

chicago manual of style citing geospatial data

The Chicago Manual of Style (CMOS) provides a framework for academic and professional writing, but its application to increasingly complex data types like geospatial information requires careful consideration. As digital mapping, GIS, and spatial analysis become integral to research across disciplines, understanding how to properly cite geospatial data in Chicago style is paramount. This article delves into the nuances of citing various forms of geospatial data, from shapefiles and geodatabases to online map services and satellite imagery, ensuring clarity, accuracy, and adherence to CMOS principles. We will explore the core components of a geospatial citation, address common challenges, and offer practical guidance for researchers navigating this specialized area of academic citation.

Table of Contents

- Understanding Geospatial Data in the Context of Citation
- Key Components of a Geospatial Data Citation in Chicago Style
- Citing Different Types of Geospatial Data
- Challenges and Best Practices for Citing Geospatial Information
- Geospatial Data Sources: Online Maps and Web Services
- Citing Proprietary and Open-Source Geospatial Datasets
- Metadata and its Role in Geospatial Citation

Understanding Geospatial Data in the Context of Citation

Geospatial data, by its nature, is inherently complex, combining locational information with attributes that describe phenomena. When incorporating this data into research, it is essential to attribute its origin and provide sufficient detail for readers to locate and verify it. The Chicago Manual of Style, while not always explicitly addressing every conceivable data format, provides a flexible system that can be adapted. The fundamental goal of any citation is to give credit where it is due and to allow others to access the source material. For geospatial data, this means identifying not only the creator or provider of the data but also its specific nature, scale, and any transformations it may have undergone.

The increasing prevalence of Geographic Information Systems (GIS) and remote sensing has made geospatial data a critical component in fields ranging from environmental science and urban planning to history and social sciences. Properly citing these datasets ensures the rigor of your research and acknowledges the efforts of data producers. Without standardized citation practices for geospatial information, the reproducibility and verifiability

of studies relying on spatial analysis would be severely hampered. Therefore, adapting existing citation principles to encompass the unique characteristics of geospatial data is a necessary evolution in academic discourse.

Key Components of a Geospatial Data Citation in Chicago Style

When constructing a citation for geospatial data, several core elements are generally required, mirroring the principles of citing other forms of media but with specific adaptations. These components ensure that a reader can identify the data, understand its context, and potentially retrieve it. Think of it as building a breadcrumb trail back to the original source. The Chicago Manual of Style emphasizes clarity and completeness, and this approach is particularly vital when dealing with datasets that might be extensive, derived, or updated frequently.

The essential elements for citing geospatial data typically include:

- **Author or Creator:** This could be an individual, an organization, a government agency, or a research group responsible for creating or distributing the data.
- **Title or Description:** A clear and specific title or descriptive name for the dataset. This should be as precise as possible to distinguish it from similar datasets.
- **Publication or Distribution Information:** Details about when and where the data was made available. This might include a date, a repository, or a website.
- **Version or Date:** For datasets that are updated or revised, specifying the version number or the exact date of access is crucial for reproducibility.
- **Access Information:** How the data can be accessed, including URLs, DOI (Digital Object Identifier) if available, or other persistent identifiers.
- **Coordinate System/Projection (Optional but Recommended):** For spatial data, noting the projection system used can be vital for understanding and replicating analyses.
- **Scale or Resolution (Optional but Recommended):** Providing information about the spatial scale or resolution of the data offers context for its applicability.

Author or Creator Identification

Identifying the author or creator of geospatial data is the first step in establishing the origin of the information. This is analogous to naming the author of a book or a journal article. For publicly available datasets, the creator is often a governmental body, a research institution, or a data repository. For instance, the U.S. Census Bureau is a common creator of demographic geospatial data, and the United Nations is often credited with global spatial datasets. It is important to use the most formal and recognized name of the entity responsible for the data.

Title and Descriptive Naming Conventions

The title or descriptive name of the geospatial dataset should be specific enough to avoid ambiguity. If the dataset has an official title, use it. If not, provide a concise and accurate description that reflects the data's content. For example, instead of "U.S. Cities Data," a more precise title might be "United States Cities Geographic Data, 2023." This level of detail helps researchers quickly understand what the dataset represents and if it is relevant to their work. Consistency in naming, especially when citing multiple datasets, is also a key best practice.

Publication, Distribution, and Versioning

Pinpointing when and how the geospatial data was made available is critical. This might involve citing the year of publication, the specific release date, or the date the data was downloaded if it is a dynamic dataset. If the data comes from a specific repository, such as a university library's data archive or a government portal, this information should be included. Furthermore, many geospatial datasets are updated periodically. Citing the specific version or the date of acquisition is essential for ensuring that any reader attempting to replicate your work is using the exact same data. Without versioning, analyses could yield different results due to data updates, compromising the scientific method.

Citing Different Types of Geospatial Data

The vast array of geospatial data formats necessitates a flexible approach to citation. Each type of data may present unique challenges and require specific details to be included in the citation. Whether you are using vector data like shapefiles, raster data such as satellite imagery, or accessing data through web services, the underlying principles of attribution remain the same, but the particulars of how you present that information will vary.

Citing Vector Data (e.g., Shapefiles, GeoJSON)

Vector data, which represents geographic features as points, lines, or polygons, is commonly found in formats like shapefiles (.shp) or GeoJSON. When citing a shapefile, for instance, you would typically identify the creator, the specific name of the shapefile (or the collection it belongs to), the year of creation or release, and the repository or source from which it was obtained. If the shapefile is part of a larger dataset, ensure that the overall dataset is clearly identified. For example, a citation might look like: "World Boundaries, 2022, Esri, World Topographic Map, accessed October 26, 2023, <https://services.arcgisonline.com/...>"

Citing Raster Data (e.g., Satellite Imagery, DEMs)

Raster data, which represents geographic information as a grid of cells (pixels), includes satellite imagery, aerial photographs, and Digital Elevation Models (DEMs). Citing raster data requires specifying the sensor or source of the imagery, the acquisition date (if known), the provider, and any processing or transformations applied. For example, citing Landsat imagery would involve identifying the specific Landsat mission, the sensor, the path/row or geographic extent, and the data distribution center (e.g., USGS EarthExplorer). If you have processed the raw imagery, you should note this in your citation or methodology section.

Citing Geodatabases and Spatial Databases

Geodatabases and other spatial databases, often containing multiple layers of related geospatial information, require citations that reflect the structure and source of the data. If you are citing a specific layer or table within a geodatabase, you should identify the database itself, the specific table or layer name, the owner or creator of the database, and the access date. For proprietary databases, citing the vendor and version is also important. The key is to be as precise as possible about the exact subset of data used from a larger database system.

Challenges and Best Practices for Citing Geospatial Information

Citing geospatial data presents unique challenges due to its dynamic nature, varying formats, and the potential for extensive derivation. Overcoming these challenges requires a commitment to detail and an understanding of the underlying principles of academic integrity. Best practices are essential for

ensuring that your citations are not only compliant but also genuinely useful to your readers.

Some common challenges include:

- **Data Updates and Versioning:** Geospatial datasets are frequently updated, making it crucial to cite the specific version or acquisition date.
- **Data Transformation:** Researchers often modify, reproject, or aggregate data. Citing these transformations is vital for reproducibility.
- **Complex Licensing and Access:** Proprietary data or data with complex licensing agreements require careful attention to usage rights and proper attribution.
- **Lack of Standardized Metadata:** While metadata is critical, it is not always consistently available or comprehensive for all datasets.

Reproducibility and Data Provenance

The cornerstone of good scientific practice is reproducibility. When citing geospatial data, the ultimate goal is to enable another researcher to access and use the same data to arrive at similar conclusions. This means meticulously documenting not only the source data but also any steps taken to prepare or analyze it. Provenance, the record of the origin and lineage of data, is paramount. If you have clipped a dataset, merged it with another, or performed a spatial analysis, these steps should be clearly articulated, ideally in your methodology section, and referenced in your citation if they significantly alter the data's form.

Leveraging Metadata for Robust Citations

Metadata, or data about data, is an invaluable resource for citing geospatial information accurately. Comprehensive metadata often includes information about the data's creator, publication date, spatial resolution, coordinate system, data lineage, and intended use. When available, metadata should be consulted to gather the necessary details for your citation. If metadata is incomplete, you may need to rely on documentation provided by the data distributor or make educated inferences, clearly stating any assumptions made. The more detailed your metadata, the more robust your citation can be.

Geospatial Data Sources: Online Maps and Web Services

The accessibility of geospatial data through online platforms and web services has revolutionized research, but it also introduces new considerations for citation. Many researchers now rely on dynamic map services or online data portals. Citing these sources requires attention to the specific service, the provider, and the date of access.

Citing Web Map Services (WMS, WFS)

When using data accessed via Web Map Services (WMS) or Web Feature Services (WFS), the citation should identify the service provider, the name of the service, and the specific layer or feature class accessed. Crucially, the URL to the service endpoint and the date of access are essential, as these services can be dynamic and their content may change. For example, citing a layer from a government agency's WMS might involve providing the agency name, the service name, the layer title, the service URL, and the access date. The Chicago style manual's principles for citing online resources can be adapted here, emphasizing clarity and retrievability.

Citing Online Data Repositories and Portals

Many organizations maintain online repositories or portals where geospatial data can be downloaded or accessed. These platforms, such as government data portals (e.g., data.gov), university data archives, or specialized research data repositories, act as publishers. When citing data from such a source, identify the repository as the publisher, provide the specific dataset title, and include a persistent identifier like a DOI if available. If no DOI exists, provide a stable URL and the date of access. The original creator of the data should still be credited if it differs from the repository.

Citing Proprietary and Open-Source Geospatial Datasets

The distinction between proprietary and open-source geospatial data influences how you cite it, particularly concerning licensing and commercial use. Both require proper attribution, but proprietary data may also necessitate adherence to specific terms of service or licensing agreements.

Proprietary Data Citations

Proprietary geospatial data, often produced by commercial entities like Esri, DigitalGlobe, or HERE Technologies, usually comes with licensing agreements. When citing such data, it is important to acknowledge the commercial provider as the creator. Include the product name, version number, and the date of acquisition. If the data was obtained through a subscription or specific license, it may be beneficial to note this in your methodology or acknowledgments, although typically not in the formal citation itself unless it pertains to data usability or licensing restrictions being discussed.

Open-Source Data Citations

Open-source geospatial data, which is freely available for use and redistribution (often under licenses like Creative Commons), still requires clear attribution. The creator or steward of the data should be identified, along with the dataset's title and source. Many open-source datasets are hosted on platforms that provide DOIs or unique identifiers, which should be prioritized in the citation for maximum clarity and persistence. Examples include data from organizations like NASA, NOAA, or various open data initiatives.

Metadata and its Role in Geospatial Citation

Metadata is the backbone of robust geospatial data citation. It provides the essential context and descriptive information necessary for others to understand and utilize the data effectively. Without proper metadata, a citation can be incomplete and less useful for future research or verification. Chicago style, in its emphasis on providing all necessary information for a reader to locate a source, implicitly supports the thoroughness that good metadata enables.

Key metadata elements that directly inform citation practices include:

- Data Originator/Creator: Who produced the data.
- Publication Date: When the data was released or last updated.
- Data Title: The official or descriptive name of the dataset.
- Spatial Reference System: The coordinate system and projection used.
- Data Dictionary/Attribute Information: Descriptions of the variables and their meanings.

- **Data Lineage:** Information on how the data was collected, processed, and transformed.
- **Contact Information:** For the data producer or distributor.

When citing geospatial data, always consult any accompanying metadata files (e.g., .xml, .txt). This information is not only crucial for constructing a complete citation but also for understanding the limitations and appropriate use of the data. Properly citing data with its associated metadata demonstrates a high level of scholarly rigor and respect for the data's origin and intended application.

In conclusion, adhering to the principles of the Chicago Manual of Style when citing geospatial data requires a detailed and adaptable approach. By carefully identifying the creator, title, publication details, version, and access information, and by leveraging available metadata, researchers can ensure clarity, accuracy, and reproducibility in their work. This meticulous attention to detail not only upholds academic integrity but also contributes to the broader scientific community's ability to build upon existing research effectively.

FAQ

Q: What is the most crucial piece of information to include when citing geospatial data in Chicago style?

A: The most crucial piece of information is the source of the data, including the creator or responsible entity and a specific identifier for the dataset (like a title or file name), along with how and when it was accessed. This ensures that readers can identify and potentially retrieve the exact data used.

Q: How should I cite satellite imagery that I have downloaded and processed?

A: When citing processed satellite imagery, you should credit the original source of the imagery (e.g., NASA, ESA), mention the specific satellite mission and sensor, the date of acquisition if known, and the data distribution center. In your methodology section or potentially in a note, you should detail the processing steps you undertook, such as atmospheric correction, georeferencing, or mosaicking, as this is part of the data's provenance.

Q: What if the geospatial dataset does not have a clear author or title?

A: If a dataset lacks a formal author or title, you should provide the most descriptive name possible for the dataset and identify the organization or entity that produced or hosts it. If it's an internal dataset or a collection of publicly available layers compiled by an individual or group, describe it as clearly as possible based on its content and origin.

Q: Is it necessary to include the coordinate system or projection in a geospatial data citation?

A: While not always explicitly mandated by all citation styles, including the coordinate system and projection for geospatial data is highly recommended in Chicago style for clarity and reproducibility. This information is vital for understanding the spatial context and for users who may need to reproject the data for their own analyses.

Q: How do I cite data accessed through a web map service (WMS) or web feature service (WFS)?

A: When citing data from a WMS or WFS, you should identify the provider of the service, the name of the service, the specific layer or feature class accessed, the full URL to the service endpoint, and the date you accessed the data. As these services are dynamic, the access date is critical.

Q: Should I cite the software used to create or analyze geospatial data?

A: Yes, it is good practice to cite the software used for data creation or analysis, especially if specific versions or plugins were employed that might affect results. This information is typically included in the methodology section of your work rather than in the formal citation list, but it is crucial for reproducibility.

Q: What is a DOI, and why is it important for citing geospatial data?

A: A DOI (Digital Object Identifier) is a persistent, unique identifier for a digital object. For geospatial data, a DOI is incredibly valuable because it provides a stable link to the dataset, even if its URL changes. If a dataset has a DOI, it should always be included in the citation as it offers the most reliable way for readers to locate the data.

Q: How does Chicago Manual of Style handle citations for datasets that are constantly updated?

A: For constantly updated datasets, the best practice is to cite the most recent version available at the time of your research and to clearly state the date of access. If specific versions are labelled or have historical archives, citing the particular version used is even more precise. The goal is to capture a snapshot of the data as it existed when you used it.

[Chicago Manual Of Style Citing Geospatial Data](#)

Chicago Manual Of Style Citing Geospatial Data

Related Articles

- [chicago manual of style for book chapter citations](#)
- [chicago manual of style for academic journal publishing editors](#)
- [chicago manual of style capitalization rules for books](#)

[Back to Home](#)