Departmental Data Policy 2021-2: Department of Labor Geospatial Data Strategy

Executive Summary

Section 759(a)(1) of the Geospatial Data Act of 2018 (GDA) requires covered agencies to "prepare, maintain, publish, and implement a strategy for advancing geographic information and related geospatial data and activities appropriate to the mission of the covered agency." [1] As a covered agency, the Department of Labor (DOL or Department) with guidance from the Federal Geospatial Data Committee (FGDC) and in support of the National Spatial Data Infrastructure (NSDI) strategic plan, has developed a plan to ensure that any geospatial data it defines, collects, or uses is fit for purpose, interoperable with other public, private and Federal data, and offers the broad utility to all potential users.

While DOL's mission does not involve or require geospatial data generation, management, or use, the Department would directly benefit from developing more systematic data governance of any address or location data it does collect, and aligning such governance around an enterprise data strategy. The Department's primary approach to implementation is to create a shared service for address standardization, address validation and geocoding of location data, and start integrating this service into legacy and prospective data collections to begin building geospatial data capacity. [2] As the Department establishes data that can be leveraged in geospatial application and analysis, DOL's data governance body will work to identify training, opportunities, and tools to leverage the strategic value in that data for program administration, policymaking, and resource allocation.

Background

Federal agencies typically develop data strategy around the types of data their programs collect, documenting how they hope to better manage and use such information, and the associated goals for achieving efficient, consistent, and effective data governance across their collected and managed data.

Section 759(a)(1) of the GDA requires covered agencies to "prepare, maintain, publish, and implement a strategy for advancing geographic information and related geospatial data and activities appropriate to the mission of the covered agency." The GDA identifies DOL as a "covered agency," and while DOL does not have a mission that involves or requires geospatial data generation, management, dissemination, or use, the Department would directly benefit from developing more systematic data governance of any location data it does collect and aligning such governance within a data strategy.

Thorough understanding of the laws, related policies, and compliance framework is essential to ensure that any data strategy resulting from these mandates maintains relevance for the

[1] Final text: https://www.fgdc.gov/gda/geospatial-data-act-of-2018.pdf, US code: https://www.fgdc.gov/gda/43-usc-ch-46-geospatial-data-geospatial-data-act.pdf

[2] Geocoding, within this document, refers to the process of transforming a description of a location or address into a pair of geographic coordinates that describe a location on the earth's surface, and can be placed on a map.

Department, complies with requirements in law, and establishes a framework for building the necessary capacities, including producing actions and processes consistent with other agencies.

Overview of the Geospatial Data Act of 2018

The GDA was signed into law on October 5, 2018 as part of the Federal Aviation Administration Reauthorization Act (P.L. 115-254, Subtitle F). The GDA identifies organization and processes to develop and manage the NSDI and aligns agency geospatial data management with Federal efforts to make such work more rigorous, consistent and interoperable. The GDA also identifies governance processes related to geospatial data and clarifies policy and process to harmonize geospatial data management across Federal agencies.

Overview of key requirements in the law

The GDA includes numerous requirements that are assigned to three levels of entities: (1) the FGDC, (2) lead covered agencies, and (3) covered agencies. The FGDC oversees broad Federal implementation of the GDA and geospatial data governance across agencies. It includes requirements to prepare and maintain a strategic plan and lead development and implementation for the NSDI. Lead covered agencies also have requirements to establish goals that support the strategic plan for the NSDI prepared under section 755.

The Department is a covered agency under the GDA, and the requirement for covered agencies is to "prepare, maintain, publish, and implement a strategy for advancing geographic information and related geospatial data and activities appropriate to the mission of the covered agency, in support of the strategic plan for the NSDI prepared under section 755(c)."

Overview of Federal Guidance

Federal agencies implementing changes to geospatial data governance in response to the GDA must also be aware of supplemental guidance issued by the FGDC and the the Office Management Budget (OMB). The FGDC has issued the NSDI Strategic Plan 2021–2024 to organize Federal geospatial data management around a common set of goals and to harmonize Federal approaches to geospatial data management. As part of its broader leadership role, the FGDC also coordinates lead covered agency and covered agency strategic planning and reporting to Congress required by the GDA and provides template materials for agencies to develop strategic plans, data strategies and specific data management and governance actions.[3] In addition to the FGDC, OMB issues guidance to all agencies on geospatial data management under Circular A-16. It is important to note that this guidance has been pending revision since the GDA was signed into law in 2018 and, though drafts of revised guidance have been circulated for comment, no formal revision has been issued.

[3] The terms "lead covered agency" and "covered agency" are specifically described in the GDA at Section 752(9) and 752(3) respectively, and detail the roles and responsibilities these agencies have.

OIG Audit of the Department of Labor's Compliance with the GDA

In late fiscal year (FY) 2020, DOL's Inspector General conducted and completed an audit[4] of the Department's compliance with the GDA. The audit recognized that DOL does not produce geospatial data assets, but also identified the lack of enterprise governance for the location and address data the Department collects for individuals and businesses. The audit recommended DOL implement appropriate governance actions and management processes to support broader analytical use of resources, such as address data, and to ensure these actions support the broader Federal efforts identified by the GDA and FGDC.

Geospatial Strategy Planning Approach

Overview of Department of Labor Geospatial Data

DOL data are an invaluable asset to the American labor force, businesses, policy-makers, and the public. Effective strategic planning for this data requires aligning proposed governance, data management policies, and planned capabilities appropriately with agency mission, the types and condition of data collected, and the roles that data will play. Location data constitutes a very small portion of the data collected by the DOL, and virtually all of that is in the form of addresses for individuals, establishments or entities who form the regulated community for DOL's programs. This strategy focuses on developing and implementing enterprise governance for address data to begin building capacities to better leverage this data and make it a more useful asset to develop geospatial capabilities in the future, and to ensure it will support the widest possible set of stakeholders.

Types of geospatial data

Virtually all DOL programs collect location data as part of their routine administration of their programs and virtually all of the locations collected are in the form of addresses. This address data represents the only meaningful instance of location data the Department collects. Addresses may represent locations for a wide variety of places including but not limited to sites at which DOL-funded services are provided, establishments at which inspections or compliance actions have been taken, physical addresses for entities that are part of DOL's regulated community, locations to mail compensation from Federal programs, and addresses of record for enforcement and litigation.

Uses of data

Address data is principally used at DOL for its intended purpose: providing Federal materials through mail or delivery services to stakeholders and members of the regulated community. There are limited cases of geocoding for address to plot activities on maps, estimate distance between events, and assess the alignment of Federal activities. An example of this is where Federal grants are provided, along with where individuals who have needs for the services offered through the grant are located, to assess the alignment of service delivery with need.

Without a common Federal establishment identifier, address data and components of address data are often used to perform probabilistic linkage of records. This represents a common and

[4] Summary: https://www.oig.dol.gov/public/reports/oa/viewpdf.php?r=23-20-004-01-001&y=2020&t=b, Full report: https://www.oig.dol.gov/public/reports/oa/viewpdf.php?r=23-20-004-01-001&y=2020

important use for addresses in many Federal agencies and allows programs to find prior instances of inspection or enforcement records for specific locations, or to track compliance and enforcement actions across the different programs of the Department.

DOL currently shares data from a variety of programs with the public as part of DOL's open data efforts. Known use cases for these data by the public are minimal, principally due to the uneven quality of the information. Efforts to improve these data would be expected to lead to greater number of use cases by other Federal, public, and private entities and a greater probability that promising use cases for programs at DOL might emerge.

Data acquisition approaches

Address data is collected from many different sources and through many different systems. These include but are not limited to:

- Internal administrative data and case management systems;
- Data sets provided by grantees or Federal partners documenting services delivered;
- Required reporting from the regulated community;
- Intake systems that document the characteristics of participants in Federal programs; and
- Surveys of individuals, establishments, business owners or entities who are responding with data to the Federal government or describing interactions with Federal programs.

With a few exceptions, there are not currently any standards established within DOL for the definition, collection, maintenance, and use of address data. While there are some instances of *ad hoc* address validation or standardization, the vast majority of address data collected within DOL is through free text entry fields without any meaningful quality control measures. Due to these factors, there is little consistency or predictability in the format and content of address data collected across DOL's programs. This is an especially prominent issue in cases where public or private entities are reporting directly into Federal systems, and concerns about validation processes negatively affecting response rates leads to relaxing quality control measures. Performing validation and standardization on data provided to the Federal government from external partners, grantees or respondents who have defined, collected, and managed their data in a parochial or proprietary manner, is especially important to ensure that the resulting information is fit for purpose and interoperable.

Geospatial data management

Adopting address standards for location data is the simplest and most direct way to ensure that Departmental programs create and manage data that are fit for purpose, can be leveraged by the broadest set of users and applications, and are interoperable with other Federal, public, and private data. Implementing standards for data content and format creates data management processes that are more predictable, consistent and effective. Issuing the standards as Departmental Data Policy will signal that the department has adopted a voluntary consensus standard and will codify consensus around the standard and its implementation.

DOL's Chief Information Officer (CIO) will work to stand up a shared service within the Department to support address validation, address standardization, and geocoding. The CIO will

also work to establish complementary capabilities for agencies to integrate this service into their prospective and legacy data systems, and to provide support to agencies in the storage and retrieval of additional data resulting from address validation and geocoding. The DOL Data Board, the Department's representative data governance body, will work in close collaboration with the CIO to create organizational support and resources to help agencies successfully implement data validation procedures and best support compliance with, and adoption of data governance policies for address data.

Implementing Departmental data standardization and validation programs has a tremendous potential to improve data quality and interoperability, and further increase the value and utility of DOL data. Standardization and validation will substantially reduce errors associated with the free-text entry of addresses or components of addresses, increase analytical possibilities for DOL data, and increase the efficiency of functional and analytical work with address data for all Departmental data users.

Infrastructure supporting geospatial interoperability & data sharing

DOL's CIO is currently planning to create a shared service within the Department to support address validation, address standardization, and geocoding. The CIO will also work to establish complementary capabilities for agencies to integrate this service into their prospective and legacy data systems and to provide support to agencies in the storage and retrieval of additional data resulting from address validation and geocoding.

Creating a framework for address data validation and standardization should allow DOL's existing open data and data sharing capabilities to provide this seamlessly. DOL currently places much of its data into the public domain, and the challenges it faces are not related to data sharing but to the quality and capacity within its data. As DOL moves to institute data quality and format standards, various Federal, public, and private entities should benefit from the improved data quality and bring additional value to DOL through increases in the amount and types of innovation and analytical development using the location data available.

Administration/governance

The DOL Data Board, the Department's representative data governance body, will take principal responsibility, in close collaboration with the CIO and Chief Data Stewards (CDS) from the agencies, for establishing standards for address data. The DOL Data Board will continue to work in close collaboration with the CIO to create organizational support and resources to help agencies successfully implement data validation procedures and best support compliance with and adoption of data governance policies for address data.

Description of alignment of DOL's geospatial strategy with the DOL strategic plan and the NSDI strategic plan

Beginning in Federal FY 2021, DOL had all agencies incorporate data projects into their operating plan. Integrating the management and use of data into DOL's operating plan was beneficial for a variety of reasons, including reinforcing the importance of data for the enterprise and its critical role in the successful conduct of the agency mission,and leveraging existing planning, budget, and performance systems to minimize the need for additional process and

bureaucracy. Adding projects expanding and improving the management and use of data to operating plans provides additional methods to institutionalize processes that improve data capacity within agencies.

As DOL begins planning for next FY and the intermediate future, the Department will continue to integrate aspects of consistent and predictable data governance into planning, performance and budget processes. To build data capacity and to reach the goals of the Foundations for Evidence-Based Policymaking Act of 2018[5], the GDA, and the Federal Data Strategy[6], DOL is considering adding various data-related goals as agency priority goals and incorporating the work into broader agency level planning.

Evolving guidance such as the *Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking*[7] tasks agencies with developing annual data plans and offers additional places where governance of all data, including location and geospatial data, can be addressed, improved, and prioritized.

To the extent that the broader geospatial data user community finds benefit in the data that DOL produces and disseminates, an improved internal focus on data quality and open data should ensure that data would be available for assessment, use, and evaluation. As integration of address validation, address standardization, and geocoding becomes more prevalent at the Department, the DOL Data Board will continue to assess whether an expanded role for geospatial data would benefit DOL programs, and will commit to taking appropriate actions. DOL will continue to meet its obligations to fund the Federal geospatial data platform and evaluate whether offering data to the platform or leveraging the data on the platform would bring greater benefit to Departmental efforts to improve policymaking, strategic planning, or program administration.

Resourcing

The GDA did not provide any resources to Federal agencies to assess the current state of their data assets, design approaches to embrace the spirit of the law, or build important capabilities that help agencies comply with the law. It will be challenging to stand up the necessary services and begin integrating them into prospective and legacy systems given the current state of resources and the minimal role that geospatial data plays in the administration of DOL programs. It will be extremely challenging to achieve full compliance and institutionalize these changes to maintain higher levels of quality, in addition to wider benefits through transparency and data sharing, for Federal geospatial data.

The Chief Data Officer (CDO) and CIO will continue to collaborate to expedite this work among the many other critical IT priorities that DOL agencies face. The Department will work with agencies to identify opportunities such as planned system modernizations, resource re-allocation, and re-programming existing sources of revenue where these actions can be reasonably

[5] https://www.congress.gov/bill/115th-congress/house-bill/4174
[6] https://strategy.data.gov/
[7] https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/

accomplished, but the pace of adoption will rely on agencies prioritizing this work alongside a number of other pressing technological needs.

GDA Roles and Responsibilities

Overview

Data governance and implementation of data strategy requires considerable effort, diverse resources, and numerous individuals from across the enterprise to operate towards common goals in a cohesive way. Effective data governance requires coordination across all levels of the organization, particularly when organizations are trying to increase their sophistication and data management maturity. Beyond the specific roles that individuals or teams play in improving geospatial data management and achieving effective governance, advocates are also needed for this work as well as a willingness to lead agencies through the cultural changes necessary to give greater priority to data quality and utility.

Departmental Leadership

Every successful data governance system has strong executive support. While adopting policies is relatively simple, culture change is extremely difficult. It requires persistent attention, the allocation of appropriate resources, prioritization of this work among many competing priorities, and a continuous emphasis on enterprise approaches to problem solving. With respect to implementation of geospatial data policy, Departmental leadership, including the Secretary, Deputy Secretary, and the political and career leadership within DOL component agencies, play a vital role in ensuring that issues of data management are appropriately prioritized in agency budget and strategic planning, reflected in their staff's priorities and working plans, and generally aligned with departmental goals.

Chief Data Officer

With respect to implementation of geospatial data policy, the CDO is responsible for identifying compliance requirements in the GDA and coordinating the development of geospatial data policy that is appropriate for the data DOL collects and relevant to the mission DOL conducts. Once the policy is developed and implemented, the CDO works with the other partners to help Departmental programs implement the policy and improve the quality of data governance and data management. The CDO, in conjunction with all of the other partners, tracks the progress of implementation and the ways in which data are being used towards the goals identified in the GDA and DOL data management plans. The CDO leads efforts to develop and propose changes to geospatial data management policy as needed.

Chief Information Officer

With respect to implementation of geospatial data policy, the CIO is responsible for providing strategic leadership and full management and execution for information technology related to geospatial data management, validation, and geocoding services. This work includes but is not limited to planning, budget, development, maintenance, implementation, and ongoing operation of shared services and technical solutions identified in DOL data policy. The CIO oversees the development, utility, and maintenance of services to ensure that developed systems constitute the full range of support necessary for appropriate geospatial data management for DOL data.

Agency Chief Data Stewards

With respect to implementation of geospatial data policy, CDSs are responsible for working with Departmental leadership, the CDO, and CIO to lead the data collection programs within their agencies toward compliance with the GDA and Departmental data policy. They have principal responsibility for coordinating necessary training, stakeholder communication, and administrative changes to their data collections resulting from the integration of shared services for data validation, standardization, and geocoding. CDSs should work to ensure that the process of adding validation and standardization, and leveraging the new data produced by geocoding into analytical products is harmonious, efficient, and effective. CDSs shall also be responsible for regular reporting to the DOL Data Board on progress within their programs implementing Departmental data policy, communicating the importance of this work to their staff, and building capabilities that expand the utility of their data and the ways in which they can leverage data as a strategic asset.

Individual Data Set Data Stewards

With respect to implementation of geospatial data policy, Data Stewards (DS) should work in close coordination with their CDS to ensure that technical changes made to data collection systems, and the data resulting from those changes, are clearly documented, appropriately implemented, and communicated to stakeholders. Their work should ensure that data are readily accessible and interpretable, interoperable with other Departmental data, shared appropriately, and leveraged to the greatest extent possible. They should coordinate with other DS to identify implementation challenges and best practices to make this work as efficient and cost-effective as possible.

DOL covered agency responsibilities

DOL, for the purposes of the GDA, is a covered agency, and the 13 covered agency responsibilities are described in section 759. As noted previously, the DOL Inspector General has evaluated and reported on DOL's compliance with the 13 responsibilities, and identified some areas in which DOL has complied, some in which there has been partial progress, and some in which there has been no progress.

The GDA does not draw distinctions between agencies who have significant responsibilities, involvement with, and dissemination tasks for geospatial data (e.g. Departments of Interior, Commerce, Transportation, etc.) and agencies such as DOL that do not produce, procure, manage, or use much geospatial data in the routine conduct of the mission. In spite of this, DOL will take actions to meet its responsibilities under the law that are most appropriate for the data it generates and uses, the available budget and technology it can access, and the perceived benefit the public sees in DOL data. The CDO and DOL Data Board will assume the responsibility of ensuring that progress is made on all 13 items by establishing teams with responsibilities in areas of information technology, budget, performance, process improvement, and innovation.

DOL GDA reporting responsibilities

The CDO, in close collaboration with the CIO and DOL Data Board will take principal responsibility for addressing current and emerging reporting requirements from the FGDC, OMB, Federal Data Strategy, and other Federal guidance. For FY 2021, the FGDC has designed

a survey-based questionnaire that agencies are being asked to use to provide responses on their compliance with the GDA. The FGDC will summarized Agency responses and compile them into a report, which the FGDC will submit to Congress on behalf of all covered agencies. As additional reporting responsibilities materialize, the CDO and Data Board will work to update enterprise data strategy and data governance plans to reflect goals accomplished, as well as new or shifting requirements.

Goals and Objectives

Overview

The general objectives of this geospatial data strategy are to ensure that geospatial data collected by the Department is consistently fit for purpose, managed to allow for optimal benefit to the widest set of stakeholders and for the broadest set of purposes and can be consistently interoperable with other data of the same kind.

Specific Departmental Objectives

DOL's goal is to implement a strategy for building geospatial data capacity that is appropriate to the agency's mission, current and prospective use of geospatial data, and the amount and types of geospatial data collected. The only notable geospatial data that DOL collects, maintains, and uses in the routine conduct of its mission is address and location data. Departmental geospatial strategy goals and objectives should appropriately be focused entirely on improving the quality, consistency, and utility of this data.

Consistent with guidance from the FGDC, this strategic plan is conceived and designed to cover the time period from FY 2021 through FY 2024.[8] The Department will update this strategy to reflect any necessary re-alignment or re-prioritization resulting from publication of the NSDI strategic plan, or changes to Federal guidance issued through the FGDC, OMB, or the Federal Data Strategy. The Department will also update this strategy if implementation of the primary actions described above changes substantively, or sufficient challenges arise that modifications to the broad approach must be made.

The broad objective is to guide DOL programs towards data generation, management, and use that is more broadly in compliance with the requirements outlined by the GDA. The CDO and DOL Data Board anticipate formal release of new standards for geospatial data and metadata, and this plan allows the Department to adapt to any evolving standards and ensure compliance.

Implementation

Implementation approach

During FY 2021, DOL will seek to establish a shared service for address validation, standardization, and geocoding. As the shared service is being designed and implemented, DOL will begin to work with agencies to identify appropriate instances where this shared service can be integrated into prospective or legacy data systems. Agencies will need to work through Departmental leadership and the CIO to appropriately prioritize this work and identify funding sources as the GDA did not provide a supplemental source of funding.

[8] See FGDC Covered Agency Geospatial Strategy Template FINAL, dated November 19, 2020.

The CDO and DOL Data Board will coordinate the agencies in gathering experience with integration of the shared service, and production of any supplemental guidance that can help with successful transition to address data that is validated and standardized. The CDO and DOL Data Board will also work to see how emerging geospatial data at DOL can benefit the wider Federal community through geospatial analytics within DOL, sharing data as appropriate through open data programs, and participation in the Federal Geospatial Data Platform.[9]

There are numerous challenges to overcome for the actions described in this strategy to produce the intended changes. The principle challenges will be finding adequate resources to develop the shared services and integrate them into data collections, and reprioritizing scheduled work to expedite this work, and integrating this and other data projects into existing departmental budget and capital planning work. The extent of progress towards making location data collected by the Department more interoperable and useful will be directly related to the pace at which systems integrate shared services for address standardization, validation, and geocoding.

Description of alignment with OPEN Government Data Act requirements

The activities identified in this plan were designed to complement DOL's current efforts to make data more widely available to the public, and to be fully aligned with requirements for open data in the Evidence Act. Location data collected by agencies now is highly heterogeneous in format and quality, and this increases burden for people seeking to use the data analytically, visually, or spatially.

The Department currently makes a large proportion of the data it collects public through agency web sites, topic-specific data portals, and data clearinghouse services. The areas in which the Department seeks to make improvements is not in sharing or disclosing data, but in making the data it shares easier to work by standardizing formats and more useful and meaningful by validating content and adding geocodes.

The Department's most substantial investments in open data over the next two years will go towards creating a new application programming interface[10] (API) service to better support open data. The department is aligning all geospatial data governance to better enhance the quality of the data offered by DOL through this service and ensure that the data can reach the widest possible set of users, in the easiest possible format.

Reporting

DOL will continue to follow FGDC guidance on reporting the outcomes of planned actions and progress towards more consistent and effective geospatial data governance. DOL anticipates support revisions to the reporting process through pending reissuance of OMB Circular A-16, the release of metadata standards for geospatial data, and additional FGDC guidance.

[9] https://www.geoplatform.gov/

[10] APIs provide data-as-a-service and allow users to query subsets of records, in machine readable format, and supported by machine readable metadata.

The DOL Data Board will track progress with implementation of geospatial data governance, how agencies are able to leverage the resulting data, and how data are accessed by users through the emerging API service. DOL has anticipated this requirement and should be able to report quantitative metrics on integration of validation including percent/number of data sets, amount of legacy data geocoded, and access/usage statistics for location data being served through the API.

Revision History

Date	Revision	Name/Entity	Description
01/15/2021	1.0	DOL Data Board	Initial Issuance of Geospatial Data Strategy