



OPEN DATA GUIDE

City of Las Vegas Open Data Guide

Contents

Background.....	2
Open Data Portal	2
Purpose	4
GOVERNANCE.....	4
Open Data Program Manager	5
IT Data Architect	5
IT Business Partners	5
Department Directors	5
Open Data Coordinators	6
Internal data review	6
Sensitive Information and Personally Identifiable Information	7
Community Engagement	8
Communications	9
Evaluating Reach.....	9
APPENDIX A.	10
Step 1: Identifying Data Sources.....	10
Step 2: Identify Datasets.....	11
Step 3: Dataset Inventory	11
Step 4: Dataset Prioritization.....	12
Step 5: Reviewing Data for Completeness and Accuracy.....	13
Step 6: Publishing Data	13
Step 7: Adding Metadata	14
Step 8: Publication.....	16
Step 9: Open Data Steering Committee Review	18
APPENDIX B.	20
Step 1: Identifying Datasets for Archiving.....	20
Step 2: Request to Archive.....	20
Step 3: ODSC Review	20
Step 4: Archived Data Storage	20
Step 5: Permanent Removal of Datasets	20
APPENDIX C.	21
GLOSSARY	21

City of Las Vegas Open Data Guide

Background

The City of Las Vegas's open data initiative got its start in 2013 when the Code for America Fellowship in emphasized the importance of open data. The City introduced its Open Data portal when it first participated in National Day of Civic Hacking the same year.

As part of What Works Cities (WWC) engagement in 2016, City Council issued a resolution promoting open data <https://opendata.lasvegasnevada.gov/dataset/Open-Data-Resolution/t246-yhfk>. Also, as part of WWC, the City's Open Data policy was re-issued the same year (the original was issued in 2014) <https://opendata.lasvegasnevada.gov/dataset/Open-Data-Policy/fzgv-7d28/data>. The policy drives increased government efficiency and civic engagement, leading to social and economic benefits such as:

- Empowering citizens through democratization of information
- Fostering citizen participation
- Supporting early-stage entrepreneurship
- Contributing to workforce development and job creation
- Increasing a positive business environment
- Promoting public-private partnerships

The policy directs all departments to make their data open, which both sets the expectation that public data will be public and makes clear that the City has a responsibility to protect privacy. Under the policy, each department works with IT to publish its data. Each department Director appoints an Open Data Coordinator (ODC) to lead the department's efforts, and this guide is intended primarily for those ODCs. Any procedures put in place per the Open Data policy, will be documented and included as an appendix to this guide.

A question staff often ask is, "*Why should my department make its data available to the public?*" There are a number of reasons why releasing open data can benefit your department and those it serves. Opening up data to the public is a priority to City Council as evidenced by the aforementioned resolution and policy. Releasing data that has been collected and/or maintained through the operation of City departments enhances the transparency of City government, and can help citizens become better informed about the operation of the City and more engaged with their elected representatives.

Releasing open data has many practical benefits for city departments. Open data releases can be an effective way of responding to requests for data through the Nevada Public Records Act. One Open Data release may address multiple requests for information that can be repetitive and costly to respond to if addressed on an individual basis.

City departments may help to stimulate new and innovative ideas from our local technology community. There is great potential for Open Data to act as the fuel for new solutions and even new businesses that can address common problems or challenges facing those that live in, work in or visit the City.

Open Data Portal

The City's Open Data portal is the central location for published City data. As of 2018, the portal now includes 118 datasets (aka Data Catalog) which provide information on publicly regulated activities that affect the community, plus Open Checkbook and Open Budget. The IT department, specifically the IT Data Architect, is responsible for operation of the portal by managing the following activities:

- Publish datasets with reasonable, user-friendly registration requirements, license key requirements, or restrictions on the use and distribution of dataset.

City of Las Vegas Open Data Guide

- Ensure that updated datasets retain the original data structure, i.e., the number of data elements per record, name, formats, and order of the data elements must be structurally consistent with the originally approved submission.
- Use open, non-proprietary standards when feasible.
- Include an on-line forum to solicit feedback from the public and to encourage public discussion on Open Data policies and public dataset availability.
- Forward open data requests to the assigned ODC.
- Take measures to ensure access to public datasets while protecting the Data Catalog from unlawful abuse or attempts to damage or impair use of the website.
- Make department data available according to their customer needs. For example, high-volume datasets of interest to developers should be released using bulk downloads as well as Application Programming Interfaces (APIs).
- Maintain an Open Data product and licensing to ensure that if data are made public there are no restrictions on copying, publishing, distributing, transmitting, adapting, or otherwise using the information for non-commercial or for commercial purposes.
- Ensure datasets, and any associated metadata, that are outdated, no longer utilized or updated are properly archived while continuing to be available to the public.

The City published datasets are placed in the public domain with no restrictions or requirements placed on use of these datasets. The City assumes no liability for errors, omissions or inaccuracies in the public data. Data published on the Open Data portal is subject to terms of use developed by the IT Department posted conspicuously on the portal. The City may discontinue availability of content on the Open Data Portal at any time and for any reason.

City of Las Vegas Open Data Guide

Purpose

The purpose of this guide is to outline the work being done to implement the City's policy, as well as provide guidance to City, Open Data Coordinators. Open Data Coordinators should use this data guide to assist in their roles. That policy directs City staff to develop and implement practices that allow the City to proactively release all publishable City data, to publish it with documentation, and to encourage use of the data by City departments, outside entities and the public. The policy also directs City staff to minimize limitations on the disclosure of public information while appropriately safeguarding protected and sensitive information. This guide, which provides practical instructions, is a work in progress and will be updated on a regular basis

GOVERNANCE

The following is intended to define roles to help the ODCs and others when carrying out their responsibilities.

Open Data Steering Committee

Development and implementation of best practices are overseen by the Open Data Steering Committee (ODSC). The ODSC provides direction to all City departments, offices, administrative units, boards, commissions, and committees on Open Data policy. The ODSC is comprised of persons designated by the City Manager from internal City departments and external stakeholders, who will work with the City's departments to:

- Identify a lead open data coordinator for each City department who will work with the Information Technologies Data Architect and will be responsible for managing that department's participation in the Open Data Initiative;
- Oversee the creation of a comprehensive inventory of datasets held by each City department which is published to the Open Data Portal and regularly updated;
- Develop technical requirements for publishing and archiving public datasets by departments;
- Develop and implement a process for determining the relative level of risk and public benefit associated with potentially sensitive, non-protected information so as to make a determination about whether and how to publish it;
- Develop and implement a process for prioritizing the release of datasets to the Open Data Portal which takes into account new and existing signals of interest from the public (such as the frequency of public records requests), the City's programmatic priorities, existing opportunities for data use in the public interest, and cost;
- Establish processes for publishing datasets to the Open Data Portal, including processes for ensuring that datasets are reviewed for use-appropriate formats, quality, timeliness, and exclusion of protected and sensitive information;
- Develop and oversee a routinely updated, public timeline for new dataset publication; and
- Ensure that published datasets are available for download, and archived datasets continue to be available for download.

The ODSC submits an annual Open Data Report to the City Council in March of each year. This report includes the following:

- Assessment of progress towards achievement of stated goals

City of Las Vegas Open Data Guide

- Assessment of the current scope of department compliance
- List of datasets currently available on the Open Data Portal
- Timeline for datasets expected to be published in the coming year

In the report, the ODSC may suggest ways to improve the City's Open Data management processes to ensure movement towards the achievement of the policy's goals.

Open Data Program Manager

The Information Technologies (IT) Department is led by the IT Director who provides leadership by liaising with the City Manager, the Mayor's Office and City Council as necessary. The IT Director designates an Open Data Program Manager (ODPM) to oversee the process for inventorying and publishing datasets, including setting timelines for publication. The ODPM is responsible for ensuring publication of open data is consistent with the Open Data policy, including privacy protections. The ODPM will develop a process for determining the relative level of risk and public benefit associated with potentially sensitive, non-protected information so as to make a determination about whether and how to publish it (See Appendix). The ODPM will establish a process for prioritizing the release of datasets to the Open Data Portal which takes into account new and existing signals of interest from the public, the City's programmatic priorities, existing opportunities for data use in the public interest, and cost (See Appendix). The ODPM will maintain a timeline for the public announcing new dataset publication. The ODPM is responsible for upkeep of this guide, which details actions necessary to implement the Open Data policy and standards of when datasets are appropriate for public disclosure.

IT Data Architect

The IT Data Architect is responsible for day-to-day management of the open data program, including process management and quality control for all datasets published to the Open Data portal, as well as the open dataset inventory. The Data Architect will ensure data is in the correct schema for automating upload process; application and use of tools and standards on Open Data platform; builds datasets on the Open Data portal; develops technologies required to capture data from third-party sources; properly archives datasets and all corresponding metadata; and implement privacy, data licensing, metadata and other standards and practices. The Data Architect will follow the direction set by the ODSC to implement the Open Data policy, and actively seek feedback from consumers or users on published datasets and their use/usefulness.

IT Business Partners

The IT Business Partners are responsible for assisting with creation of an inventory of datasets held by each City department. The Business Partner will follow processes put in place by the ODPM for publishing datasets to the Open Data Portal, including processes for ensuring that datasets are reviewed for use-appropriate formats, quality, timeliness, coordinating dataset publication, and exclusion of protected and sensitive information. The Business Partner also communicates data releases to involved and/or impacted departments.

Department Directors

Department Directors guide their department's participation in the Open Data initiative, setting open data-related performance goals for their teams, updating policies and procedures to reflect the open data policy, and making sure their staff have the time and resources to participate in the program. As needed, they set priorities, engage stakeholders, and work through any sensitivities in their datasets prior to publication. They consult with department staff at the start of new projects and software implementation to ensure that they facilitate compliance with the City's Open Data policy.

City of Las Vegas Open Data Guide

Open Data Coordinators

Each department's Open Data Coordinator (ODC) is responsible for managing their department's participation in the initiative. The ODC works with the department Director to set priorities, coordinate publication, keep the open dataset inventory up-to-date, request archiving of datasets, and contribute to annual reports. The ODC works in partnership with their IT Business Partner to inventory department datasets; to establish a plan and timeline for publishing datasets; to implement privacy, data licensing, metadata and other standards and practices; and to provide reports on progress in implementing the open data plan. The ODC is the point of contact in a department for the IT Data Architect, and when called upon, answers questions from the ODSC about datasets.

The ODC for each department is responsible for carrying out the following actions:

- Oversee implementation and compliance with the Open Data policy.
- Inventorying datasets and data sources
- Working with ODPM to prioritize datasets for publication
- Regularly review progress on providing access to datasets requested by the public.
- Make reasonable efforts to make available *all* datasets under the department's control.
- Follow a process to engage with customers to help facilitate and prioritize data release.
- Receive from the Data Architect, public feedback on the department's published datasets.
- Respond to information requests from the public related to a dataset through the ODPM.
- Notify the IT Data Architect prior to any structural changes to datasets when updating data.
- Request the archiving of datasets.
- Assist the Open Data Architect with the archiving approval process (Exhibit B).

The ODC must ensure their department adheres to the following open data standards/requirements with future data collection:

- Use machine-readable and open formats for information as it is collected or created.
- Ensure privacy and confidentiality are fully protected and that data are properly secured.
- Determine if there are any restrictions that would prevent making data publicly available.

Internal data review

Open data has many stakeholders (e.g., those impacted by the data directly or indirectly; those who use or consume the data). At each stage in the process, gauge how stakeholders react to the proposed data, identifying areas of potential confusion or unnecessary limitations on its usefulness before it is published. Direct Stakeholders are people and institutions who are represented directly in the data, whereas Indirect Stakeholders are people and institutions who may be impacted by the release of the data or analysis conducted on it. Users are people and institutions who will use the raw data, Consumers are anyone who reads or uses the information.

An important outreach task to undertake when releasing open data is inside the ODC's own department. Most importantly, make sure the department Director is aware of any planned data release. Some questions the ODCs should ask themselves:

- Have people who can provide insights into the structure of the data been interviewed?
- Have planned data releases been communicated to areas that may be impacted?
- Does the data you plan to release contain personally identifying information?
- Will removing such information diminish the value of the data for consumers?
- Are there any restrictions on how the data may be used?
- Has the City Attorney's Office approved the release of such information?

City of Las Vegas Open Data Guide

Having an internal conversation with others in a department about data that is planned for release is a necessary precursor to external communication with outside data consumers. This dialogue will also inform the terms of use that accompany your data release. It is also a good idea to identify people in a department through these discussions knowledgeable about the data to be released. Identify both administrative contacts (i.e., how often is the data updated, where can it be downloaded, etc.) and technical contacts (i.e., what does this field in your dataset mean?) for the data, in the event that the ODC will not fill these roles.

Sensitive Information and Personally Identifiable Information

Public information or publishable information can be or currently is released to the public. It does not need protection from unauthorized disclosure, but does need integrity and availability protection controls. This would include general public information, published reference documents (within copyright restrictions), open source materials, approved promotional information and press releases.

Sensitive Information is information that could potentially impact the privacy or welfare of an individual, the security of a business or organization, and trade-secrets of a business. Additionally, at the personal level, several acts related to internet fraud, identity theft, and credit fraud are precipitated by misuses of sensitive information. This information is often released in an aggregate format, reducing the ability to trace or link information to a specified individual.

Sensitive information may not be specifically protected from disclosure by law and is for official use only. It is generally not released to the public unless specifically requested. Although most all of this information is subject to disclosure laws because of the City's status as a public entity, it still requires careful management and protection to ensure the integrity and obligations of the City's business operations and compliance requirements. It also includes data associated with internal email systems and City User account activity information.

- *High* information is considered sensitive by stakeholders, leadership; including Council Members and the Mayor's Office, such that it would likely require outreach to those groups prior to publication.
- *Medium* information has an impact on stakeholders that should be taken into account, but is unlikely to be disruptive to ongoing processes.
- *Low* information is already public in some form or does not contain data that would be surprising to stakeholders.

Some confidential information is specifically protected from disclosure by law and subject to strict handling requirements dictated by statutes, regulations, or legal agreements. Serious consequences could arise from unauthorized disclosure, such as threats to critical infrastructure, increased systems vulnerability and health and safety, or legal sanctions. Departments handling this category of information must demonstrate compliance with applicable statutes, regulatory requirements and legal agreements.

City of Las Vegas Open Data Guide

Personally identifiable information (PII) refers to information that can be used to distinguish or trace an individual's identity, either alone or when combined with other personal or identifying information that is linked or linkable to a specific individual, like sensitive information. Because PII is not anchored to any single category of information or technology, it requires a case-by-case assessment of the specific risk that an individual can be identified. Before disclosing potential PII or other sensitive information, all Departments must recognize that non-PII can become PII whenever additional information is made publicly available (in any medium and from any source) and is combined with other available information. The data created through this mosaic effect could allow for the identification of an individual or present additional security concerns. As departments consider whether or not information may be disclosed, they must also account for the "mosaic effect" of data aggregation. The mosaic effect occurs when the information in an individual dataset, in isolation, may not pose a risk of identifying an individual (or threatening some other important interest such as security), but when combined with other available information, could pose such risk.

PII may include the following:

- Name and initials in any combination
- Identification number (e.g., Social Security #)
- Birthdate
- Age
- Gender
- Home address
- Home telephone number
- Personal cellular, mobile or wireless number
- Personal e-mail address
- Drivers' license number
- Information on medical or health conditions
- Financial information (credit card/PCI, billing info, account info)
- Health information
- Marital status
- Nationality
- Sexual behavior or sexual preference
- Physical characteristics
- Racial or ethnic origin
- Religious, philosophical or political beliefs
- Trade union membership
- Biometric data
- Household information
- Consumer purchase or billing history
- Unique device identifiers (IP/MAC addresses)
- Location information (GPS)

Community Engagement

Publishing an open dataset for outside consumers isn't the end of the process - it's just the beginning. Las Vegas has a broad and vibrant community of different users interested in working with City data. The members of this community, and all of the smaller communities, cliques and collectives that it is made up of, are the city's most important asset for turning data into value. Reaching out to and actively engaging with this community is a process that will take place long after your data is initially released.

There are many people and institutions who are impacted by the data in different ways. The more you can map them out as you think about your datasets, the easier it will be to make decisions about what to include (or what not to), what format to provide the data in, and what activities to undertake around

City of Las Vegas Open Data Guide

its release to ensure the data is well understood. All this should be viewed through an equity lens. The Open Data initiative provides outreach opportunities and can help to connect with users of data as well as to support efforts to engage those impacted by the data in the process, from publication to promotion, but subject matter expertise is necessary to consider the possible impacts and start brainstorming about who needs to be engaged.

The City has participated in national efforts to release more and better data, such as the What Works Cities initiative. The City also collaborates with regional jurisdictions, such as the county and state, and share best practices. These connections allow the City to share and adopt the best approaches, creating a community of practice that makes all this work easier in the long run. The ODSC will work to facilitate better communication between the City and the technology community and participate in, or host, events such as hackathons.

Communications

The ODSC and the Communications office will develop a campaign that involves the following elements:

- Links to Open Data Portal across city internet site, where appropriate
- Use embed codes to insert captivating open data visualizations on the Open Data/Transparency Page
- Include links to Open Data Portal in blog posts where appropriate
- Post information on social media when new datasets are added
- Use hashtag #opendata to increase reach on social media
- Create a regular feature on social media to share unique or popular data
- Post interesting data visualizations on existing government social media accounts
- Create a distribution list and send regular open data updates
- Press releases/op-eds connecting open data to some important initiatives
- Encourage stakeholders to use our open data in their meetings
- Consider gamification techniques to incentivize open data posts by city departments
- KCLV videos of datasets

Evaluating Reach

Over the next few years, the Open Data Initiative will focus on quality over quantity, prioritize datasets based on the public interest, improve governance of our open data through improved metadata and regular updates to datasets, and engage our users proactively throughout the process. We want to publish meaningful and usable data, maximizing the benefit to the public of the resources spent to manage the program.

As the initiative progresses, older, outdated, and underutilized datasets will be archived to make room for newer and more refined datasets to be published.

Measures

The open data platform provides performance data about the published datasets, including: dataset name, category, date created, date last updated, number of views, number of downloads, and number of rows.

The ODC will evaluate reports, provided by IT, outlining the trends and analysis of the open data datasets. The data presented during the ODC meetings will allow the committee to determine dataset feasibility and consideration of the following actions: replacement, removal, or general reevaluation or revision.

City of Las Vegas Open Data Guide

Datasets categorized to highlight:

- Top 5 Datasets Most Viewed
- Top 5 Datasets Most Downloaded
- Top 5 Datasets Least Viewed
- Top 5 Datasets Least Downloaded
- Top 5 Datasets Oldest “Date Last Updated”

Qualitative data will also be provided based on feedback from the Explorer Applications, which are applications developed by external users utilizing open data. Additionally, any feedback posted on the open data portal, or gathered by data driven initiatives sponsored by the city, local educational institutions, and others (i.e. hackathons, civic events, and tradeshow).

APPENDIX A.

The following is a process to follow when publishing department datasets on the Open Data portal:

Step 1: Identifying Data Sources

Data may be housed in a variety of places including on the systems of outside vendors. To discover and identify major data sources, the ODCs ask the following questions:

- What information systems does the department use?
- What databases does the department use?
- What applications capture information or are used in department business processes?
- Where does the data reside?
- What information is the department already publishing; where did that information come from?

For each data source, the following information is entered in the [Data Sources Template](#):

- Name of data source
- Data source description
- Name of data administrator
- Date range included in data source
- Comments

Departments will repeat this process as new datasets are created, discovered or requested.

City of Las Vegas Open Data Guide

Step 2: Identify Datasets

There are a number of easy ways to identify data that is suitable for release. Review existing requests for data and information such as formal Open Record requests or less formal requests sent via e-mail or some other channel. Check the pipeline of datasets being prepared for publication (those associated with Results Vegas, for example). Look at what other cities have published. Ask the public for ideas. Look at the City's website for documents with data such as annual reports, financial statements, office locations, etc. Any web-based application that has a database behind it, like a search feature, is worth examining more closely for a dataset that might be appropriate to release. And check for "scrapers" which are programs written by people who want to extract information, usually in bulk, from a website.

Every dataset will be reviewed and approved before it's published. At this point, the ODC does not exclude any datasets based on privacy, security, sensitivity or quality concerns. Some of information sources may be fairly simple, while others may be complex. To discover and identify datasets, the ODCs ask the following questions:

- What data populates department monthly or quarterly reports?
- What data does the department use for internal performance and trend analysis?
- What information is used for performance measures?
- What data is reported to federal, state or local agencies?
- What data do other departments ask for or is already shared with other departments?
- What information is sought via the public disclosure process and/or open data requests?
- What departmental data is publicly available on the Open Data Portal or elsewhere online?

An at-a-glance assessment for the inventory is intended to assist with prioritization. It is not a substitute for the in-depth data quality and privacy assessments that are required prior to publication. In the template itself are definitions for each assessment category as well as examples. Most fields are a drop-down.

Step 3: Dataset Inventory

The Open Data program's dataset inventory serves as a way to see what data the City has that might be appropriate for publication as open data under the open data policy. The inventory is a tool to keep track of all the potential open data in the City to best manage the program down to the department level. Each department is different in terms of how it is managed, how much data it owns and what type of data it owns. The ODCs meet with department leadership to share what open data is, what the ODSC's expectations are for the inventory, and how they might best go about soliciting input from data owners. The ODCs may consult with IT for guidance and support. The ODCs do not exclude any datasets based on privacy, security, sensitivity or quality concerns. Every dataset is reviewed for such elements and approved by the ODSC before it is published. To determine data suitable for release, the ODCs may also work with Communications to use site analytics on the City's Internet site and to conduct online surveys.

For each dataset identified, the ODC enters the following information on the [Dataset Inventory Template](#):

- Name of dataset: What is this dataset called?
- Dataset description: Provide a brief description of the contents of the dataset. What is its purpose? What kinds of information does it contain?
- Data Source: Does this dataset come from a larger database or system? If so, what system or database?
- Data Administrator: Who maintains this dataset? Who should be contacted with questions related to this dataset?
- Data Start Date: first date this dataset is available

City of Las Vegas Open Data Guide

- Frequency of data change: How often is this data updated? (e.g. continuous, hourly, monthly, annually)
- Departments/Divisions using this data: Which departments, or divisions/workgroups within the department use this dataset?
- Data Quality: Are there any concerns with regard to accuracy, completeness, or consistent entry? Any limitations to this dataset?
- Demand: Do you receive requests for this data or is it used for one of your performance measures? Rate the demand as low, medium or high.
- Data Classification: Classify the data as protected, sensitive, or public.
- Protected Details: Provide details regarding what information is protected.
- Sensitive Details: Provide details regarding what information is sensitive
- Priority Level: Categorize the data as priority 1, 2, 3, or 4. See instructions in Step 4 below.
- Comments: Do you have any other comments or ideas for how the make the data useful?

Step 4: Dataset Prioritization

Unlike the City's comprehensive data inventory, which includes all data collected and stored by the City, the Open Data initiative's dataset inventory is focused on existing datasets that are already in use. This is a first step that may or may not lead to publication of the dataset, in full or in part. The Open Data Initiative's dataset inventory serves as:

- A way to see, at a glance, what data the City has that might be appropriate for publication as open data under the new open data policy
- A tool to keep track of all the potential open data in the City to report to the City Manager and out to the public
- A tool for ODCs to keep track of their department's open data and make sure it is managed well on an ongoing basis

Each department is different in terms of how it is managed, how much data it owns, what *kind* of data it owns. Depending on what works for the ODC, they may start with any or all of these steps:

- Meet with the department Director and managers to share what open data is, what the ODSC's expectations are for the inventory, and how to solicit input from data owners
- Meet with the department's person responsible for performance management reporting
- Meet directly with data owners to talk to them about what they own

As the ODCs assign a priority level to each dataset, they take into account the following factors:

- The data is related to one of the department's measures for Results Vegas, the City's Strategic Priorities or has been identified as a priority.
- Publishing the data would make it easier for the City to comply with existing reporting requirements.
- Existing reports already automate the data extractions making it easy to add to the Open Data portal.
- It is easy to publish this dataset because a related dataset is already being published.
- Migration to a new backend system would create additional work by automating the publication twice.
- There are major data quality concerns, making the data unusable.
- Data is not available in a structured manner. (e.g., not in a database or well-designed spreadsheet).

Each priority level is discussed in more detail below:

Priority 1	Priority 2
Public datasets that are part of the new performance measures or are in high demand	Datasets range from high to low demand and may have contain private and sensitive information

City of Las Vegas Open Data Guide

Priority 3	Priority 4
These datasets will be scheduled after all Priority 1 and 2 data has been scheduled.	These datasets do not need to be scheduled until all other datasets are scheduled for publication.

Assigning Priority to Datasets

The ODCs use the criteria for demand for the dataset (Low, Medium, High) and for classification (protected, sensitive, public) to determine the priority for publication based on the following matrix:

Classification	Public	P2	P2	P1
	Sensitive	P3	P2	P2
	Protected	P4	P3	P2
		Low	Medium	High
		Demand		

Step 5: Reviewing Data for Completeness and Accuracy

Every public dataset varies in completeness and quality, and prior to releasing data for outside users in every department should strive to ensure their data is as accurate, complete and up to date as possible. However, there is no such thing as “perfect” data and using data perfection as a prerequisite for releasing data can become an impediment. The following are definitions of data quality:

- *Good* This dataset is used routinely, is managed in a way that makes it legible to users besides the owner, and is reasonably accurate.
- *Acceptable* This dataset is used routinely but not necessarily legible to non-owners and/or has some gaps or discrepancies that limit its usability.
- *Poor* This dataset exists but is not regularly used, updated, or otherwise managed in a way that makes it valuable.

When releasing datasets, be explicit about any limitations that were encountered in preparing it for release and add caveats that will help data consumers understand the limitations of your data if any exist. If data is subject to revision, if portions have been redacted, if it only covers a limited time period - make sure to clearly state these limitations as part of your data release. Clearly stating limitations and caveats will make your data more usable because consumers will have a more thorough understanding of what the data represents. It is also worth noting that most communities of data consumers will provide feedback on the quality of data or any perceived inaccuracies.

Step 6: Publishing Data

After departments have submitted their data inventories, the next step is to develop a plan and timeline for publishing datasets. Publishing each dataset requires work to standardize the data, make it machine readable, write good metadata, and take steps to protect privacy and mitigate security risks.

City of Las Vegas Open Data Guide

Each department will need to prioritize which datasets to publish based on a variety of factors, including input from the public.

The ODC works with the person or people who manage the data and/or work related to it, who understand what it is used for, who can provide the context/metadata that ensures the public understands what it means, and who can lead stakeholder engagement. The ODC also works with their department's IT Business Partner and the IT Data Architect to help extract datasets from databases and other sources, prepare the data for publication and ultimately publish the data on the Open Data portal.

Step 7: Adding Metadata

If you don't tell recipients of your data what it is, how it can be used, or anything about its timeliness, how can they be expected to use it effectively? Metadata, or descriptive information about data, is common among geospatial data (i.e., GIS layers) and other datasets due to their complexity. Typically included in geospatial metadata is the following content: a descriptive summary file and its primary uses; methods of creation; changes (if the file is routinely updated); contact information; geographic references, and most critically - dates indicating when the data was captured and/or prepared.

To submit a dataset for publication, the metadata for each dataset must be detailed. This information will be accessible to the public alongside the dataset itself. This is where you will want to put any information that does not belong inside the dataset but is important for understanding what it is and how to use it. Relevant links may be included, as long as you take care to keep them up to date.

The basic tenants of metadata (what the data is, how to interpret the data for use, how the data was prepared) qualify for inclusion within all digital databases be they geographic or otherwise. How metadata is included in a data release may vary from format to format. The ODC and IT prefer csv as the standard.

Basic content should include:

- Data description and common usage;
- Department Owner;
- Source of Data
- Results Vegas Measures
- Descriptions of key fields and/or field codes, and;
- Limitations, disclaimer (if appropriate) and terms of use.
- Data Classifications and Standards
 - Publisher
 - Technical Contact (Developer, DBA, CDO)
 - System of Record
 - Data Collection Method (API, URL, CSV)
 - Business Rules (CLV, Other Agency, Ad-hoc)
 - Data Source Location (On-Premise, Cloud, Hybrid)
 - Data Refresh Frequency

City of Las Vegas Open Data Guide

For each column in a Dataset, the ODC will document:

- Column (field name)
- Data type (text, date/time, geocode, etc.)
- Sample value
- Column metadata description (what the data in the column represents or means)

The ODC will provide a one-record example of dataset with all fields populated (i.e., field name, data type, value, metadata description). The ODC will also include a one-record example of dataset with all fields populated. Here is an example:

Field Name	Data Type	Value	Metadata Description
EQUIP_ID	Number	1404	Vehicle equipment identification number
YEAR	Text	2000	Vehicle manufacture year
MAKE	Text	PLYMOUTH	Vehicle make
MODEL	Text	NEON	Vehicle model
DESCRIPTION	Text	SUBCOMPACT SEDAN	Type of vehicle
DEPT	Text	SPD	City department
SALE_PRICE	Number	3700	Vehicle sale price
SOLD_BY	Text	BIDADOO ONLINE	Sales agent
DATE	Date	CCYYMMDD	Sales Date

- **Metadata for Datasets on the Open Data Portal**

A catalog of datasets is available on the Open Data Portal. For each dataset, there is a menu called “About” which includes the following information:

- Dataset Title: It is brief, but descriptive.
- Brief Description: Begins with “See ‘About’ for attachment with field definitions,” so users know where to find the field definitions. It should also include a reference to the master dataset if the table is a summary of a larger dataset.
- Category: Results Vegas datasets are categorized according to the dashboard they are associated with.
- Tags/Keywords: This section includes different terms that a user may utilize to search for the dataset.
- Row Label: Not used
- Licensing Type: Not used
- Data provided by: Source, such as department, of the data provided.
- Source Link: This is the link to the entity that provided the data.
- Semantics and RDF: Not used
- API Endpoint: Not used
- Thumbnail Image- Not used
- Attachments: Each dataset has a .csv attachment (Step 7) with a list of field names and a description for each field name. It includes the name of the department, the name of the indicator, the name of the dataset, and a list of fields and associated descriptions. If a field can have specific values or categories, those categories are defined.
- Contact information: Email opendata@lasvegasnevada.gov

Each dataset has an owner, the ODC for the department if it is in-house, or a contact at the external agency if the dataset comes from outside of the City. A master database with information on all datasets (both Results Vegas and non-Results Vegas datasets) will be maintained by the Data Architect, which include the following:

City of Las Vegas Open Data Guide

- Data owner (ODC)
 - Data source
 - Frequency the dataset is refreshed/updated
 - Whether the update of the dataset is automated or not
 - Who is responsible for checking to see if the data is coming in correctly and how often that should occur
 - .CSV file with field definitions for every dataset
-
- **Metadata for charts (Open data Platform Requirement Only)**
Description field includes a reference back to the full dataset because many charts are filtered views of the data and users need to know where they can find the full dataset.
-
- **Metadata for maps**
 - The description field includes a link to the map service.
 - The sections titled basic and Drawing Info were self-populated.
 - The attachment (Step 7) includes a list of fields and field descriptions.
 - For each map, clicking on the map allows the user to see the data associated with that point. The field descriptions align to the fields displayed on the map.

Step 8: Publication

There are three methods for publishing datasets to the Open Data portal:

- *Manual* is generally used for datasets that are relatively small and uncomplicated, and where the refresh rate is quarterly or less infrequent.
- *Automated* is recommended when datasets will be updated more frequently than quarterly. (Evaluated on a case-by-case basis, depending on source system capabilities)
- *Programmatic (API to API)* is recommended for real- or near-real-time data transfers (Preferred Method).

For any of the methods above, the following steps are required:

- *Extract* export data from source system, and since each is different, the methods will vary.
- *Transform* if necessary, modify the exported data so the data structure matches the data schema.
- *Upload* the data into the Open Data portal using one of the publication methods described above.
- *Privately* publish the dataset initially (password-protected) for stakeholders to review using the open data platform sharing tool.
- *Open* the data to the public once the data is deemed ready for publication by the ODSC.

Numerous efforts are underway to create and implement enterprise data standards, allowing for comparing data and interoperability of applications built for cities. The ODSC will maintain an awareness of standards with the potential improve the potential for useful analysis and drive down costs of solutions. Open data is a global movement that is helping governments improve their operations, better communicate with the public, power analysis that makes cities “smarter,” and stimulate the development of civic technology tools that improve people’s quality of life.

An important step in releasing open data is determining how it will be made available. To some extent, this decision will be a function of what kind of data you want to release and the format it is currently in (i.e., how it is used or maintained internally). Every data release should be accompanied by a clear statement providing terms of use to end users. If your data release is subject to special restrictions or caveats, the terms of use should state these very clearly for users.

City of Las Vegas Open Data Guide

One of the most common types of city data is *geospatial data*, or data that relates to a specific place. This may be things like the locations of points of service, like a library or police station, or of physical assets like fire hydrants or street lights. It is common for city data to have a geospatial component - the location of a pothole that is reported. The easiest and most efficient way to make geospatial data available to outside consumers is to work through the GIS team in the IT department.

If department data changes somewhat infrequently, or will be accessed only periodically it may be appropriate to release such data as a static download in a commonly used format (e.g., comma separated value). However, if your data is updated frequently or subject to changes - or if it will be targeted for specific user communities like software developers - it may make sense to consider provisioning an Application Programming Interface (API) for the data.

All Open data published by the City will be consistent with the following principles:

- **Public** - City departments must adopt a presumption in favor of openness to the extent permitted by law and subject to privacy, confidentiality, security, or other valid restrictions.
- **Accessible** - Open data are made available in convenient, modifiable, and open formats that can be retrieved, downloaded, indexed, and searched. Formats should be machine-readable. To the extent permitted by law, these formats should be non-proprietary, publicly available, and no restrictions should be placed upon their use.
- **Described** - Open data are described fully so that consumers of the data have sufficient information to understand their strengths, weaknesses, analytical limitations, security requirements, as well as how to process them.
- **Reusable** - Open data are made available under an open license that places no restrictions on their use.
- **Complete** - Open data are published in primary forms with the finest possible level of granularity that is practicable and permitted by law and other requirements. Derived or aggregate open data should also be published but must reference the primary data.
- **Timely** - Open data are made available as quickly as necessary to preserve the value of the data. Frequency of release should account for key audiences and downstream needs.
- **Managed Post-Release** - A point of contact must be designated to assist with data use and to respond to complaints about adherence to these open data requirements.

Although the City strives to ensure that such public data are accurate, the City makes no representation or warranty of any type as to the content, accuracy, timeliness, completeness or fitness for any particular purpose or use of any public data provided on such portal; nor shall any such warranty be implied, including, without limitation, the implied warranties of merchantability and fitness for a particular purpose. The City shall assume no liability for the following:

- Errors, omissions or inaccuracies in the public data provided on the CLV Data Catalog regardless of how caused.
- Decisions made or actions taken or not taken by anyone using or relying upon such public data.
- Any virus or other damage to any computer that might occur during or as a result of accessing such portal or the public data provided therein.

The CLV Data Catalog and all public data contained on such catalog shall be subject to Terms of Use developed by IT. Such Terms of Use are posted by IT in a conspicuous place on the CLV Data Catalog. The City reserves the right to discontinue availability of content on the CLV Data Catalog at any time and for any reason. If a dataset is made accessible by a city department on the CLV Data Catalog and such city department is notified or otherwise learns that any dataset or portion thereof posted on the data portal is factually inaccurate or misleading or is protected data, the city department shall, as appropriate, promptly correct or remove, or cause to be corrected or removed, such data from the data portal and shall so inform the chairperson of the designated DC.

City of Las Vegas Open Data Guide

Step 9: Open Data Steering Committee Review

The ODSC will review each dataset for privacy, security, and quality prior to publication. If there is a significant risk of privacy harm or if they notice security issues, they will work with the ODC to reduce the risk of harms prior to publication. If security issues are identified; the risk of harm will be mitigated by the IT Security Team prior to publication. The ODSC will flag any quality concerns so they can either be addressed prior to publication or described in the metadata, so that users are aware of the dataset's limitations. The following matrix provides a guideline for ODSC members when evaluating a dataset for publication:

CRITERIA	DOES NOT MEET	MEETS	EXCEEDS
Open Data Standard	Data is not available online, closed format, legal encumbrances on use and reuse	Data is available online, in open format, no legal encumbrances on use or reuse	Data is available in multiple formats in addition to the meets qualifications criteria
Open Format Standard	Data presented in a propriety, non-standard format, format is not machine readable, processes are manual and non-automatable	Data presented in widely accepted, nonproprietary, platform independent, machine-readable method for formatting data and process automatable	Data is accessible through developed APIs, web services, and other advanced technologies in addition to meets qualifications criteria
Protected Information	A department allows access to any dataset or portion, contrary to state law or city ordinance, or any other law or rule or regulation	Any dataset and all portions that a department may <u>deny</u> access pursuant to state law or city ordinance, or any other law or rule or regulation	Not applicable
Sensitive Information	Data published on the open data portal that could raise privacy, confidentiality or security concerns or have the potential to jeopardize public health, safety, or welfare greater than the benefit of publication	Data is <u>not</u> published on the open data portal that could raise privacy, confidentiality or security concerns or have the potential to jeopardize public health, safety, or welfare greater than the benefit of publication	Not applicable
Security	CLV information technology assets, both electronic information systems	CLV Information Technology Security Team validates the protection of	In addition to meets qualification criteria, ensuring CLV Open Data Portal is not

City of Las Vegas Open Data Guide

	infrastructures, and open data portal are exposed, abused, damaged or impaired	information technology assets, encompassing both electronic information systems and infrastructures	abused, damaged, or impaired
Data Quality	Data published is low quality, no metadata, unidentified source system, unsustainable	Data published is high quality, contains metadata, references source system, and sustainable	Data published is superior quality, contains metadata, reference source system, sustainable, data dictionary, entity relationship diagram
Automation/Timeliness	Source systems, file transfer mechanisms, and technologies are outdated without automation options, jeopardizing the timing of portal updates	Source systems, file transfer mechanisms, and technologies are modern or upgradable with automation options, increasing the timeliness of portal updates	Source systems, file transfer mechanisms, and technologies are fully automated using the latest technologies providing optimal timeliness of portal updates
Financial Impact	Data published on the open data portal creates an unintended fiscal impact for CLV	Data published on the open data portal does <u>not</u> create an unintended fiscal impact for CLV	Not applicable

Once the dataset has been published, each ODC is responsible for keeping their department's data up-to-date, for handling questions about the dataset from the public, working with the ODSC to resolve any issues that emerge, and submitting requests for archiving as needed. At least once a quarter, the ODC should review all the department's published datasets to ensure they are being updated on the basis promised in the metadata. This is true for datasets updated automatically as well as those that require a manual update. The dataset inventory acts as a tool to keep track of what has been published to verify that the data are up-to-date.

City of Las Vegas Open Data Guide

APPENDIX B.

Data archiving is essential to the structural integrity and effectiveness of the open data platform. Historical data can aid in identifying information on trends, project successes, revenue, expenditures and other key performance indicators, leading to future data-driven decisions. The following process outlines the steps to follow when archiving department datasets on the open data portal:

Step 1: Identifying Datasets for Archiving

There are numerous ways to identify datasets capable of being archived, including datasets:

- that have a low visit/download performance rate.
- that are redundant; other datasets contain the same information.
- that are outdated (over 365 days without update).
- that will no longer be updated.
- where data sources or systems are no longer available or has been migrated, upgraded, changed, etc.
- that were operational but are no longer in use.
- other issues, not listed above, expressed by the ODSC, ODC, Directors, and/or community stakeholders.

Datasets that are utilized by other city applications and/or identified platforms or software will be excluded from archiving even when any or all of the conditions above are met.

Step 2: Request to Archive

A request to archive can be started once a dataset has been identified. Initial archive requests are to be submitted through the city of Las Vegas service desk. The ODC will receive the request and will begin to assist the requestor in providing the necessary documentation and metadata need for ODSC review.

Step 3: ODSC Review

Completed archiving requests are reviewed by ODSC for approval. Approved requests will proceed to the next step. Denied requests will be returned to requestor with cause and steps for corrective action.

Step 4: Archived Data Storage

All approved archived datasets and related metadata will be stored in a dataset of datasets, stored on the open data portal and accessible to the public.

Step 5: Permanent Removal of Datasets

Permanent removal of data from the open data portal will only take place when legal and/or privacy circumstances are present. Permanent removal of datasets must receive approval from the ODSC and City Clerk's Office.

City of Las Vegas Open Data Guide

APPENDIX C.

GLOSSARY

The following are frequently used terms related to the aforementioned processes:

Data	“Data” means statistical, quantitative, or qualitative information that is regularly maintained, created, or obtained by or on behalf of a City department.
Dataset	“Dataset” means a named collection of related records, with the collection containing data organized or formatted in a specific or prescribed way, often in tabular form.
Department	“Department” refers to all departments, divisions, boards, commissions, agencies, and internal or external organizations responsible for providing and disseminating data.
Open Data	“Open data” means data that is available online, in open format, with no legal encumbrances on use or reuse.
Open Format	“Open format” means any widely accepted, nonproprietary, platform-independent, machine-readable method for formatting data, which permits automated processing of such data and facilitates search capabilities.
Protected Information	“Protected information” means any dataset or portion thereof to which a department may deny access pursuant to state law or City ordinance, or any other law or rule or regulation.
Publishable Data	“Publishable data” means data which is not protected or sensitive and which can be prepared for release to the public.
Sensitive Information	“Sensitive information” means any data which, if published on the Open Data Portal could raise privacy, confidentiality or security concerns or have the potential to jeopardize public health, safety or welfare to an extent that is greater than the potential public benefit of publishing that data.

REFERENCES

City of Raleigh Open Data Catalog, retrieved 5/16/13 from, <https://data.raleighnc.gov/>
City of San Francisco Open Data Policy, retrieved 5/16/13 from, <https://data.sfgov.org/>
Socrata’s Open Data Field Guide, retrieved 5/16/13 from, <http://www.socrata.com/open-data-field-guide-chapter/about/>

City of Las Vegas Open Data Guide

New York City's Open Data Tech Standards, retrieved 5/16/13 from,

http://nycopendata.pediacities.com/wiki/index.php/NYC_Open_Data

City of Chicago Open Data Policy, retrieved 5/16/13 from, <https://data.cityofchicago.org/>

Executive Office of the President of the United States Memorandum on Open Data Policy, retrieved 5/16/13 from, <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>