data means that the dataset or replica dataset is:

• Free to download, use, and share

• Discoverable and accessible to all over the internet

~~3.~~ Published to ~~**City Departments.**~~

~~a. Develop an access control policy to determine who may be granted access to~~minimize time between the ~~data based on a risk assessment~~creation and ~~consistent with appropriate privacy protections.~~

~~b.~~• ~~Develop and manage a standard, clear and visible process for granting access to~~ dissemination of the data ~~and consistent with the access control policy. Minimally make this process available through a discoverable URL.~~

~~c. Make the data available on any citywide data sharing platforms, including the open data portal if the data can be made publicly available and consistent with the access control policy.~~

d. Provide the data as a service preferably through a documented Application Programming Interface or other web service.

4. **Monitor and report on data.**

a. Establish a data quality monitoring program and report on data quality metrics. The Data Quality resource collection provides guidance on establishing a data quality monitoring program.
b. Establish and monitor service availability metrics and report on these in a consistent, preferably automated manner.

Requirements for Interdepartmental Data Standards. The following requirements apply to Interdepartmental Data standards:

- Stewards for Interdepartmental DataDocumented

1. Provided under terms that permit re-use, redistribution, and the mixing with other datasets standards must make the standard available as a service preferably through a documented API. Publishing the data on the open data portal meets this requirement.
2. Interdepartmental Data standard stewards are responsible for the management and quality control of their respective data standards.

- Interdepartmental Data standard stewards must set clear, visible feedback mechanisms for reporting errors or issues with the data standard.

3.- Provided in an open format that is machine-readable and accessible via API