

San José Digital Privacy Manual

Maintained and updated by:
San José's Digital Privacy Office
digitalprivacy@sanjoseca.gov

Last updated:
August 11th, 2023

PURPOSE

This document provides guidance on matters related to Digital Privacy for the City of San Jose (commonly referred to as the “City”). It serves as the foundational policy document for digital privacy decisions, enabling a more predictable review process based on previous review. It also serves as the general guide for collecting, managing, sharing, processing, or otherwise using personal information (defined in “Personal Information” section below) for the City.

Refer to this document for the following purposes:

1. To understand how to initiate a Privacy review (required for all new technology procurements and data initiatives) see “Digital Privacy Review”
2. To review the current Digital Privacy Policy, which is the foundational document for the City of San José’s approach to privacy, see “Digital Privacy Policy”
3. To understand what “Personal Information” is—i.e., information that presents a Privacy Risk and therefore will require additional review by the Digital Privacy Office—see “Understanding Personal Information”
4. To find additional guidance on collecting data, anonymizing data, and handling sensitive data, see “Data Collection Guidance”, “Guidance for Anonymizing Data”, “Survey Guidance”, and “Special Considerations for Sensitive Data” respectively
5. To find an introduction to digital privacy, refer to “San José’s introduction to Digital Privacy”

This document will continue to be updated to provide the latest information on the City’s Digital Privacy Policy and practices.

TABLE OF CONTENTS

Purpose.....	2
Table of Contents.....	3
Digital Privacy Review.....	4
Background.....	4
Standardized Privacy review protocol.....	4
Initiating privacy review	6
Threshold Analysis – Identifying a Project’s Privacy Risk.....	7
Privacy review process	10
Data Usage Protocols and ongoing monitoring.....	13
Digital Privacy Policy	14
Context.....	14
Policy.....	14
Understanding Personal Information.....	15
Data Collection Guidance – Minimizing Data	17
Collection practices - Video recording and images.....	18
Collection practices – Audio recordings.....	19
Collection practices – Other sensors / activity tracking.....	19
Collection practices – Direct data submissions.....	20
Guidance for providing notices.....	21
Guidance for Anonymizing data	22
Data anonymization detailed through a brief example	22
Survey Guidance	26
Special considerations for sensitive data	27
Health information.....	27
Education information.....	29
Financial Information	30
San José’s introductory Training to Digital Privacy	30
Privacy Fundamentals	31
Protecting Privacy	37
City Staff and Privacy	40
Appendix B – More exhaustive list of Personally Identifiable Information (PII).....	43

DIGITAL PRIVACY REVIEW

Required for all technology and data initiatives

Background

The Mayor and City Council approved the City of San José’s Digital Privacy Policy on December 8th, 2020 (see “Digital Privacy Policy” section below). In October 2021, the City founded its Digital Privacy Office (DPO) to implement the Digital Privacy Policy and protect community members in how the City collects, processes, and uses information.

The DPO has coordinated with City departments and peer governments in developing and implementing a formal process or “protocol” for Privacy reviews, DPO approvals, and defining data usage. **A DPO review is required for all new technology and data initiatives. Existing projects that use personal information** (see “Personal Information” section below for a full definition) **may also require a privacy review if they have not been reviewed by the DPO previously.**

For more background on the privacy review process, see the Digital Privacy Program Status Report presented to the Smart Cities and Service Improvements Committee on February 3rd, 2022 (item CC 22-011).¹

Standardized Privacy review protocol

The DPO provides guidance on developing effective and responsible PII usage,² building public and employee trust in new PII usage, and complying with our City’s Privacy Principles and (in consultation with legal) relevant Privacy Law.

The Digital Privacy Office outlines seven possible steps to the City’s privacy protocol, which include review, defining data usage, and monitoring. While not all steps are necessary for every project, the protocol provides the general framework for any project.

1. **Data Initiative Proposal**: Departments first engage the Digital Privacy Office (DPO) to discuss their project. Departments should provide existing material related to the project, including details on project purpose and data collected.
2. **Initial Privacy Assessment**: DPO conducts a “threshold analysis” to determine if project requires additional review, a Data Usage Protocol (DUP), and public engagement. Low-risk projects are approved at this step. Medium-risk projects can often be handled through an abbreviated version of the steps.
3. **Data Usage Protocol (DUP)**: The DPO determines if the proposed project is covered under an existing DUP (e.g., DOT plans to install a traffic monitoring camera when a general usage policy for traffic monitoring cameras has already been approved). If

¹ <https://sanjose.legistar.com/LegislationDetail.aspx?ID=5381557&GUID=6AF73AE0-79E7-453E-A365-B9D984474BBB>

² “PII usage” includes any action—data collection, sharing, storage, usage, etc.—that involves Personally Identifiable Information or “PII” (see definition in “Personal Information” section)

so, the project can generally move to step 5, otherwise the departments and DPO draft a DUP.

4. **Public Engagement:** The draft DUP (explained in next section) is published online³ for public comment. If the project is considered high-risk and applies City-wide, the DPO conducts in-person outreach targeting communities with limited access to online comments (either due to language or internet access issues). Community feedback is then incorporated into the DUP.
5. **Privacy Review:** The department(s) fills out the required review form for the DPO. The DPO provides assessment, approval/denial, and/or recommendations. High-risk projects require a full Privacy Impact Assessment, covering in detail all 7 components of the City's Digital Privacy Policy. Medium-risk projects only require an abbreviated review. A separate review on the project's cybersecurity from the Cybersecurity Office is often required as well.
6. **Submittal:** High-risk projects without a previously approved DUP are submitted to the City Manager's Office (CMO). Pending CMO review the DUP rises to a relevant Council Committee. Separately, the review and DUP are published.⁴ Note that the published DUP and review contain no personal information or other confidential information.
7. **Ongoing Monitoring:** Departments report annual metrics defined in the DUP and Privacy Impact Assessment. Reports usually require metrics on data usage and project effectiveness. Public can comment on data usage and annual updates online.

³ As of writing, Data Usage Protocols can be found online at <https://www.sanjoseca.gov/digitalprivacy>

⁴ As of writing, these can be found online at <https://www.sanjoseca.gov/digitalprivacy>

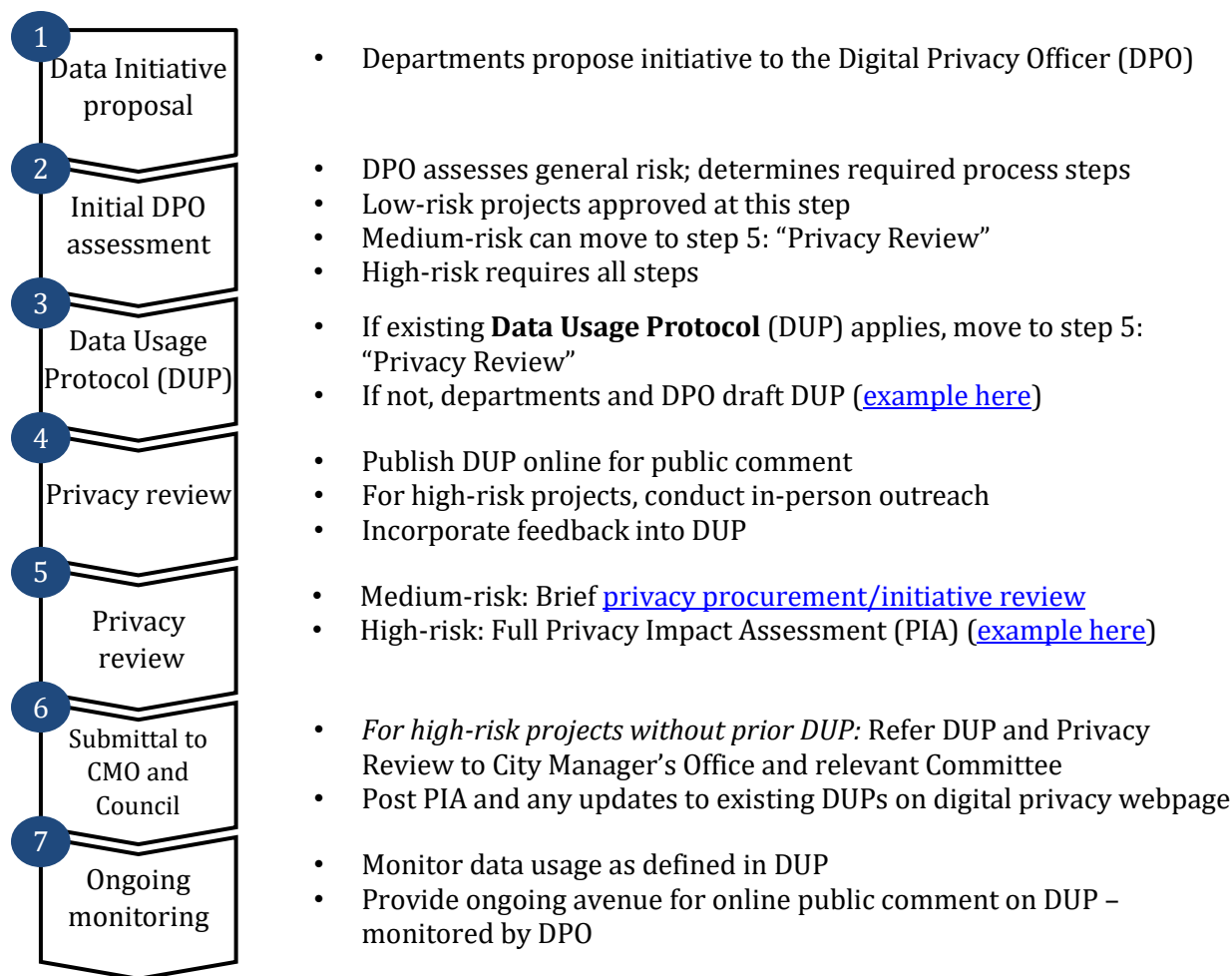


Figure 1: City privacy protocol developed from best practices and use-cases. Steps 3-7 may not be required pending assessed risk in step 2. The image lists the seven potential steps of the DPO’s privacy process.

Initiating privacy review

The formal review process can be initiated through the [City’s helpdesk site](https://csjhelpdesk.sjcity.net/) (<https://csjhelpdesk.sjcity.net/>) by going to the Request tab. For informal guidance, or help starting a review, contact the Digital Privacy Office at digitalprivacy@sanjoseca.gov. When filing a helpdesk ticket, you will be required to enter a “Request Type”. The type depends on the action:

1. To procure a service, technology, or other system that involves PII usage, select “Technology Procurement”. The DPO review is automatically included on the same ticket used for general technology procurement reviews done by the Information Technology Department (ITD).
2. For new PII usage that does not involve a procurement, select “Technology Inquires” and add the words “PRIVACY REVIEW” at the start of the Subject (see figure below)

The screenshot shows a web form titled "Help Request" with a navigation bar at the top containing "Request", "History", "Assets", "FAQs", "Messages", and "Profile". The form fields are as follows:

- Request Type:** A dropdown menu with "Technology Inquiries" selected.
- Subject:** A text input field containing "PRIVACY REVIEW: project name here".
- Hint:** A green callout box with the text: "Use this if you could not find a suitable category for your request and have a technology related question or any issue".
- Request:** A large text area containing the instruction: "Details on request. Should include purpose, data you plan to collect, and who will be able to access it (City departments, outside organizations, etc.). Be sure to attach relevant files, such as a project charter, RFP, and third-party privacy reviews."
- Alternate phone number:** A text input field with the label "Please provide an alternate phone #: (enter '0' to bypass)*" and an information icon.
- Carbon Copy (Cc):** A text input field with a checkbox labeled "Enabled".
- Attachments:** A section with an "Add File" button.

Figure 2: Screenshot of privacy review request through the [Helpdesk portal](#). When the initiative does not require a procurement, mark the request type as "Technology Inquiries", and add "PRIVACY REVIEW" to subject name. If it does require a procurement, pick request type "Technology Procurement"

When contacting the DPO, provide information on the project's purpose, the data that will be involved, and who can access the data (City staff, public, third parties). Attach relevant files, such as a project charter / description, a Request for Proposals (RFP), or other foundational documents.

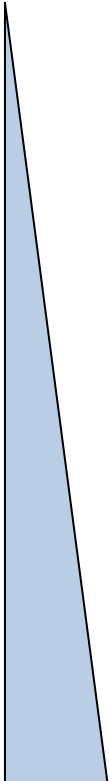
For a consultation on potential projects, to initiate an initial "threshold analysis", or to ask any questions, you can contact the Digital Privacy Office directly at digitalprivacy@sanjoseca.gov.

Threshold Analysis – Identifying a Project's Privacy Risk

Privacy Risk is evaluated in Step 2 of the review, otherwise known as the Privacy Threshold Analysis. Privacy Risk is evaluated on three key factors: type of data collected, purpose of data collected, and notice/consent provided. The figure below provides a summary of Privacy Risk, with details to follow. Privacy Risk is generally classified into 3 levels: Low, Medium, and High. Each factor contributing to Privacy Risk can also be defined at a low, medium, or high risk level.

1. **Low Privacy Risk** involves **anonymized information** used to **provide general improvements** to the City. **Notice is provided** upon collection if any personal information is involved
2. **Medium Privacy Risk** involves **identifiable information not traditionally kept hidden** (e.g., name, email, phone #) to provide **targeted government services desired** by the data owner. **Notice is provided** at time of collection and often requires written consent

3. **High Privacy Risk** involves **identifiable information traditionally kept hidden** (e.g., SSN, credit card #) that may provide **targeted services or punitive services**. **Notice may not be provided** upon collection or provided in a very limited format.



	Data collected	Purpose	Notice/Consent
Low-risk	Non-PII / anonymous survey	N/A	N/A
Mid-risk	“Indirect” PII: Zip code, age, etc.	General improvement to government processes	Notice provided at time of collection; active consent required
	Direct PII: Name, email, picture	Targeted, desired government services	
High-risk	Sensitive PII: Health, criminal records, SSN, etc.		Limited notice or no consent required
	Direct & Sensitive PII	Punitive in nature	

Figure 3: Privacy Risk can be defined by type of data collected, purpose and notice

Below we separate risk into its three components: type of data, purpose, and notice/consent provided.

1. **Type of data collected:** measured on the ability for the data to be connected to an individual and sensitivity of information
 - a. **Low-risk data collected includes anonymized information or information with no person data** (e.g., a database of City roads, sewage, or other infrastructure). Some “Indirect” PII can be considered low-risk if it cannot be reasonably tied to an individual. For example, collecting one’s zip-code may be considered low-risk unless it can be connected to additional data—such as birth date and gender—that can be used to identify an individual directly. In general, low-risk data can be used and shared without concern.

- b. Mid-risk data collected includes “Direct PII” that can be connected to an individual but is not traditionally kept hidden.** This could include one’s name, address, photograph (non-explicit), email address, etc. “Indirect” PII can also classify as mid-risk if it is collected in certain combinations to reasonably identify an individual. In general, mid-risk data is not made public until it is anonymized.
 - c. High-risk data includes “Sensitive PII” that can be tied to an individual and is traditionally kept hidden.** This could include one’s financial information (credit/debit card, bank account number, etc.), government ID information (Driver’s License number, Tax ID, Passport, birth certificate, etc.), immigration status, biometric/health information, education records, or criminal records. High-risk data should not be collected unless necessary to perform a service or comply with a law/regulation and should never be made public.
- 2. Purpose of data collected:** measured by the capacity for the purpose to benefit or harm an individual or community.
 - a. Low-risk purposes include those that strictly benefit a general community** or the entire City with no reasonable negative impact on any citizen or community. This could include data that helps inform which neighborhoods should receive the most resources for youth programs, informing how to address food deserts, or how to improve intersection safety/efficiency.
 - b. Mid-risk purposes include those that strictly benefit an individual** with no reasonable negative impact on the individual or others. This could include identifying the specific children in a classroom that could benefit the most from being offered a service or program, identifying a family in need of broadband access, or tracking the life outcomes of individuals before and after engaging in a City service or program.
 - c. High-risk purposes include those that are punitive in nature** or present a **significant negative impact on an individual or community.** This could include data collected for law enforcement purposes, or directly impacting one’s credit score or other financial prospects.
- 3. Notice/consent provided:** measured by the availability of notice to the data subject.⁵ Also accounts for the understandability and level of consent provided regarding the notice.

⁵The data subject is the individual that is subject of the data. For example, if “Mary Lee” filled out an application for a City permit, the data subject for that application (and the datasets that include this information) would be Mary Lee.

- a. **Low-risk notices require active consent and easily understandable language.** The data subject should be given detail, in a language they understand, on what data is being collected, how it is being used, who can access it, and how long it will be kept. The data subject should also provide recorded consent (written, recorded audio, checking a box on an online form, etc.) before submitting the data to the City.
- b. **Mid-risk notices provide clear notice but may not require active consent.** This could include a notice on what tracker cookies a website might be using, or a sign at an intersection that informs people that video recording is in progress, and link where they can learn more about the data usage.
- c. **High-risk notice implies little to no clear notice or consent** provided upon data collection. This may include data collected during an emergency response (such as identifying the location of a 911-call) or some law enforcement activity.

Privacy review process

Following the threshold analysis, the DPO will guide you through the necessary steps of the review process. The review steps vary by the level or Privacy Risk determined by the DPO, and are outlined in the figure below with details to follow.

	Data collected	Purpose	Notice/Consent
Low-risk: quick approval	Non-PII / anonymous survey	N/A	N/A
Mid-risk: Privacy procurement review; may open for public comment	“Indirect” PII: Zip code, age, etc.	General improvement to government processes	Notice provided at time of collection; active consent required
	Direct PII: Name, email, picture	Targeted, desired government services	
High-risk: Privacy impact assessment; public comment or direct outreach	Sensitive PII: Health, criminal records, SSN, etc.		Limited notice or no consent required
	Direct & Sensitive PII	Punitive in nature	

Figure 4: Privacy review requirements based on level of assessed risk, which can be determined by the data collected, purpose, and notice/consent required

For all levels of review, be prepared to provide information on:

1. **What data fields are being collected** (e.g., full name, age, video, audio) and at what granularity (e.g., if data is stored by individual or if aggregated by zip code)
2. **How the data will be used** (e.g., to deliver meals). In some cases, especially with law enforcement, it will help to explicitly say what the data will not be used for (e.g., data will be used to solve crimes, but not to solve minor drug offenses or violations)
3. **What notice is provided** at the time of collection and if consent is required.

Low-risk actions can usually be approved with minimal review. The brief review will focus on confirming the information provided, which usually can be done by reviewing the vendor contract (in the case of a procurement) or reviewing the details of the initiative. Review may also involve a live discussion with the project owners.

Mid-risk actions require a Privacy initiative and procurement review,⁶ and may require additional live discussions. Following discussion, the privacy details (what is being

⁶ An example Privacy initiative and procurement review can be found on the digital privacy webpage ([sanjoseca.gov/digitalprivacy](https://www.sanjoseca.gov/digitalprivacy)) at - <https://www.sanjoseca.gov/home/showpublisheddocument/81734/637792259089970000>

collected, how it’s used, and what notice is provided) may be published for public comment on the [digital privacy webpage](#).

High-risk actions require a more thorough Privacy Impact Assessment,⁷ live discussions with the project owners, and a public comment period. For the highest risk projects, targeted public outreach may be required (e.g., public meetings and listening sessions), and a status update may need to be reported to Council and/or a subcommittee meeting (e.g., Smart Cities committee).

Following the review, the DPO may approve the project, reject the project, or require adjustments to the Data Usage Protocol before providing approval. This may require shifts in what data can be collected, how it can be collected and how it can be used.

Ultimately, the DPO’s approval depends on whether the project’s Civic Value clearly outweighs the Privacy Risks (see figure below). Rarely will the DPO reject a project outright. Projects of low and medium risk are routinely approved without much issue, and most high-risk projects can be negotiated.

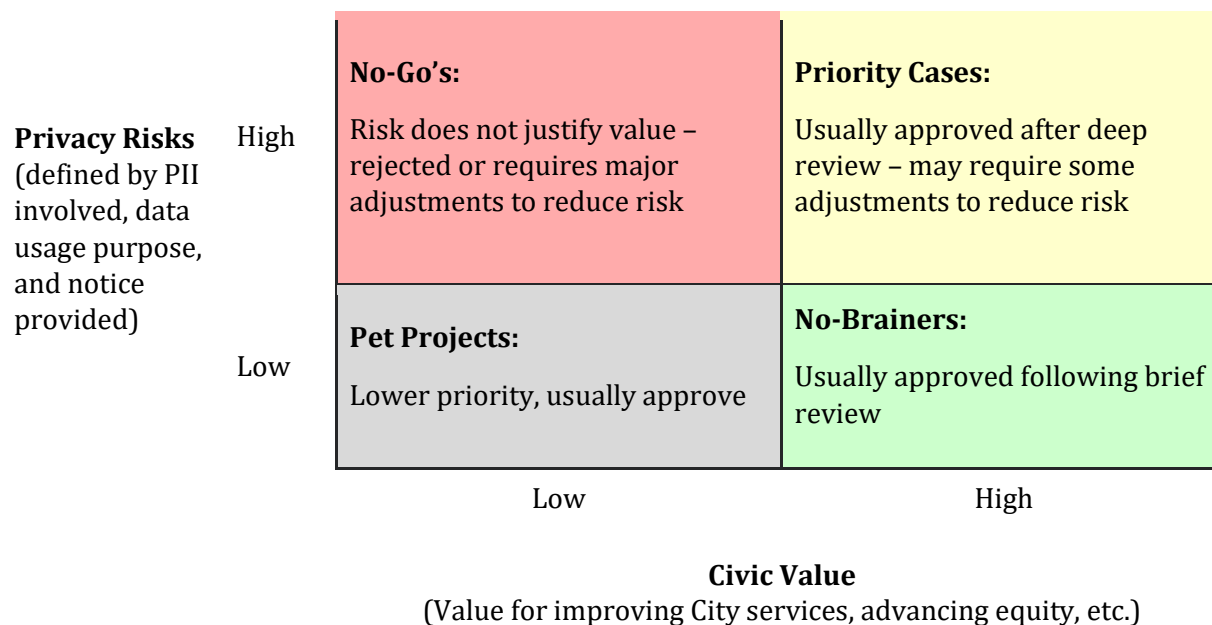


Figure 5: Relationship between Privacy Risk and Civic Value when determining project approval. Approval of a project with Privacy Risk requires that the Civic value clearly outweighs the Privacy Risk

⁷ An example Privacy Impact Assessment (PIA) can be found on the digital privacy webpage ([sanjoseca.gov/digitalprivacy](https://www.sanjoseca.gov/digitalprivacy)) at - <https://www.sanjoseca.gov/home/showpublisheddocument/81842/637793404561700000>

Data Usage Protocols and ongoing monitoring

Permissible data usage for a project is defined in its relevant Data Usage Protocol (DUP). The DUP is typically 1-3 pages and covers a type of data usage (e.g., video traffic monitoring). It is typically composed of six key elements:

1. Project summary
2. Authorized and prohibited uses of data
3. Practices for data collection, storage, notice, retention, sharing, access, and corrections
4. Summary of security practices and approval from the City's Cybersecurity Office
5. Required training for personnel using systems and technologies (if any)
6. Oversight and annual compliance expectations, written in an Annual Data Usage Report

The Annual Data Usage Report mentioned in item 6 is typically 1-2 pages, drafted by the applicable departments and details:

1. Project summary
2. Required performance metrics as defined in the Data Usage Protocol (e.g., accuracy, effectiveness, cost)
3. Plans for the data initiative (e.g., expansion, shift in usage)

This approach was adapted from best practices in privacy programs, including key collaboration with Santa Clara County's Privacy Office.

Once a Data Usage Protocol is drafted, it will be made available to the public. Projects with the highest risk require public notification through public forums, such as City Council meetings and in-person outreach to specific communities. The public can also comment on active DUPs online or by contacting the DPO directly at 408-793-6878 or through other public meetings on privacy.

DIGITAL PRIVACY POLICY

Context

The Digital Privacy Policy serves as the overarching document guiding all City privacy practices. It sets forth the framework for City departments to observe when information systems or other applications and forms collect, process, share, or otherwise use the public's personal information. It was initially approved by City Council on December 8th, 2020⁸ and went into effect on July 1st, 2021. The latest version can be downloaded [here](#) (last updated December 8th, 2020).

Policy

It is the Policy of the City to protect the privacy of individuals and the digital form of any Personally Identifiable Information that is collected, used, shared, or stored by the City. In addition to this Policy, the City is also subject to laws and regulations that govern the information collected.

To the extent permissible by law, City departments will adhere to the Privacy Principles,⁹ and this Policy based on those Privacy Principles including the following elements to protect individual privacy:

- **Notice:** Providing notice about the collection, use, and sharing of personal information at the time such information is collected. The City will make every reasonable effort to provide a privacy notice when basic municipal services are requested or delivered.
- **Retention:** Developing, maintaining, and following the City data retention schedule.¹⁰ Departments must ensure that identifying information is deleted or deidentified after the retention period expires. In the event of a conflict between this Policy and the Public Records Act, Sunshine Act, or other law governing the disclosure of records, other applicable law will determine our obligation in support of open and transparent government. See California Public Records Act.¹¹
- **Minimization:** Minimizing the collection and processing of identifying information and limiting collection to only what is necessary to provide services and to conduct business. When personally identifiable data is required to deliver or improve a service, departments must anonymize, de-identify, pseudonymize, or otherwise mask this information.
- **Accountability:** Maintaining documentation, available for public review and third-party monitoring, to evidence compliance with our privacy principles and Policy. If any information under our control is compromised or if residents are impacted due to a breach of security or negligent maintenance of information systems, the City will take reasonable steps to investigate the situation and notify those individuals whose information may have been impacted.

⁸ Link to council memo and attachments here:

<https://sanjose.legistar.com/LegislationDetail.aspx?ID=4700505&GUID=D2CE2587-F29F-4807-AB58-DB0E6B50A4F1&G=920296E4-80BE-4CA2-A78F-32C5EFCF78AF&Options=&Search=>

⁹ Link: <https://www.sanjoseca.gov/home/showpublisheddocument/79367>

¹⁰ Link: <https://www.sanjoseca.gov/home/showdocument?id=41127>

¹¹ Link:

https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=7.&chapter=3.5.&lawCode=GOV&itle=1.&article=1.

- **Accuracy:** Making every reasonable effort to provide the public with information on how predictive or automated systems are used and will institute processes to correct inaccurate information or methodologies in those systems. City Departments may use predictive or automated systems and technologies to support decision making, but some degree of human input and oversight into decision making is also required.
- **Sharing:** Following clear data governance procedures and instituting information sharing agreements when sharing information with outside entities, which shall strive to enable effective information sharing while following the City’s Privacy Principles and this Policy.
- **Equity:** The City is mindful of the populations it serves and how data about members of the public, including vulnerable populations, can and should be used. The City will strive to advance equity in a data-driven way while ensuring that PII is used only in accordance with this policy. The City will work to mitigate the impact of algorithmic and data bias.

UNDERSTANDING PERSONAL INFORMATION

“Personal Information” or “Personally Identifiable Information” (PII) refers to information that can directly or indirectly identify an individual. The City classifies PII into 5 categories of data:

- Personal data: information relating to an individual, such as a full name, street address, email address, Social Security Number, Credit card number, and personal computer or mobile device IP address.¹²
- Sensitive or demographic data: subsets of personal data that require extra security and care, such as biometric or genetic data, racial or ethnic origin, and religious or political affiliations. Note: Sensitive or Demographic data is not considered PII if it is only shared/collected/used in aggregate of a population larger than 1,000¹³ (e.g., # of registered voters in San José).
- Image data: digital pictures or photographs that can identify an individual by their face or other contextual information
- Recording data: audio or video information that can identify an individual by their face, voice, or other contextual information.
- Geolocation data: information affiliated with a computer, device, or vehicle that can be used to identify an individual based on physical location or on aggregate location patterns.

A non-comprehensive table of subcategories of Personal Information is shown below. A more exhaustive list of PII can be found in Appendix B.

Category of PII	Sub-categories
Personal Data	General: Full name; Home address; Date of birth; Place of birth Technology: Email address; Phone number; Phone, laptop, or other device IP address; Vehicle make, model and year

¹² An Internet Protocol address (IP address) is a numerical label such as *192.0.2.1* that is commonly associated with a device connected to the Internet. An IP address serves two main functions: network interface identification and location addressing.

¹³ Based on reporting requirements used for anonymity by the US Department of Health and Human Services [AFCARS Foster Care Dataset](#); refer to the [2021 codebook, element #6](#)

Category of PII	Sub-categories
	<p>Government-issued ID: Driver’s License; Passport; Social Security Number; Federal Employer ID or Tax ID; Employee ID number; License Plate</p> <p>Financial data: Credit or debit card information; Bank account, brokerage account or other financial information; Income; Wealth/Assets</p> <p>Other written or scanned information that can directly tie to an individual or household</p>
Sensitive PII or demographic-related PII	<p>Health data: Biometric data; Genetic data; Physical identifiable characteristics; Accessibility concerns (e.g., Mobility, Hearing, Vision); Other health records</p> <p>Race/Ethnicity: Race or ethnic origin; Nationality; Immigration status</p> <p>Religion/Politics: Religious affiliation; Political affiliation; Voter status</p> <p>Gender/Sex: Gender; Sexual Orientation; Sex assigned at birth</p> <p>Sensitive personal records: Education records; Criminal records</p> <p>Other sensitive written or scanned information traditionally kept confidential</p> <p>NOTE: Not PII if data is only shared/collected/used in aggregate of a population larger than 1,000¹⁴ (e.g., # of registered voters in San José)</p>
Image data	Picture that can identify an individual by their face or other physical and contextual information ¹⁵
Recording data	<p>Video that can identify an individual by their face or other physical and contextual information</p> <p>Audio that can identify an individual by their voice or other contextual information</p>
Geolocation data	Data affiliated with a vehicle, computer, or other device that can be used to identify an individual’s physical location

While the list above provides a good idea of what can be considered Personally Identifiable Information, what’s considered PII may change as technology evolves to collect more data. For example, tracking one’s eye movement may become a common form of PII if eye-tracking technology becomes more widely available. If you feel some new information may be considered PII, contact our Digital Privacy Office at digitalprivacy@sanjoseca.gov for guidance and support.

¹⁴ Ibid (same footnote as above).

¹⁵ An example of “contextual information” being used to identify someone could include a picture of a license plate, car make model and year, or a picture of someone’s backside next to a house with a visible address.

DATA COLLECTION GUIDANCE – MINIMIZING DATA

The City’s Privacy Policy applies to all data collection done by the City and by third-parties on behalf of the City.¹⁶ This section provides some general guidance and considerations when collecting data to comply with our Privacy Policy and to build public trust in our data usage.

Data collection practices can generally be classified into four types of data collection:

1. Video recording / images
2. Audio recording
3. Other sensors and activity tracking (e.g., GPS location tracking, online activity tracking)
4. Direct data submissions, such as submitting an application or filling out a survey

Each of these collection types present their own expectations for Notice and risk of unintentionally collecting more data than needed (i.e., going against the City’s Data Minimization requirement). The figure below provides a general overview of expected data collection practices by method.

	Video recording / images	Audio recording	Other sensors / activity tracking (e.g., GPS, online activity)	Direct data submissions (e.g., surveys, applications, interviews)
Notice / Consent expected ¹⁷	Notice through visible sign around camera Consent implied via notice Applies to Public/City spaces only; private locations require express consent	Express consent usually required, even in public areas ¹⁸ <i>Note: Express consent requires explicit agreement of collection, usually in writing</i>	Notice provided at point of collection (e.g., site home page, near sensor) Consent usually implied through usage (e.g., using an app) with opt-out controls ¹⁹ N/A if data is anonymous	Notice provided with data submission Consent varies by data collected and Privacy Risk, but usually implied through submission

¹⁶ With some exceptions outlined in the Privacy Policy, such as during an emergency call (e.g., 911 call)

¹⁷ The two main forms of Notice and consent are express consent and implied consent. Express consent requires explicit agreement (usually in writing) to the data usage outlined in the Notice. This can be done via a signature, a check box online, or other explicit method. Implied consent only requires that the Notice be easily accessible to the data subject, and the data subject is made aware of the Notice and given the opportunity to read the Notice before data collection

¹⁸ Based on existing law under CA Penal Code 632

¹⁹ Opt-out controls allow the user to control how much data they provide. Opting out of some collection may impact the amount of services the data subject can receive.

Minimization of data collected	High risk of unintended collection May require scrubbing, prompt deletion, etc.	Mid risk of unintended collection May require scrubbing, voice anonymization, etc.	Low risk of unintended collection Can explicitly target only data needed	Low risk of unintended collection Can explicitly target only data needed
--------------------------------	--	---	---	---

Figure 6: Collection practices and innate Privacy Risks vary by collection method. Projects are expected to provide at least the minimum consent/notice outlined above, and should do as much as possible to minimize the amount of unintended data collected. For example, if a motion sensor can collect the necessary data, then a camera should not be installed.

Collection practices - Video recording and images

Video and image recording presents a wealth of information but also opens a large door for Privacy Risk. For example, if a camera is set to record at an intersection for monitoring traffic safety, that camera will also capture PII such as faces and license plates (see figure below).

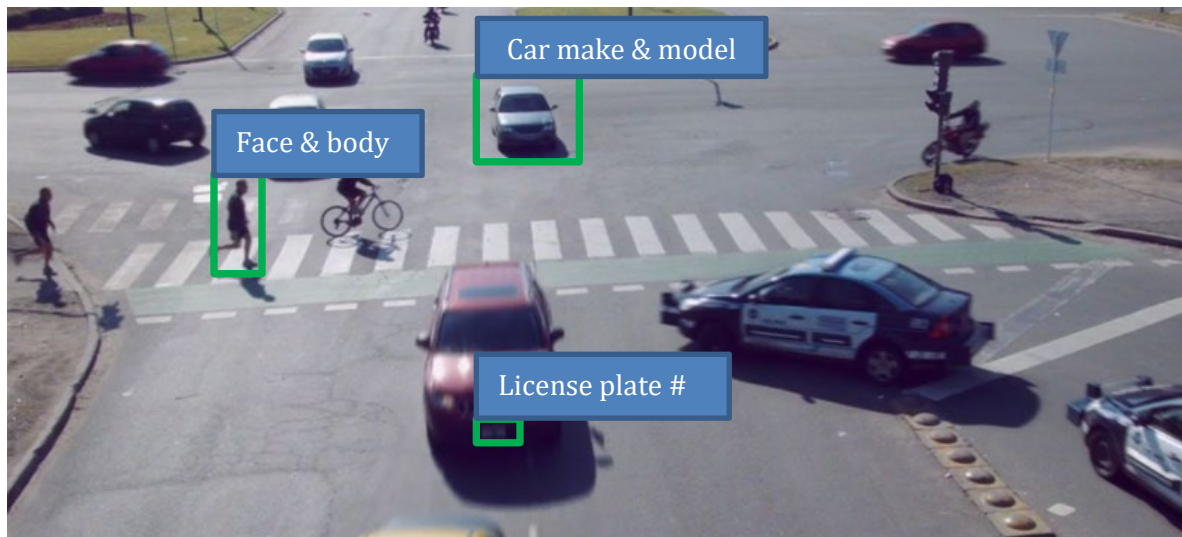
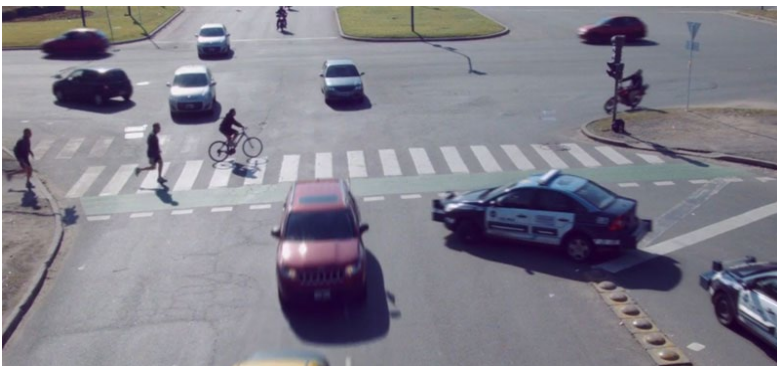


Figure 7: Example footage from a traffic camera. Unnecessary PII collected from a camera, such as one's face and license plates, raise the Privacy Risk of a project.

The potential for “over-collecting” data through video/images presents an innate Privacy Risk. Ideally, video can be collected and immediately deleted, or only brief video clips are stored.

For example, say the City’s Department of Transportation (DoT) needs to count the number of cars that pass an intersection, so they install cameras above the intersection. If DoT can record the number of cars that pass and then delete the footage, this drastically reduces Privacy Risk of the project, turning it from mid-high to low Privacy Risk (see figure below).

Footage collected: high amounts of PII



Data stored: no PII

- Intersection name
- Date & time
- # cars
- # pedestrians
- # bikes
- # motorcycles
- # near-collisions

Figure 8: By translating video footage into anonymized metrics (and deleting the footage immediately after collection), a department can eliminate most, if not all Privacy Risk.

When collecting video footage and images from a public space, such as a public park, public road or public spaces within City-owned property, the department should provide Notice in the form of a sign near the camera, or in the broader City area if near the camera is impractical. Simple signs can say “recording in progress”. For some higher-risk projects, additional signage may be required pending Digital Privacy Office review.

Recording in private areas, such as private offices in a City building, bathrooms, or people’s private residences / businesses requires express consent (i.e., written consent) from the individuals being recorded.

Collection practices – Audio recordings

Best practices for recording audio are like those of video recordings, but California imposes additional expectations for what is considered a “private space”. A conversation can still be considered as “private” in a public area if someone could have a “reasonable expectation of privacy” such as two people quietly talking in a library or while walking outside.²⁰ In general, projects should assume they have to get express consent from an individual before recording their audio. Exceptions can be discussed with the Digital Privacy Office and City Attorney’s Office.

Collection practices – Other sensors / activity tracking

This category of collection method includes other activity tracking that has a much smaller risk of collecting unintended PII. For example, the City’s website or digital app may track a user’s site activity, IP address, etc., but the City can easily choose which elements of information are stored or not stored. This minimizes risk of unintentionally collecting data.

Notice for tracking site activity should be available upon site entry, like on all pages of sanjoseca.gov (see figure below). Certain types of activity tracking, such as location tracking, require express consent, which can be done by checking a box on a website, pressing a button on a phone, etc.

²⁰ See CA Penal Code 632 for more information

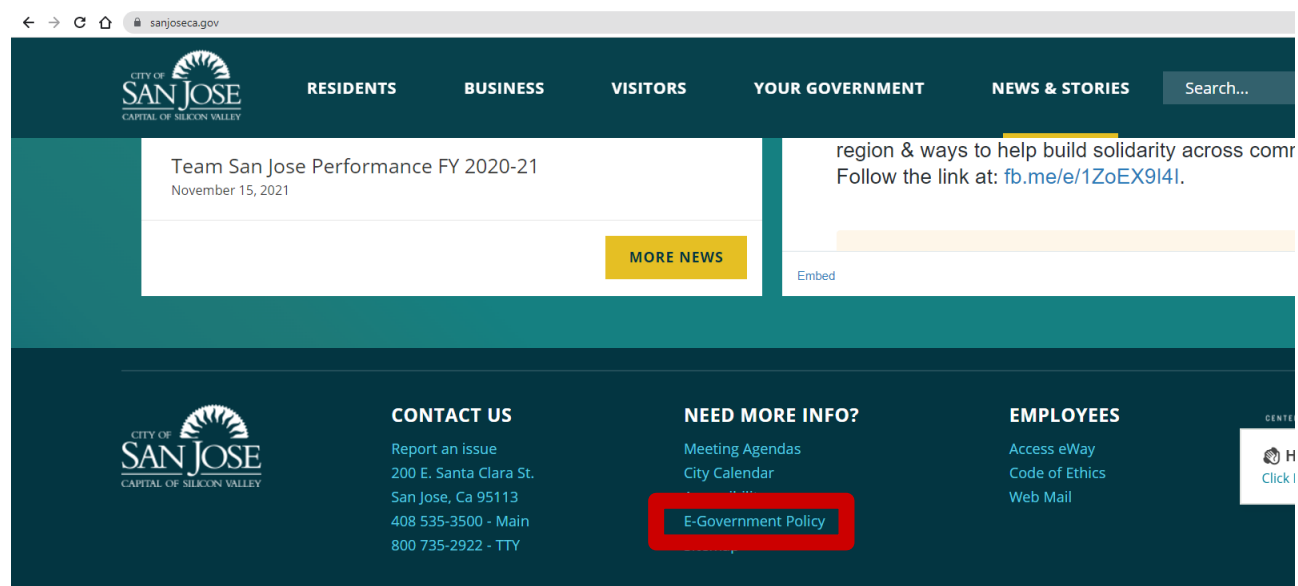


Figure 9: San José's website provides a link to its Privacy Notice (here called "E-Government Policy") on site pages.

Collection practices – Direct data submissions

Direct data submissions include applications, surveys, interviews, and any other means where an individual intentionally provides data to the City. Because the individual is providing the data themselves, there is a low risk of unintentionally collecting data.²¹

Best practices when collecting data through direct submission include:

1. Providing clear Notice to the individual on what is being collecting, how it will be used, who can access it, and how long it will be stored.
2. Ensuring the information is accurate and about the intended individual. For example, it's best practice to have a form confirm key elements of information such as an email address or password.
3. Confirming the information is about the intended individual. For example, a nurse may confirm their patient by asking for name and birthdate before discussing medical history.
4. Only collect the Personally Identifiable Information (PII) that is necessary for the purpose. Only collect sensitive information when legally required or functionally necessary to provide a service, such as Social Security Numbers, scans of passports, credit card numbers, etc.

Whenever possible, record only anonymous or indirect PII (such as zip code or birth month) to minimize Privacy Risk. Guidance for anonymizing data is provided in the "Guidance for Anonymizing data" section. Additional methods for anonymizing data can be discussed with the Digital Privacy Office.

²¹ There may be some instances where the act of providing information provides unintentional PII. For example, one's handwriting might be identifiable on a hand-written survey even if the survey is meant to be anonymous. A solution could be moving the data to a digital form and shredding the hand-written copy.

GUIDANCE FOR PROVIDING NOTICES

When personal information on residents is collected, notice should be provided. Notice, and the needs for notice can vary depending on the use case. For example, it may not be practical to provide notice alongside a drone, but notice should be accessible somewhere on the City website. Notice may not be provided in emergencies, such as during a 911 call.

When collecting information directly from a resident, such as through a form, notice should meet the following criteria:

1. Short and to the point – we have the [E-Government policy](#) as our robust terms and conditions. The goal of the notice is to lay out in plain terms how the data will be used.
2. Brief information on who the resident should contact to update or remove their data (if applicable). Alternatively, just allow residents to update their own information online through a portal
3. Reminder that the data will not be sold or used to investigate immigration status. This has been a core tenant of all our data usage, and will continue to be unless we get different direction from Council.
4. Link to the [E-Government policy](#) for more information.

For example, a notice for housing services could be:

“The information you provide will be used for support in the requested housing services and associated government services. All personal identifying information is kept confidential unless required by law. You can update your information by going to your account profile. Your data provided will not be sold for any purpose and will not be used to investigate matters related to immigration status. For more information on the City of San Jose’s data handling, see the [E-Government policy](#)”

A notice when collecting confidential information can be:

“The personal information you provide, such as your address, email, and demographics, will be treated as confidential by the City. However, we cannot guarantee your information will stay private in the event of a data breach. If the City becomes aware that this information was released to an unauthorized party (e.g., not working for the City), the City will notify the public. For more information on the City of San Jose’s data handling, see the [E-Government policy](#).”

A notice when conducting an anonymous survey can be:

“By checking this box, you consent to providing this anonymous information to the City of San José. The City may use this information to inform decisions. The data will not be sold and will not be used to connect back to an individual. For more information on the City of San Jose’s data handling, see the [E-Government policy](#).”

A short general notice can be:

I consent to submitting this information to the City of San José in accordance with the City's [E-Government Policy](#). The data I provide will not be sold or shared with any third party unless required by law.

GUIDANCE FOR ANONYMIZING DATA

One of the best ways to reduce Privacy Risk is to anonymize the data, or strip away personally identifying information. There are many ways to anonymize data, some of which are detailed below. The Digital Privacy Office is available to support in anonymizing data, contact us at digitalprivacy@sanjoseca.gov.

Data anonymization detailed through a brief example

Often the City needs data to study the impact, or effectiveness of its programs. For example, the City might want to know how a job training program has impacted the students' incomes. The City might collect data on students' incomes before and after the program in a database like this:

Table 1: Example table showing three hypothetical participants of a job training program, their incomes before the program and after the program

Name	Monthly income before program	Monthly income after program
Chase Garcia	\$2000	\$2303
Ty Kirk	\$2936	\$3244
Kevin Hsu	\$1677	\$2070

Because we have an identifying characteristic (full name) paired with sensitive information (monthly income), this database presents a privacy concern. However, the data is still valuable because it provides evidence that the program boosts income. Below we will walk through three ways to anonymize this data while still showing the impact of the program.

Option one: Remove the identifiers or replace with a pseudonym

The most straightforward option is to simply remove the personal identifiers, in this case the full name. The new dataset would now look like this:

Table 2: By removing the names, we can still see that monthly incomes tend to increase after the program. This still shows program impact while minimizing Privacy Risk

Monthly income before program	Monthly income after program
\$2000	\$2303
\$2936	\$3244
\$1677	\$2070

While the income information is still sensitive, it can be considered anonymous because it is not tied to a person. Alternatively, you can replace the “Name” column with a unique ID that is not tied to person. For example, replacing “Chase Garcia” with “ID: 1” or something similarly anonymous. Example of this “pseudonymization” is below:

Table 3: Another option would be to replace the names with unique IDs that cannot tie back to personal information

ID	Monthly income before program	Monthly income after program
1	\$2000	\$2303
2	\$2936	\$3244
3	\$1677	\$2070

This anonymization method is preferred whenever possible. It eliminates Privacy Risk by removing the personally identifiable information entirely.

Two factors to keep in mind when using pseudonyms:

1. Confirm that your dataset cannot be linked to another dataset to identify the person

Continuing the example above, if there is a separate dataset available that uses the same ID system, it may be possible to connect the two datasets and identify the person. For example, linking datasets A and B below could be enough to identify a person even if Dataset A alone presents no privacy concern.

Table 4: Even though Dataset A alone does not present a Privacy Risk, by linking it to Dataset B we can now identify individuals and their income

Dataset A		Dataset B	
ID	Monthly income	ID	Name
1	\$2000	1	Chase Garcia
2	\$2936	2	Ty Kirk
3	\$1677	3	Kevin Hsu

2. Confirm that multiple columns in your dataset cannot uniquely identify a person

Often the connection to identify a person is not as obvious. For example, if the dataset used zip code, gender, and birthdate rather than a random ID number, it could be possible to identify the person through a public source like a social media platform. Instead, use more vague identifiers, such as birth month + zip code or a uniquely assigned ID number like an assigned library card number.

Table 5: Zip code, birthdate, and gender can uniquely identify over 80% of US residents. In this scenario we can confidently say that the first row in dataset A is referring to Chase Garcia based on the zip code, gender, and birthdate, which we were able to find on twitter.

Dataset A

Zip code	Birthdate	Gender	Monthly income
95112	9/12/1998	He/him	\$2000
95115	7/05/1996	She/her	\$2936



Option two: Aggregate the data

“Aggregating” the data involves combining multiple records using some method that still allows us to get the information we need. In this example we want to show the impact of the job training program on monthly income, so we could aggregate the data by averaging income before and after the program. The averaged dataset is now anonymized and presents no Privacy Risk, while still showing the impact of the program.

Table 6: Averaging income before the program and after the program from Dataset A allows us to study impact via Dataset B without a Privacy Risk

Dataset A

Name	Monthly income before program	Monthly income after program
Chase	\$2000	\$2303
Ty	\$2936	\$3244
Kevin	\$1677	\$2070

Dataset B: Average income

Average income before program	Average income after program
\$2204 $= (2000+2936+1677)/3$	\$2539 $= (2303+3244+2070)/3$

Alternative ways to aggregate data include but are not limited to:

1. Summing – adding all values in a column. Useful to compare totals
2. Max change – select the max difference between columns. Useful to show the largest change before and after the program
3. Minimum change – select the minimum difference between columns. Useful to show the smallest change before and after the program. Combined with “Max change”, we can see the observed range of possible impacts

Option three: Encrypt the data or otherwise “mask” it

If your team needs to store the personally identifiable information, you will want to encrypt the data or mask it so that only authorized individuals can see the real data. Masking data involves altering the data so even if someone accesses the dataset, they can only understand it if they know how it was altered.

For example, a very simple masking would be to adjust the letters in the names by one character in the alphabet, so “A” becomes “B”, “B” becomes “C”, and “Z” becomes “A”.

Table 7: a very simple masking would be to adjust the letters in the names by one character in the alphabet, so “A” becomes “B”, “B” becomes “C”, and “Z” becomes “A”. Simple masking is not recommended because it could be unmasked by someone determined to access the data.

Dataset A

Name	Monthly income
Chase	\$2000
Ty	\$2936
Kevin	\$1677

Dataset B: Names masked

Name	Monthly income
Dibtf	\$2000
Uz	\$2936
Lfwjo	\$1677

However, a simple masking method can usually be unmasked by someone determined to access the data. A better option would be to encrypt the data using a much more complicated masking technique. Some database software offer encryption services built-in, so only those authorized by a password or (preferably) a device (like two-step authentication) can access the unencrypted data. An encrypted version of the example dataset using the standard SHA-256 method is shown below.

Table 8: A better option for masking data would be to encrypt the data, which is available in some database software. This way only those with a password or two-step authentication can view the unencrypted data.

Dataset A

Name	Monthly income
Chase	\$2000
Ty	\$2936
Kevin	\$1677

Dataset B: data encrypted

709a23220f2c3d64d1e1d6d18c4d5280f8d82fca	7585b52ed5ebaab27ce0fed7acdfc645023902ed805758b617e200dcc4567b5b
1ae6e6416a7ad654c3bc276ec2ee2a048f7a37b42970f3140d21d0abb625780d	81a83544cf93c245178cbc1620030f1123f435af867c79d87135983c52ab39d9
7c596f8cf2973c105b7b7229578c06e7877918f757c3d7b1564415456b81dfb9	a93706e865c271f4741a4a5583981818a6697fc88c0b2f15d8781c10eb21e431
0e4dd66217fc8d2e298b78c8cd9392870dcd065d0ff675d0edff5bcd227837e9	88b82308d570d71f02c5aa6a14ffe29a7ffbe14969a63094eb19e8e34256d9bf

For support on any data anonymization, you can always contact the Digital Privacy Office at digitalprivacy@sanjoseca.gov.

SURVEY GUIDANCE

The following six guidelines should be used when developing surveys for San José:

1. **Anonymize:** Anonymize surveys, and anonymize or do not collect IP addresses if feasible.
2. **No PII:** Do not use personally identifiable information (PII) if you do not want to go through the Digital Privacy Office (DPO) for review. If the PII is needed, then we will need to contact the DPO. Refer to Appendix A for a list of PII examples.
3. **Replace PII with Alternatives:** Replace PII with alternatives. For example, use ZIP code instead of home address or first name and last initial instead of full name. Refer to Appendix B for common examples of swaps.
4. **Sensitive Information:** Sensitive demographic information can be used if it's not connected to personal information (PII). In other words, both datasets should not be accessible and or used together. Refer to Appendix C for a list of sensitive information examples.
5. **Releasing datasets publicly:** Before releasing a dataset publicly, such as through the City's Open Data Portal, contact the DPO for review.
6. **Complexity:** If the request involves linking across surveys with varying degrees of personal information, or other kind of broad effort that involves many different surveys or data collection methods, the DPO needs to be contacted.

Personally Identifiable Information (PII) Examples

Personally Identifiable Information (PII) is information that may be used to identify an individual, trace their activities, or check their status. Be mindful of the personal information collected in surveys.

Types of PII commonly collected in surveys or forms	
<p>General</p> <ul style="list-style-type: none"> • Full name (first and last) • Home address • Date of Birth, including year <p>Contact info</p> <ul style="list-style-type: none"> • Email address • Phone number • Online profile (e.g., LinkedIn, Instagram) • Other means of direct contact 	<p>Issued IDs</p> <ul style="list-style-type: none"> • Driver's License • Passport • Social Security Number • License Plate • Other government or business-issued ID <p>Financial</p> <ul style="list-style-type: none"> • Credit card • Bank account number • Other financial information <p>Photographs, video, or audio of individuals</p>

Alternatives to PII

If you want to collect...	Collect ___ instead to better anonymize the data.
Full name	First name or initials
Home address	Zip code
Birthdate (day, month, year)	Age
Exact income (e.g., \$34,309.23)	Income range (<\$25k, \$25-50k,...)
IP address	"Masked" IP address or subset of IP address (e.g., last 4 digits)

Sensitive Information Examples

Sensitive information is a type of personal information. If revealed, it can leave an individual vulnerable to discrimination or harassment. Laws protect personal information as a whole, but add extra focus to sensitive information because of possible impacts to a person's livelihood, quality of life, and ability to participate in daily activities. The following are examples of sensitive information:

- Race or ethnic origin
- Religion
- Political affiliations
- Sexual orientation, gender identity, and expression (SOGIE)
- Criminal history
- Trade union or association memberships
- Biometrics, genetics, or medical history
- Languages used

SPECIAL CONSIDERATIONS FOR SENSITIVE DATA

Data related to health, education records and financial information require their own special considerations when processing and managing. A brief overview of applicable rules by sensitive information type is available below.

Health information

Health information includes:

- Biometrics such as heart rate, infection status and chronic diseases
- Medical history, pre-existing conditions
- Health care ID and medical claims

Note that some medical information, such as height, weight, and visible conditions (such as a physical disability) may be captured by a camera and are treated as less sensitive than non-visible medical information like pre-existing conditions, chronic diseases, and medical claims.

Handling of medical records and other sensitive health information is governed by federal and state laws. A list of relevant laws is provided by the [California Office of Health Information Integrity](#),²² and a summary is shown below. Usually HIPAA-specific training is required before handling sensitive information:

Federal laws

- **Health Insurance Portability and Accountability Act (HIPAA)** – HIPAA establishes national standards for the administration and protection of individuals' health information (e.g., medical or health records, personal health information). HIPAA training is typically required before handling sensitive health information. Educational materials to begin learning about HIPAA can be found at [HealthIT.gov](#).²³ Chapter 2 of the "[Guide to Privacy and Security of Electronic Health Information](#)"²⁴ provides a great entry point.
- **Confidentiality of Substance Use Disorder (SUD) Patient Records** – 42 C.F.R. Part 2 applies to federal assisted SUD treatment programs that meet the definition of a program within the regulation. These regulations apply to information that would identify a patient as having a SUD and allow very limited disclosures of information without a patient authorization.
- **Genetic Information Nondiscrimination Act (GINA)** – GINA protects individuals against discrimination based on their genetic information in health coverage and in employment.

State laws

- **Confidentiality of Medical Information Act (CMIA)** – This law protects the privacy of an individuals' medical information (in electronic or paper format) from unauthorized release by limiting disclosures by providers of health care, health plans, and contractors. CMIA was amended to further define administrative fines or civil penalties for any person or entity including licensed health care professionals who knowingly and willfully obtains, discloses, or uses medical information in violation of the CMIA.
- **Physical Safeguards from Health and Safety Code § 1280.18** – This law requires health providers to establish and implement administrative, technical, and physical safeguards to protect the privacy of patient's medical information. Each health

²² <https://www.chhs.ca.gov/ohii/health-laws/>

²³ <https://www.healthit.gov/topic/privacy-security-and-hipaa/health-it-privacy-and-security-resources-providers>

²⁴ <https://www.healthit.gov/sites/default/files/pdf/privacy/privacy-and-security-guide.pdf>

provider shall reasonably safeguard confidential health information from any unauthorized access, use, or disclosure.

- **Patient Access to Health Records – Health and Safety Code § 123100 and § 123111** – With minor limitations, this law gives patients the right to see and copy information maintained by health care providers relating to the patients’ health conditions. The law also gives patients the right to submit amendments to their records, if the patients believe that the records are inaccurate or incomplete.
- **Consent by Patient for Lab Results via Internet or other Electronic Means – Health and Safety Code § 123148** – If the patient requests, a health care provider shall provide the results of a laboratory test to the patient in written or oral form. Consent must be obtained (consistent with CMIA) to deliver results via electronic means. Electronic delivery or results shall be consistent with applicable federal law or state law. HIV antibody test, hepatitis infection tests, abusing the use of drugs, and tests related to routinely processed tissues revealing malignant results may not be conveyed by electronic means, unless the specific requirements of subdivision (f) of Health and Safety Code section 123148 are met. Test results and health information may not be used for commercial purpose without patient consent.

Education information

Education information includes “educational records” defined under the Family Educational Rights and Privacy Act (FERPA) which includes:

- Grades, test scores, etc.
- History of disciplinary action
- Other records that are directly related to a student and is maintained by an educational agency or institution (e.g., school) or by a party acting for the agency or institution²⁵

The main law that governs education information is the Family Educational Rights and Privacy Act (FERPA). According to the [California Department of Education](#),²⁶ FERPA typically requires that a child’s “educational records” are only accessible by:

- The parents/guardians of the child
- The child after reaching the age of 18 or attending school beyond high school
- The school, school district, and parties acting for the school or district
- The California Department of Education

Additional information on FERPA is provided by the [US Department of Education](#).²⁷

²⁵ 20 U.S.C. 1232g(a)(4)(A); 34 CFR § 99.3 “Education Record”

²⁶ <https://www.cde.ca.gov/ds/ed/dataprivacyferpa.asp>

²⁷ <https://studentprivacy.ed.gov/>

Financial Information

Financial information includes:

- Credit card and debit card information (number, security code / pin, expiration date, etc.)
- Bank account and brokerage account information
- Personal check data or scanned images of checks
- Income/Salary/Wage data

In general, financial information should be treated as highly sensitive and only requested when necessary. **When looking to handle financial transactions (especially online), it is highly advised to find a third-party vendor to process payments securely (e.g., PayPal, Stripe, Zelle).** This removes the need to navigate several state and federal laws.

Information on income/salary/wages is sensitive when paired with identifying information such as one's name or address. This information can usually be stored in aggregate (such as the average income of a neighborhood) or disconnected from other identifying information without introducing much Privacy Risk. For guidance, contact the Digital Privacy Office before collecting data on an individual's income/wage or socioeconomic status (digitalprivacy@sanjoseca.gov).

Some of the relevant financial privacy laws and regulations include:

- Fair and Accurate Credit Transactions Act (FACTA) of 2003 – covers identity theft through the handling of consumer account information
- Gramm-Leach-Bliley Act (GLBA) of 1999 – establishes rules for financial institutions to create and implement privacy policies, and allow customers to opt-out of having their private information disclosed
- Fair Credit Reporting Act (FCRA) – governs access to credit reports, accuracy and fairness in credit reporting, collection of credit information, and other aspects related to credit information
- Payment Card Industry Data Security Standard – industry-imposed security standards that an organization must meet to process payments of major credit cards (American Express, Visa, Mastercard, etc.)

SAN JOSÉ'S INTRODUCTORY TRAINING TO DIGITAL PRIVACY

Introduction

The City of San Jose strives to be the most innovative City in the country. As we collect, use, and share more information to drive that innovation, we require that the **data collected on our communities is used to support those communities**. Our communities should understand what data is being collected, how it's being used, and who can access it.²⁸

The City of San José interacts with personal information (names, home addresses, credit cards, etc.) every day. As you read through this training, consider how your department or office may collect or use personal information in support of the City's mission. Even if you do not interact with the public daily, your job may still require you to protect personal information.



It is everyone's responsibility to protect privacy, including City employees, contractors, volunteers, vendors, and others working on behalf of the City. Throughout this course , we will collectively refer to these individuals as **“staff.”**

Objectives

This training is designed to improve your understanding of the following topics:

- The fundamentals of privacy, including the definition of personal information;
- The importance of protecting privacy, including why it needs to be protected, and the consequences when privacy is not protected;
- The responsibility of all City staff to protect privacy;
- And the City of San José Privacy Office and its functions.

Course Instructions

Please read the following pages carefully and participate in the Knowledge Check activities. Report any questions or comments to the Privacy Office (digitalprivacy@sanjoseca.gov)

Privacy Fundamentals

History of Privacy

Privacy is deeply rooted in history, as evidenced in ancient texts, art, and even architecture. The concept of privacy has evolved in response to changing social, legal, and other factors. But the concept of privacy, as we know it today, can be traced back to its origins in British legal traditions from the 1700s.

²⁸ While these three components form the most essential elements to inform our communities, we should also be prepared to explain where the data exists (when being stored, used, or shared), how long it will be kept, and who they can contact to request adjustments to inaccurate data and (where applicable) request deletion of data.

Privacy in the United States

In the United States, there is no single overarching law that governs privacy or sets requirements for how personal information should be handled. Instead, privacy is regulated through a variety of federal, state, and local laws that apply to specific sectors, such as healthcare, finance, and consumer reporting, among many others.

Today, the concept of privacy is enshrined either implicitly or explicitly in several different sources.

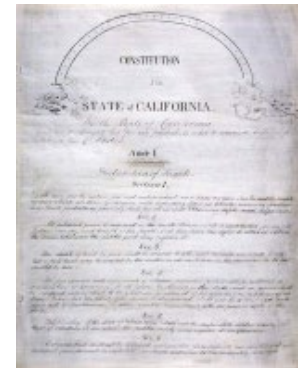


For example, although the U.S. Constitution does not contain an express right to privacy, certain amendments support its protection, such as the protection against unlawful search and seizure and the provision of a right to due process.

At the state level, some state constitutions include explicit provisions relating

to a right to privacy. For example, California's constitution guarantees each citizen an "inalienable right" to pursue and obtain "privacy."

At the local level, the City's Digital Privacy Policy, Privacy Manual, and Digital Privacy Office guide the handling of personal information. We will learn more about some of these later in the course.



What is Privacy?

Privacy can be interpreted in a variety of ways, depending on the social and legal environment. It is a complex concept, and its precise meaning is the subject of ongoing debate and discussion. For our purposes, privacy can be described as **the protection of personal information**. Protecting privacy means supporting the privacy rights of constituents and staff and being thoughtful in how personal information is collected, managed, and used in conducting City business. The Digital Privacy Office works to strike a balance between the City's need to collect and process personal information with its responsibility to protect personal information and treat it with integrity and respect. At

the City of San Jose, it is everyone's responsibility to protect privacy. Next, we will discuss personal information and the need to protect it.

Privacy is the protection of **personal information**.
Privacy supports the privacy rights of constituents and staff.
It is *everyone's* responsibility to protect privacy.

What is Personal Information?

Personal information is any information about an individual that may be used to identify them directly or indirectly. Some common examples of personal information include:

- *Name*, such as full name, maiden name, mother's maiden name, or alias.
- *Contact information*, including telephone numbers, email addresses, or screen names.
- *Numerical identifiers*, such as Social Security number (SSN), driver's license number, or immigration number.
- And *personal characteristics*, including photographic images, fingerprints, or other biometric data.



The City has many reasons to collect, store, use, process, and share many different types of personal information. City residents entrust us with their personal information and expect us to use it responsibly in the provision of City services and to protect it against those who do not need it or should not have it.

Direct and Indirect Personal Information

Personal information not only includes information that can be used to directly identify a person, like their Social Security number, but also includes categories of information that may apply to many people.

For example, several people in a group could have the same birthday. When alone, it would be difficult to identify a specific person based on birthday alone. However, when combined with other pieces of information, such as their name, telephone number, or photograph, this information would only apply to one individual.



Antonio Smith
SSN: ###-##-####

Direct personal information uniquely and directly identifies an individual. Examples of direct personal information include name, Social Security number (SSN), email address, and biometric data, like fingerprints, among others.

Indirect personal information, however, may need to be combined with other data to identify someone. Examples of indirect personal information include age, ethnicity, political affiliation, place of birth, or other similar pieces of information that could apply to many different people but could also be used to identify a specific person.



Born on October 2
(Indirect Personal Information)

What constitutes personal information is not anchored to any single category of information or technology. Rather, it may call for a case-by-case assessment of the specific risk that an individual could be indirectly identified. In other cases, the definition of what may be considered personal information varies depending on the law or policy in question.

Regardless of whether personal information directly or indirectly identifies a person, it should be protected, and treated with respect. **If you have any questions or concerns about personal information or other potential privacy issues, please reach out to us at the Digital Privacy Office - digitalprivacy@sanjoseca.gov.**

The Difference between Privacy and Security

Privacy and the Digital Privacy Office are separate from security and the Information Security Office. However, **privacy and security are separate but mutually supportive disciplines.**

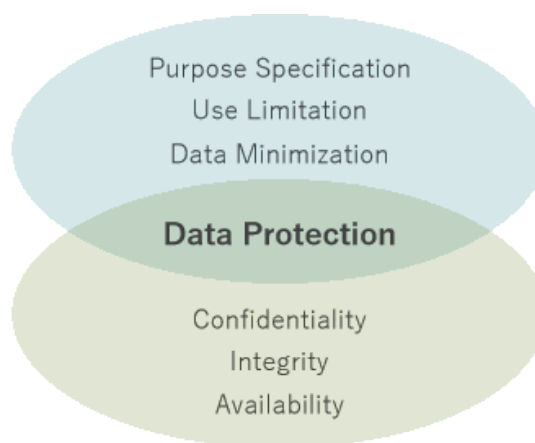
Privacy is concerned with the protection of personal and sensitive information that can directly or indirectly identify an individual. Privacy often involves policy decisions regarding the appropriate collection, sharing, and use of personal information. Privacy not only considers how personal information is secured but also the specific purpose for collection, defining uses of that information, and reducing risk to individuals through minimizing what data is collected and maintained.

Security, on the other hand, is concerned with all forms of data, not just personal information. Security is a set of strategies for managing the policies, processes, and tools necessary to maintain the confidentiality, the integrity, and the availability of an organization's information systems, regardless of the type of data. Security often focuses more on the practical implementation of policy decisions by using physical, technical, and administrative means to protect systems and data from malicious attacks or limiting opportunities to steal data for profit.

Simply stated, security is about protecting systems, while privacy is about protecting personal information. Privacy and security share their commitment to the goal of data protection. Although security is necessary for protecting personal information, it is not sufficient on its own to meaningfully address privacy. The Privacy Office and the Information Security Office work closely together on many projects to protect personal information processed by the City.

Privacy – The protection of personal information; concerned with how personal information is collected, shared, and used.

Security – The protection of all forms of data; concerned with the implementation of physical, technical, and administrative controls to protect systems and data from attacks.



General categories of Personal Information

“Personal Information” or “Personally Identifiable Information” (PII) refers to information that can directly or indirectly identify an individual. While there are many ways to categorize PII, the City classifies PII into 5 categories of data:

- **Personal data:** information relating to an individual, such as a full name, street address, email address, Social Security Number, Credit card number, and personal computer or mobile device IP address.²⁹
- **Sensitive or demographic data:** subsets of personal data that require extra security and care, such as biometric or genetic data, racial or ethnic origin, and immigration status. Sensitive or demographic data is not considered PII if it is only shared/collected/used in aggregate of a population larger than 1,000³⁰ (e.g., # of registered voters in San José).
- **Image data:** digital pictures or photographs that can identify an individual by their face or other contextual information

²⁹ An Internet Protocol address (IP address) is a numerical label such as *192.0.2.1* that is commonly associated with a device connected to the Internet. An IP address serves two main functions: network interface identification and location addressing.

³⁰ Based on reporting requirements used for anonymity by the US Department of Health and Human Services [AFCARS Foster Care Dataset](#); refer to the [2021 codebook, element #6](#)

- **Recording data:** audio or video information that can identify an individual by their face, voice, or other contextual information.
- **Geolocation data:** information affiliated with a computer, device, or vehicle that can be identify an individual, their location, or general location patterns.

A non-comprehensive table of subcategories of Personal Information is shown below. A more exhaustive list of PII can be found in Appendix B.

Category of PII	Sub-categories
Personal Data	<p>General: Full name; Home address; Date of birth; Place of birth</p> <p>Technology: Email address; Phone number; Phone, laptop, or other device IP address; Vehicle make, model and year</p> <p>Government-issued ID: Driver’s License; Passport; Social Security Number; Federal Employer ID or Tax ID; Employee ID number; License Plate</p> <p>Financial data: Credit or debit card information; Bank account, brokerage account or other financial information; Income; Wealth/Assets</p> <p>Other written or scanned information that can directly tie to an individual or household</p>
Sensitive PII or demographic-related PII	<p>Health data: Biometric data; Genetic data; Physical identifiable characteristics; Accessibility concerns (e.g., Mobility, Hearing, Vision); Other health records</p> <p>Race/Ethnicity: Race or ethnic origin; Nationality; Immigration status</p> <p>Religion/Politics: Religious affiliation; Political affiliation; Voter status</p> <p>Gender/Sex: Gender; Sexual Orientation; Sex assigned at birth</p> <p>Sensitive personal records: Education records; Criminal records</p> <p>Other sensitive written or scanned information traditionally kept confidential</p> <p>NOTE: Not PII if data is only shared/collected/used in aggregate of a population larger than 1,000³¹ (e.g., # of registered voters in San José)</p>
Image data	Picture that can identify an individual by their face or other physical and contextual information ³²
Recording data	<p>Video that can identify an individual by their face or other physical and contextual information</p> <p>Audio that can identify an individual by their voice or other contextual information</p>

³¹ Ibid (same footnote as above).

³² An example of “contextual information” being used to identify someone could include a picture of a license plate, car make model and year, or a picture of someone’s backside next to a house with a visible address.

Category of PII	Sub-categories
Geolocation data	Data affiliated with a vehicle, computer, or other device that can be used to identify an individual's physical location

While the list above provides a good idea of what can be considered Personally Identifiable Information, what's considered PII may change as technology evolves to collect more data. For example, tracking one's eye movement may become a common form of PII if eye-tracking technology becomes more widely available. If you feel some new information may be considered PII, contact our Digital Privacy Office at digitalprivacy@sanjoseca.gov for guidance and support.

Protecting Privacy

There are many reasons to protect privacy and personal information, and by extension, protect residents and others who entrust us with their information.



As discussed earlier, the California Constitution guarantees residents a right to pursue and obtain privacy. Additionally, **many federal, state, and local laws set standards or requirements to enhance privacy**, and various City rules, policies, and ordinances require staff to protect personal and sensitive information.



Second, **public trust is essential**. People may entrust highly sensitive personal information to the City, including financial, health, or other information required to access City services. In some cases, these individuals may not have a choice in providing their information for the City to provide services, comply with the law, or otherwise serve its public function. This places a special obligation on the City to safeguard personal information and to treat it with integrity and respect. If we fail to protect personal information or proceed to use it improperly, we could undermine our citizens' faith in government.



Security breaches, or data breaches, are another important reason to protect privacy. A "security breach" is defined by state law as "the unauthorized acquisition of computerized data that compromises the security, confidentiality, or integrity of personal information." A data breach can affect any number of individuals. Even situations where only a single individual's personal information has been compromised can constitute a data breach.

Breaches of personal information are costly not only to the City but also to the person whose information has been targeted.

Examples of data breaches include, but are not limited to:

For questions contact: digitalprivacy@sanjoseca.gov

- Sending an individual's personal information to the wrong person through an incorrect email address;
- Staff inappropriately accessing the personal information of an individual for any purpose other than for legitimate City business;
- Losing a laptop computer, mobile device, flash drive, or other device that contains personal information due to misplacement or theft; or
- A malicious actor from outside the organization accessing a database containing personal information.

City staff, regardless of where they work, play a critical role in protecting privacy and the personal information of those we serve. If you suspect any of the above scenarios have occurred, contact our Digital Privacy Office immediately at digitalprivacy@sanjoseca.gov.

Privacy Harms

In addition to the City's responsibility to implement privacy standards and practices, preserve public trust, and work to prevent data breaches; it is important to understand the different types of harms an individual may experience if their personal information is not properly handled. Consider the following scenarios, which demonstrate the different types of harm an individual may experience if their personal information is not properly handled.

Financial Harm

Sarina received a past due notice for a credit card in her name. The notice stated that she had an overdue balance of thousands of dollars. Sarina had always been careful to pay her bills on time and knew she was current on all her payments.



In reality, Sarina is a victim of identity theft. **Identity theft** occurs when someone steals personal information, and then uses it to impersonate the victim or to commit fraud in their name. As a result, Sarina could potentially have her credit history damaged, face a hard time getting a loan, or have her credit limits decreased. It may also take a lot of paperwork, red tape, and money to clear her name.

Sarina's story is an example of **financial harm** that resulted from an invasion of her privacy.

Reputational Harm

Sandy recently interviewed for her dream job. Unfortunately, something on Sandy's background check concerned her potential employer, and they decided to extend the offer to a different candidate.

In reality, Sandy's background check contained incorrect information that left the employer with the wrong impression. Although she could file a request to have the information corrected, Sandy has already lost a great opportunity.

Sandy's story is an example of **reputational harm**.



Emotional Harm

Mirae has been struggling at work lately but has committed to improving her performance. She has already taken steps to advance her skillset. However, her most recent performance evaluation reflected some of her past mistakes.

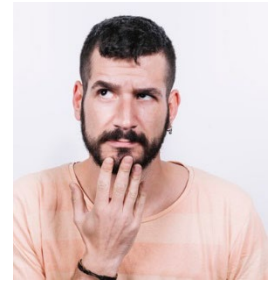


Before completing the review, her supervisor printed off a copy and accidentally left it in the break room, where several other employees saw the negative comments about Mirae's performance. Although she has recently made improvements, Mirae feels betrayed by her supervisor's carelessness and embarrassed by the coworkers who read her review.

Mirae's story is an example of **emotional harm**.

Loss of Trust

Dan is also a victim of identity theft. Unlike Sarina, he did not suffer from financial harms. After his identity was stolen, Dan became mistrusting of anyone who requested his information, including a service provided by his bank to monitor for future theft. Dan does not want to become a victim again, but he is unsure of what to do.



Dan's story is an example of a **loss of trust** resulting from an invasion of his privacy.

City Staff and Privacy

The privacy concepts that we have discussed so far lay the foundation for City staff to understand how they can improve the privacy practices of the City and its departments. **At the City of San Jose, it is everyone's responsibility to protect privacy.**

All City staff are required to do the following:

Follow All City Policies



Staff who conduct City business are governed by a variety of different policies, including the City Policy Manual,³³ City Ordinances, and Board, Administrative, and Department Policies. Many of these policies are relevant to privacy and the handling of personal information.

The following are examples of policies related to Privacy that staff are required to follow:

- City Policy Manual, primarily in sections 1.2.1, 1.7.1, 3.3.4, 6.1.2, and 6.1.5
- City of San Jose Retention Schedule
- Privacy Manual

³³ <https://www.sanjoseca.gov/your-government/city-manager/employee-relations/city-policy-manual>

Only Share Data with Those with a “Need to Know”



In general, everyone who uses or receives personal information while conducting City business should have a **“need to know.”** The principle of need to know means the person has a legitimate business purpose for accessing the information; the person has been authorized for access; and their access can be logically explained and defended.

If you believe someone should not be using or receiving personal information or other City data, discuss these concerns with your manager or consult with the Privacy Office (digitalprivacy@sanjoseca.gov).

Safeguard Personal Information



Staff must safeguard personal information, in all forms, at all times. In accordance with best practice, paper documents should be secured when not in use and should not be visible to passersby. Personal information in electronic form, such as on a laptop, cell phone, or other portable storage device, should be secured appropriately. For example, locking your computer screen when you leave your desk could prevent onlookers from viewing, editing, downloading, or sending personal information (keyboard shortcut to lock screen: **⊞** +L). City data, whether in paper or electronic form, should not be left unattended in a public space or in a vehicle, even if it is locked. You are responsible for securing these assets at all times.

Report Both Known and Suspected Data Breaches or Information Incidents



City staff are required to report both known and suspected data breaches and information incidents. A **data breach** is defined by state law as “the unauthorized acquisition of computerized data,” and examples may include sending personal information to the wrong recipient, misusing City data, losing a City asset (laptop, phone, etc.), or hacking by a malicious actor.

A breach can include any event whereby some aspect of information security could be threatened: loss of data confidentiality; disruption of data or system integrity; disruption or denial of availability.

Despite best efforts to secure information, breaches and incidents can still occur. **City staff are required to report both known and suspected breaches and incidents, whether in electronic or paper form, no matter how big or small to our Information Technology Department.** To file a report, contact the [IT Help Desk](#) via the help desk portal or email

To report a known or suspected breach or incident, contact
cybersecurityteam@sanjoseca.gov
-or-
IT Self-Service Desk

cybersecurityteam@sanjoseca.gov. Do not wait to report a data breach or information incident.

Follow Privacy and Security Best Practices



Staff should do their best to follow all privacy and security best practices, including using strong passwords, managing privacy settings, and ensuring emails are only sent to intended recipients. Take proactive steps, like taking extra care when using the auto-populate feature in email applications; being aware of, and guarding against, **social engineering** attacks, like **phishing** and **spear-phishing** emails; and ensuring data is kept safe from onlookers without a need to know.

Encourage Others to Improve Their Privacy Practices



Finally, all staff should encourage their fellow coworkers to improve their privacy practices by modeling good behavior and providing reminders, as necessary. Consider engaging the Privacy Office in a conversation with your team about how to handle personal information.

It is everyone's responsibility to protect privacy and safeguard personal information at the City of San Jose. By exercising your responsibilities and remaining diligent against attacks, you can improve the privacy practices of the City and help protect those we serve.

APPENDIX B – MORE EXHAUSTIVE LIST OF PERSONALLY IDENTIFIABLE INFORMATION (PII)

PII Reference List

This PII Reference List includes 7 categories and types of PII and subsets of PII that are included when the City refers to "personal" or "sensitive" data or information throughout the Privacy Impact Assessment.

Personally Identifiable Information (PII)

First Name
Last Name
Alias Name
Maiden Name
Full Home Street Address
Zip Code
Date of Birth
Date of Death
Email Address
Photograph
Internet Protocol (IP) Address
Marital Status
Beneficiary Name
Beneficiary Contact Phone Number
Beneficiary Contact Address
Employee ID
Identifying Marks (e.g., tattoos, birth marks, etc.)
Identifying information of children, youth, minors under 18 years old

Sensitive PII Subset

SSN (full 9 digits)
Username/ID
User Hint Question and Answer
Driver's License Number
Vehicle Information (license plate #, vehicle ID# (VIN))
Passport Number
Biometric ID Data (fingerprint, iris scan, faceprint, etc.)
Voter ID Number
FEIN (Federal Employer Identification Number)
State or City ID Number
Criminal Justice Number (arrestee or prisoner numbers)
Alien Registration Number

Demographics Subset

Citizenship Status
Nationality
Sexual Orientation
Gender Identity
Background Check/Investigation Details or Results
Drug and Alcohol Abuse Information
Criminal Offenses/Convictions
Physical Characteristics
Political Party Affiliation
Political Party Affiliation
Military / Veteran Status
Race / Ethnic Origin
Religious / Philosophical Beliefs

Other Sensor Information

Audio Recordings
Phone Call Recordings
Video Recordings
Social Network Profile, Family Network Research and/or Friends/Contacts/Followers
Computer Use or Website Tracking/ Monitoring (cookies, web beacons, web widgets)
Location Tracking (individual or vehicle, geo-location, RFID Tracking, cell tower data)
Behavioral Pattern Mapping (e.g., physical, psychological, online, etc.)
Item or Identifier Scanning (contraband recognition, license plate reader, RFID reader)
Other Electronic Signatures or Monitoring (other cell phone signal, device sensors monitoring usage not previously stated)
Other Sensory Data (visual, audio, olfactory, or biometric not previously stated)
Other uncategorized surveillance information or data

Health Information Subset

Relative / Emergency Contact Name
Relative / Emergency Contact Phone Number
Relative / Emergency Contact Email
Relative / Emergency Contact Address
Disability Description
Health Diagnosis or Condition for Physical / Mental Health (non-substance use)
Health Diagnosis (substance use)
Health Services Provided
Medical Record Number
Health Plan / Insurance ID Number or Policy (inc. Medicaid & Medicare)
Medical Payments or Health Insurance Payments (incl. Medicaid & Medicare)
Health Policy Group Number
Patient ID Number
Medical Records

Prescriptions / Medications

Financial Information Subset

Bank or Financial Account Number

Credit Card / Debit Card Number

Other Credit / Debit Card Data (e.g., Expiration date, security code)

Personal Identification Number (PIN)

Personal Check Data or Scanned Images

Income/Salary/Wage Data

Socio-Economic Status

Credit Score, Credit Grade, or Credit History

Other Sensitive Information (organizational, children, unstructured)

Intellectual Property or Proprietary Information

Budgets, Financial Statements / Forecasts

Organizational Strategy, Business Decision, or Design Info

Legal Documents, Contracts, Vendor Agreements

Other Children's Information not previously stated

Other Confidential Information not previously covered

Any Unstructured Data that might include any of the above types of information