

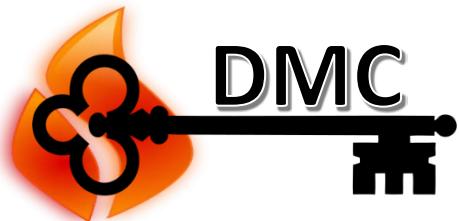
A publication of the
National Wildfire
Coordinating Group



NWCG Data Management Strategy

PMS 940

MARCH 2018



NWCG Data Management Strategy

March 2018
PMS 940

The *NWCG Data Management Strategy* establishes the roadmap by which national, interagency wildland fire data is managed to maximize data utility for efficient use across wildland fire business processes and IT systems.

This Strategy document focuses on defining opportunities and priorities to address long-standing challenges with wildland fire enterprise data.

The National Wildfire Coordinating Group (NWCG) provides national leadership to enable interoperable wildland fire operations among federal, state, tribal, territorial, and local partners. NWCG operations standards are interagency by design; they are developed with the intent of universal adoption by the member agencies. However, the decision to adopt and utilize them is made independent.

Table of Contents

Introduction.....	1
Guiding Principles.....	1
Mission	2
Goals.....	2
Goal 1: Data Standardization	3
Goal 2: Data Lifecycle Management	4
Goal 3: Data Quality.	5
Execution and Implementation.....	6
Appendix A: Metadata Dictionary Components	7

Introduction

Responsibility for data management lies primarily with the National Wildfire Coordinating Group (NWCG) Data Management Committee (DMC) in support of the Fire Management Board (FMB). The DMC collaborates with the Wildland Fire Information and Technology (WFIT) Program Management Office (PMO) and Program Board (PB), the Office of the Chief Information Officer (OCIO) organizations of the US Department of Agriculture (USDA) and Department of Interior and stakeholder groups (e.g. agency/bureaus, non-federal, state, and tribal organizations).

The DMC is responsible for maximizing data utility for efficient use across wildland fire systems. The DMC provides guidance through the following documents and products:

- *Geographic Information System Standard Operating Procedures on Incidents*, PMS 936
- *User Guide for Development and Maintenance of NWCG Glossary of Wildland Fire Terminology*. PMS 937
- *User Guide for Development and Maintenance of NWCG Data Standards*, PMS 938
- *Unit Identifiers*, PMS 931
- *Data Management Standard for Incident Complexes and Merged Wildfires*
- Data Exchange Model
- *NWCG Data Dictionary* (in development)
- *NWCG Guide for Data Management* (in development)

The DMC strives to facilitate development and implementation of data standards, yet the current state of the architecture supporting wildland fire makes data standardization at the enterprise level challenging. Although NWCG standards exist for some data elements within the wildland fire community, a majority of the data over the past 20 years were independently defined to support the application in which they were created. This ad-hoc data creation caused “data silos,” making data exchange across applications problematic and sometimes impossible. Data dictionaries, if they exist within these silos, are commonly held within the application’s development environment which makes stakeholder access challenging.

This Strategy document focuses on defining opportunities and priorities to address long-standing challenges with wildland fire enterprise data.

Guiding Principles

The following principles shall guide this Strategy:

The NWCG DMC is responsible for managing data as a business asset:

The DMC has the knowledge and authority to make decisions on how the master data and metadata is defined, maintained, archived, and how changes are authorized and audited. The DMC coordinates with other NWCG committees and IT Programs to ensure data governance is effective and aides business and project objectives.

Utilize Authoritative Wildland Fire Data:

Wildland Fire IT does not “own” wildland fire data. Instead, participating NWCG organizations own the data as a part of their overall agency resource management programs. At an enterprise level, authoritative wildland fire data sets are created by aggregating data from various agency sources, including Incident Management Teams (IMT) and local, regional, and national applications. Agencies may have Systems of Record (SOR) external to the wildland fire enterprise or may have designated wildland fire applications as their agency SOR. This relationship between agencies and the wildland fire enterprise is a fundamental driver for how data is managed.

Support the Wildland Fire Information and Technology 5 Year Investment Strategy:

This Strategy aligns with the WFIT goals for enterprise investment management. The DMC strategy is one of five foundational WFIT program elements and corresponding strategies, commonly referred to as DISAA. DISAA stands for:

- Data Management
- Infrastructure
- Security Alignment and Implementation
- Acquisition of Technology
- Applications Architecture

Adhere to OCIO Standards:

Ensure NWCG Data Management Program complies with established OCIO directives for data governance; metadata; data quality; data architecture; data modeling and design; data storage and operations; data security; data integrations and interoperability; document and content management; reference and master data management; data warehousing and business intelligence. This includes sourcing and availability of public information via data.gov and compliance with the Office of Management and Budget (OMB) Open Data initiative.

Mission

The mission of the Data Management Strategy is to define a roadmap to develop and implement national interagency standards and practices for wildland fire data management. This includes:

- Assisting the wildland fire community to identify, define, and standardize data that is reliable and accessible for planning, decision support, reporting, and research.
- Developing programmatic guidance for wildland fire data, including data requirements, data governance, and data architecture that support a data exchange environment and improved efficiency in operational work and communication processes.

Goals

Achieving the goals of the Data Strategy requires a committed, collaborative, and sustained effort. The goals are intertwined and re-enforcing. Objectives and initiatives associated with each of the three goals further define the desired outcomes of the Data Strategy. The data will serve as the foundation for the Data Cache and wildland fire architecture—increasing accessibility to and usage of data capabilities that enhance management and decision making.

Goal 1: Data Standardization: Provide standardized data to support the wildland fire program needs across all lines of business.

The target data environment includes standardized data and authoritative sources to ensure planning and response is based on the best available information. Collaboration and coordination with both internal and external business areas, primarily through existing NWCG processes, will be necessary.

Metadata, commonly defined as “data about data”, provides context for transactional data values. A Metadata Dictionary (MDD) will be created to ensure data is collected, organized, and available in applications or to business functions in a consistent manner. A data element included in the MDD as approved becomes the interagency standard. The MDD will define, structure, and catalog metadata.

- *Defining metadata:* data that defines and specifies other data.
- *Structuring metadata:* data that specifies the relationships between data elements and data groups.
- *Cataloging metadata:* data about the location, identity, history, validity, lineage, source, and destination of other data.

The number of separately supported datasets may decrease through design that is more efficient and sharing of data among systems and applications. Currently, multiple versions of similar data exist across wildland fire and it is difficult to determine the best enterprise data. The DMC plans to remedy this problem by assigning an Interagency Authoritative Data Source (IADS) or Interagency System of Record (ISOR) to each data element. A successful data environment will require data producers keep the IADS / ISOR updated and require data consumers get their data from these sources.

Identifying one IADS and ISOR for each data element will drive data integrity, accuracy, and architectural improvements.

Goal 1 Objectives/Initiatives

- 2.1. Define the data elements within the Metadata Dictionary (MDD)
 - 2.2.2. Enter existing and proposed data elements into the MDD
 - 2.2.2. Develop approval and maintenance guidelines for data elements
 - 2.2.2. Develop MDD records for data not yet standardized
 - 2.2.2. Create standardized domain tables that can be used across the enterprise
- 2.2. Develop the NWCG Guide for Data Management
 - 2.2.2. Define business practices and reporting requirements for the interagency fire community
 - 2.2.2. Modify the PMS 938 User Guide for Development and Maintenance of NWCG Data Standards to reflect updated terminology and changes to the data requirements
 - 2.2.2. Develop the migration guidance for migrating non-standardized data
- 2.3. Continue to develop and validate the Wildland Fire Logical Data Model (LDM)

Goal 2: Data Lifecycle Management: Effectively manage wildland fire data throughout the entire data lifecycle.

Data lifecycle management (DLM) is a comprehensive approach to managing wildland fire data that involves procedures and practices as well as applications. The DLM is a policy and business based approach to managing wildland fire data from creation and operational use to archiving for planning and research.

In the past, little if any transactional metadata was collected or exposed. Planning for, collecting and sharing transactional metadata will be increasingly important as Wildland Fire data is made available to the public.

Most national level performance measures for Wildland Fire are based on a rolling ten year window. This is a key component of agency funding processes. For planning and research purposes, historical data needs to be available as far back as possible. Some local units have close to 100 years of data but it is not easily accessible nationally or may even still be paper reports.

Effective DLM will enable more robust, reliable, and complete information by planning for and documenting the status and context of wildland fire data.

Goal 2 Objectives/Initiatives

- 2.1. Define the Data Lifecycle
 - 2.1.1. Identify the different stages of data within the data lifecycle. For example: Plan, Acquire, Process, Analyze, Preserve, Publish/Share and Re-use.
- 2.2. Assess/Prescribe how data moves within the Wildland Fire portfolio
 - 2.2.1. Assess how the data moves within the enterprise by mapping between the MDD and current data sources.
 - 2.2.2. Identify the Interagency Authoritative Source (IADS) and Interagency System of Record (ISOR).
 - 2.2.3. Work closely with the Infrastructure, Security, Architecture, and Acquisition strategies to ensure the data lifecycles are aligned with all foundational element strategies.
- 2.3. Manage Wildland Fire Data transactional metadata
 - 2.3.1. Define requirements for an automated process that captures metadata from IADSs and makes the metadata available to support the goals and objectives in this strategy
 - 2.3.2. Create a reporting system that flags data elements of interest (outdated/in need of lifecycle redefinition) to the Data Governance Board

Goal 3: Data Quality: Assess Quality Assurance (QA) and Quality Control (QC) across data elements.

Before trying to assess and control the quality of the data there has to be a target, a set of criteria that states what the data should be like. This takes the form of a set of definitions of each data element in the system (Goals 1 and 2). These definitions can be rigorously applied to cleansing historical data. Active data governance would ensure the definitions and standards are understood and applied by the people working with the wildland fire applications and executing day to day operations to create clean data from the start.

The DMC defines Data Quality as:

- Fit-for-use—complete and free from defects
- Relevant – the right data for the intended purpose
- Well-defined – master and metadata standards exist and data complies with the standard

This goal is focused on continuously assessing the data quality, data exchange, and how successful the data is supporting the business needs of the community.

Goal 3 Objectives/Initiatives

3.1. Develop Data Governance and Charter Data Governance Board

- 3.1.1. Create a Governance Plan that includes a Data Governance Board to review new and changing data requirements.

3.2. Develop Metrics for: Data Quality, Application Implementation, and Data Exchange

- 3.2.1. Assess national data completeness and coverage by applying initial, high level metrics on existing data sources and developing metrics to apply to the Interagency Fire Data Cache as it matures.

- 3.2.2. Create metrics to assess the implementation of data standards within enterprise applications. Assessing implementation will be an ongoing effort as applications are refreshed or added to the portfolio. The DMC will coordinate with the PB and the other DISAA teams to complete this assessment.

- 3.2.3. Create a metric to reflect the number of applications that create or edit specific data elements. Reducing the number of locations where key data is created or edited will reduce the number of errors and inconsistencies in the data and the complexity of individual applications.

3.3. Define requirements for a dashboard for QA/QC Metrics

Execution and Implementation

The NWCG Executive Board, the FMB, the WFIT PB, the OCIO and stakeholder groups (e.g. non-federal, state, and tribal organizations) may provide recommendations to the DMC to implement and achieve the Data Strategy's Goals and Objectives.

The DMC will utilize NWCG Committees, Subcommittees, Working Groups, and Task Teams to achieve each of the three goals.

Completion of the goals will be dependent on resource availability. Currently, only one member of the DMC is dedicated to this work full time. Given this, the Data Strategy is expected to span 3-5 years. The strategy will be reviewed annually at the DMC meeting and revised as appropriate.

The DMC will report progress toward the Data Strategy Goals to the NWCG Executive Board and FMB once a year or as requested. The DMC will also regularly report to the PB along with the other DISAA teams on a schedule determined by the PB.

The DMC is committed to achieving the highest quality data possible for planning, reporting, and decision making for the Wildland Fire Program. This Data Strategy is our blueprint to accomplish this.

Appendix A: Metadata Dictionary Components

Attributes for Data Elements	Required	Short Description
NWCG Data Element Name	Y	<p>A title that describes the content and facilitates usage of the data value. This title is structured for formal data management use.</p> <p>Ex: Final Fire Acre Quantity</p> <p>Some data elements may be combined or have defined adjectives (see Descriptor Domain) to better represent data values.</p> <p>Ex: Lat/Long or Point of Origin Lat/Long</p>
Data Exchange Data Element Name	Y	The specifically formatted word or phrase used by computers when sharing the data via an exchange mechanism.
Data Element Abbreviation or Acronym		The shorthand version of the data element name, if one is used.
Also Known As		<p>A pseudonym or operational term for a data element. This is the term the business uses.</p> <p>Ex: Fire Size</p>
Short Description	Y	<p>Text that briefly characterizes the data element.</p> <p>This description must align with applicable NWCG Glossary definitions.</p>
Data Type	Y	<p>A particular kind of data item, as defined by the values it can take, the programming language used, or the operations that can be performed on it.</p> <p>Ex: Text, String, Numeric, etc.</p>
Format	Y	The arrangement of data for computer input or output, such as the number, and size of fields.
Example		A sample of the values expressed based on the standard.
Subject Area	Y	<p>A high level classification of data representing a business category. Data elements are assigned a Subject Area based on the business processes that create the data values. Ideally a data element has only one Subject Area.</p> <p>Subject Areas are also applied to NWCG committees as business stewards and to IT investments defined by the WFIT Framework.</p>

Attributes for Data Elements	Required	Short Description
Business Data Group (BDG)		<p>A group of data elements that, when viewed together, provide more complete information, or knowledge than when viewed independently. A data element may be used in more than one BDG.</p> <p>BDG are made up of Data Elements and could be a Feature Class. Business stewards define these subsets of data within or across Subject Areas and/or logical data entities.</p> <p>Business Data Groups can be used to define mandatory data for specific business processes.</p> <p>Ex: Minimum data required to create an incident record, etc.</p>
Descriptor Domain		<p>A defined group of adjectives used to describe a single data element.</p> <p>Ex: a descriptor domain for acres would include: Discovery, Initial Response, Daily, Final, etc.</p> <p>A descriptor domain would not be used as a pick list in an application user interface. Instead, it is a construct that facilitates consistent use of approved “adjectives” for data elements.</p>
Logical Data Group		<p>A logical grouping of data elements that belong together. These are commonly known as entities in a Logical Data Model.</p>
Related NWCG Glossary Term		<p>Lists and links any terms from the glossary that are related to the data element.</p>
Related NWCG Geospatial Data Layer Standard		<p>Lists and links any geospatial data standards that are related to the data element. These data elements are attributes of a point, line, or polygon.</p>
Related Publication		<p>Lists and links any NWCG publications, policy, or other guidance that are related to the data element.</p> <p>Ex: Redbook, Interagency Helicopter Operations Guide, etc.</p>
Business Rules (BR)	Y*	<p>Intended to assert business structure or to control or influence the behavior of a business function.</p> <p>Ex: Every wildfire must have a Local Incident Identifier assigned</p> <p>*N/A is a valid value</p>
Supporting Business Documentation		<p>Optional detailed documents about the business discussion and decision process that led the steward to the final product. Specific to the business definition and description—could reference a policy or other guidelines, etc.</p>

Attributes for Data Elements	Required	Short Description
Technical Rules (TR)	Y*	<p>Intended to assert technical implementation structure, often to comply with metadata standards or business rules.</p> <p>Ex: Pad Local Incident Identifier with leading zeros to meet minimum field length</p> <p>*N/A is a valid value</p>
Data Element Min/Max Length		<p>The maximum and minimum allowable lengths for the data—both raw and data entry.</p>
Case Sensitivity	Y	<p>Differentiating between capital and lowercase letters—values may be treated differently depending on whether it is in capitals or lowercase text.</p>
Sharing Sensitivity	Y	<p>Indicates the level of protection needed to prevent unauthorized access of a data element or control the timing of release to the public.</p> <p>Valid Values: Public, Slow (delayed release to public), Sensitive.</p>
Valid Values Domain		<p>A pre-defined list of accepted data values used to standardize data entry, ensure data consistency, and prevent errors. Valid values domains facilitate re-use across the enterprise and could be implemented as a pick list in an application UI.</p>
Data Validation Rule	Y*	<p>Test used to ensure a value submitted complies with established Business or Technical Rules.</p> <p>Ex: Local Incident Identifier is not null.</p> <p>Local Incident Identifier has 6-10 characters.</p> <p>*N/A is a valid value</p>
Data Validation Failure Error Message	Y*	<p>A standardized message(s) provided to a user or application when a violation of business or technical rules or other metadata attribute standards has been detected by a Data Validation test.</p> <p>*Required if there is a Data Validation Rule</p>
Interagency Authoritative Data Source (IADS)	Y	<p>A product, tool, or IT application that has been designated as the trusted source for wildland fire data. This source may also create and update transactional data for use in other applications.</p> <p>There may be more than one IADS and it can change depending business process complexity and life cycle.</p> <p>An IADS may be a compilation or subset of data from other authoritative sources. The DLM process ensures sources, limitations, currency, and attributes for the IADS are documented.</p>

Attributes for Data Elements	Required	Short Description
Interagency System of Record (ISOR)	Y	<p>Agencies and bureaus may have their own SOR for their data and</p> <p>An ISORs are identified by an interagency business area as the official application source of interagency data.</p> <p>An ISOR is the source that resolves duplicate records that may arise from various IADSs and ensures the data meets defined quality standards before it is included in official historical data sets.</p> <p>An ISOR can be an external source of data used by wildland fire.</p>
Where Used	Y	Applications that utilize the data values to fulfill a business process. Where used indicates the data is manually or automatically added to the application.
Data Stewardship Group	Y	An NWCG committee or subgroup responsible for developing a data element or data layer standard or glossary entry and ensuring consistency across agencies.
Standard Source		Identifies the source of the standard if NWCG is adopting it rather than creating it, e.g. NIST, NWS, FEMA, etc.
Status	Y	The state of a data element standard in the request process—from requested all the way to archive.
Status Current As Of		Date the NWCG data element changed status.
Version		The official iteration of a data element.
Previous Versions		A historical list of all the previous iterations of a data element standard.
Data Standards and Terminology Subcommittee (DSTSC) Decision Documentation		An optional comment about decisions during the process, including why it was rejected, why it was or was not changed as a result of a request, or why it was archived and if/what replaced it. Specific to DSTSC Process.
Dictionary ID		Unique identifier assigned to each record in the data dictionary. (Primary key)

The NWCG *Data Management Strategy* is developed and maintained by the Data Management Subcommittee, an entity of the National Wildfire Coordinating Group (NWCG).

Previous editions: None

While they may still contain current or useful information, previous editions are obsolete. The user of this information is responsible for confirming that they have the most up-to-date version. NWCG is the sole source for the publication.

This publication is available electronically at <https://www.nwcg.gov/publications/940>.

Comments, questions, and recommendations shall be submitted to the appropriate agency program manager assigned to the Data Management Subcommittee. View the complete roster at <https://www.nwcg.gov/committees/data-management-committee/roster>.

Publications and training materials produced by NWCG are in the public domain. Use of public domain information, including copying, is permitted. Use of NWCG information within another document is permitted if NWCG information is accurately credited to NWCG. The NWCG logo may not be used except on NWCG authorized information. “National Wildfire Coordinating Group,” “NWCG,” and the NWCG logo are trademarks of NWCG.

The use of trade, firm, or corporation names or trademarks in NWCG products is solely for the information and convenience of the reader and does not constitute endorsement by NWCG or its member agencies or any product or service to the exclusion of others that may be suitable.