MANAGED BUSINESS SOLUTIONS
ANALYSIS OF REQUIREMENTS
DATA REPOSITORY – STORE - CACHE
FOR THE INTERAGENCY WILDLAND FIRE PROGRAM
REQUIREMENTS DOCUMENT

Rachel Smith / Al Gates / Monica Pirlot / Jay Steen
MBS, LLC
Creation Date: April 9, 2018

TABLE OF CONTENTS

PROJECT OVERVIEW

A pain point for Wildland Fire stakeholders is related to base fire information. Internal and external stakeholders have concern that there is not a single authoritative source to obtain reliable and timely fire information for all aspects of fire business. Wildland Fire wants to take an integrated service-oriented approach to developing a system that provides an integrated and coordinated process for collecting fire related data. Key outcomes of this are the ability to analyze fire related information and its impacts from an end to end functional perspective and providing a key function of Wildland Fire while reducing and eliminating errors in that data, impacting many other processes and systems of the Wildland Fire Enterprise.

The existing environment of wildland fire systems and data created by various wildland fire business functions with differing objectives and requirements has created stove pipe systems where data is stored a variety of ways that are not cohesive in nature. While the community has made progress in the area of data integration, validation, and deconfliction for incident management data and initial attack resources, there is still a need for all data in all business areas to be accessible by varying components and contributors to wildland fire. Each system, on its own, benefits its target audience; however, when the data is consolidated for production purposes, there remains a large portion of the data that has not been validated or deconflicted. As the data collection progresses, different processes or applications add data to the environment. This data needs to be available in several states of maturity and during the process of its lifecycle, from initial acquisition, to combined with other data to create a record, processed through tools to create intelligence or predictions and ultimately archived to be available for queries, research, and analysis to impact future decisions.

The initial project scope was unknown and through Joint Application Requirments (JAR) sessions with dozens of stakeholders in November, 2017, December, 2017, and January, 2018, MBS worked with the Data Cache team to determine a list of ten (10) scoped items that will comprise the end to end vision of all focus areas for the Interagency Fire Data Cache (IFDC).

During the course of this analysis, it was determined that the deliverables for this analysis should change to more closely align with the business need and to provide more value to the business. The IFDC team along with MBS agreed to a course correction from the original statement of work that meant we took a step back from a very detailed analysis in order to define and evaluate the Data Cache project as a whole.

Therefore, the Requirements Analysis document includes:

- Tables with:
  - The business reason for inclusion of each of the ten goals identified for the IFDC
  - The current, as-is state in business terms for the each of the ten goals
  - The technical definition of each of the ten goals
  - The business use for each of the ten goals
- A list of non functional, General Requirements.

As well as some documentation that was created before the course correction where the IFDC team had originally prioritized four out of the ten goals of the Data Cache. The IFDC originally

prioritized the goals of Operational Data Store, Data Warehouse, Business Intelligence, and Reference Data Database. So, for those goals, we have provided:

- A Data Flow Diagram that depicts the way data will flow to and from each of the four originally prioritized areas of the IFDC
- High level Epics and User Stories for each functional requirement supporting the Data Flow Diagram for each of the four originally prioritized areas of the IFDC
- A supporting spreadsheet that identifies a system's interaction with the areas on the Data Flow Diagram

GOAL DEFINITIONS

1. Data Warehouse (DW)	
Inclusion Reason	The Data Cache team expressed a need for access to historical and archived fire related data for operational analysis, research, and long term storage needs. Also, users have concerns about the validity of the data they access, as there is no enterprise-wide data validation process in place.
As Is State	Currently, historical data is stored in a myriad of places, including agency specific data warehouses (FAMWEB), transactional systems' databases, third-party databases, individual user's hard drives, and physical filing cabinets. Compiling this data is tedious and time-consuming.

Technical Definition	• A Data Warehouse (DW) is a database that stores static data that is uploaded from the operational systems directly or data that passes through an ODS.
	• Data may also be migrated from other databases or data warehouses into the DW.
	• The DW integrates data from multiple sources into a single database and data model, enabling an enterprise-level view.
	• Data comes into the DW via an Extract, Transform, Load (ETL)-process that uses staging, data integration, and access layers to house its key functions.
	  • Extract - The staging layer or staging database stores raw data extracted from each of the disparate source data systems.
	  • Transform - The integration layer integrates the disparate data sets by transforming the data from the staging layer
	  • Load - The integrated data is then moved to the data warehouse database, where the data is arranged into hierarchical groups, often called dimensions, and into measures, or a star schema.

1. Data Warehouse (DW)	
	• The ETL process is done on a scheduled cadence, or a batch load, therefore the data is not available in real time like an ODS. • Data in the DW can be visualized using a Business Intelligence (BI) tool for dashboards and reporting. • It is designed to handle large, long-running, analysis queries. • It maintains data history. • It Improves data quality by providing consistent codes and descriptions, flagging or even fixing bad data. • It provides a single common data model for all data of interest regardless of the data's source. • It can restructure the data so that it makes sense to the business users; and delivers excellent query performance, even for complex analytic queries, without impacting the operational systems. • It can organize and disambiguate repetitive or duplicate data. • The data is synched to the DW, but data is not passed back to operational systems.
Business Use	• For the IFDC, the DW can provide a "one stop shop" for data to support long term historical analysis of interagency data. It will reduce the data analysts' time to create information and reduce the risk of data loss. It will also allow for more consistent reporting and allow important questions to be answered more quickly and with greater credibility than the community currently can. • The daily situation snapshots from the ODS that are sent to the DW will allow users to compare what they know today about an incident to what they knew about an incident on previous days

2. Operational Data Store (ODS)	
Inclusion Reason	The Data Cache team expressed a need for access to valid, real time, operational fire related data to facilitate communication and reporting more quickly from the initial phases of a fire through its lifecycle.
As Is State	Currently, teams who need to compile this data today are required to use a variety of methods to access data. Methods include email, phone calls, File Transfer Protocol (FTP) sites, direct access to systems, EGP access, and manual entry.

2. Operational Data Store (ODS)	
	Also, users have concerns about the validity of the data they access, as there is no enterprise-wide data validation process in place.
Technical Definition	• An Operational Data Store (ODS) is a type of database designed to integrate key operational data from multiple operational source system for additional operations on the data, for reporting, controls, and operational decision support. I.E. not all data from operational source systems is sent to an ODS; only the most valuable and informative data is sent. • It contains low level or atomic (indivisible) data elements with limited history that is captured "real time" or "near real time". • It contains data that is dynamically and constantly updated. • It provides short-term storage, and is in most cases an intermediate database before data is sent to a Data Warehoue (DW) • It can be seen as a “snapshot” of current data. As new data comes into the ODS, it may overwrite older data so that only the most current data is available for reporting. • It is designed so that it can quickly perform simpler queries on smaller sets of data. • The data is synched to the ODS, but data is not passed back to operational systems. • An ODS helps clean (transform) and organize data and ensure that it meets business and reporting requirements. Data from the operational source systems is less transformed than data in a Data Warehouse would be. • Most ODS data becomes source data for the Data Warehouse for reporting (DW); the identification and cadence of data passed to the DW is based on business rules. • The snapshots of prior day’s data is sent to the DW. • An ODS tends to focus on the operational requirements of a particular business process (for example, a current fire). • Data in the ODS can be visualized using a Business Intelligence (BI) tool for dashboards and reporting.
Business Use	• For the IFDC, the ODS will have the best and most recent, near real time data from source systems, such as CAD systems, IRWIN, and ROSS, in one centralized location so that users can visualize and report on complete data across applications in less time and more efficiently than they can today.

3. Business Intelligence (BI)	
Inclusion Reason	The Data Cache team expressed a need for a vehicle to visualize, create static reports, and do ad hoc reporting on fire related data to facilitate communication more quickly, as well as to compare data from other systems for analysis and research.
As Is State	Once data is compiled today, team members use whatever tools are available to them, including spreadsheets, word processing documents, and manual manipulation, to make the data presentable. The user community today has access to several BI tools today and a business decision as to whether to retain them all should be discussed.
Technical Definition	• The Business Intelligence (BI) tool is a type of application software designed to retrieve, analyze, transform and report on data that is available in the DW and the ODS, as well as data from external source systems and databases. • The BI tool allows users to choose specific data sets and feature services to report on or to create dashboard-type visualizations, static reports, as well as ad hoc reports. • BI tools provide historical, current and predictive views of business operations. • BI technologies can handle large amounts of structured and sometimes unstructured data. • They allow for the easy interpretation of this data. • BI tools are generally easy to use for line of business users and can compile data and create desired reports and visualizations quickly. • The BI tool does not validate that data is correct; this process should be handled upstream of the BI tool.
Business Use	• For the IFDC, all business subject areas may use the BI tool in different ways. For example, the Incident Command users may create a dashboard of mission critical data that they need for real time reporting. Resource and equipment users may use a BI tool to compare resource allocations across multiple fires.

4. Document Management System (DMS)	
Inclusion Reason	The Data Cache team expressed a need for a centralized repository for historical documents, images, databases, and system outputs, as well as other artifacts. These artifacts can be used for reporting support, research, and analysis.
As Is State	Most relevant artifacts are scattered across disparate systems, including systems where the wildland fire community may pay a third party for storage, as well as individual resource's local drives. There is no "one stop shop" for these artifacts, nor is there a way to use metadata to tag all artifacts for easy categorization and organization.
Technical Definition	• A DMS is a system used to track, manage and store documents and other artifacts. • A DMS is capable of version control, or keeping a record of the various versions created and modified by different users (history tracking). • A DMS stores metadata, such as date, name, and key words, and allows for indexing of data sets for ease of location and organization. • Artifacts in the DMS can be accessible via the BI tool. • A DMS allows for validation of upload and acceptance, to ensure that failures and errors are addressed. • A DMS has security, roles, and access permissions built in to ensure that artifacts are properly stored and accessed. • A DMS can have workflow functionality that allows users to create rules for upload, download, access, and approvals.
Business Use	• For the IFDC, all business subject areas may use the DMS in order to permanently store, retrieve, and report on, all relevant wildland fire-related artifacts. • The IFDC users could realize a cost savings by internalizing the storage of some data sets, such as modeling results and remote sensing data and images. • The IFDC users could retain management of their data sets. By having all fire-related artifacts in one easily-accessible system, the availability of the data will increase speed of communication and analysis. (e.g. IAPs, forecasts and ISuite Databases)

5. Spatial and Tabular Reference Data Database (RDDB)	
Inclusion Reason	The Data Cache team expressed a need for a centralized database that contains commonly and widely used data that can be considered reference data to ensure that all wildland fire users have access and use standard and approved data for reporting, communication, and analysis.
As Is State	Reference data today resides in multiple systems, is not standardized, and may not be under the management of the wildland fire community, meaning that users may use different representations of the same data, which can cause miscommunication and errors. Also, there are systems in use today to manage reference data, such as Unit Identifiers, that have a limited lifespan.
Technical Definition	• The RDDB will allow common, relatively static data, such as Jurisdictional Boundaries, Dispatch Locations, and Unit Identifiers, to be managed in one location. • Users will be able to access RDDB data for reporting and analysis, ensuring that all users are referencing a common, standard set of data. • The RDDB will allow for editing of data, including workflow and approvals, and auditability of changed data.
Business Use	• For the IFDC, all business subject areas may use the data in the RDDB for reporting and analysis. • Use in other source systems • Users with the correct access and permissions will be able to create, edit, update, and delete (CRUD) the data in the RDDB, per business rules. • The RDDB will be the single source of truth for data considered reference data.

6. Transactional Database Layer	
Inclusion Reason	The Data Cache team expressed a desire for the IFDC to operate as the back-end database for identified systems. In essence, the systems would have a user interface (UI) that would be the front-end of the system for users to interact with the database, but the IFDC would manage and operate the database layer where data is stored. MBS would recommend having an available common database layer that applications can adopt for use rather than having a single database accessed and shared by multiple applications. Sharing one database amongst multiple applications has some serious

6. Transactional Database Layer	
	disadvantages that are discussed in more detail in section 8.2 of the Alternatives Analysis document.
As Is State	Today, each system manages and operates their own back end database. There is one notable exception, where IRWIN is acting as the transactional database layer for the new INFORM system.
Technical Definition	• A transactional database layer would allow systems, as they change or modernize, to access an updated, managed database as the back end to a new or updated systems. • All new systems could leverage the database layer, meaning a cost-savings for development efforts. • All new systems would use standard, updated databases, allowing for simpler integrations, including reusable application programming interfaces (APIs) • This database layer establishes centralized control of all data schema changes and definitions which may impact application development design decisions.
Business Use	• As systems modernize and consolidate, having a standard database available to implement as systems are redeveloped could mean a significant cost-savings for the wildland fire community. • The wildland fire management programs would have purview over ensuring that new systems are developed in an efficient way.

7. Identify Authoritative Data Sources for Incident, Resource, and Treatment Data	
Inclusion Reason	The Data Cache team expressed an interest in the Data Cache becoming the Authoritative Data Source for tabular and geospatial Incident, Resource and Treatment Data so that they have a clear source of data for downstream processing and reporting.
As Is State	Which applications own which data is unclear in today's environment. There has been an existing strategy for Wildland Fire systems to not own data but rather be stewards for it. IRWIN has been taking on an increasing Authoritative Data Source responsibility for Incident Data.
Technical Definition	• This concept involves which applications own which data fields such that downstream systems and reporting can trust that data to contain the truly recorded, audited (record of when the data was changed over time), and QA/QC'd value for data.

7. Identify Authoritative Data Sources for Incident, Resource, and Treatment Data	
	• The Data Cache can act as this data owner; or individual Source Systems can take on this role.
Business Use	• This concept pushes data conflict resolution as far upstream as possible, so that all systems downstream and reporting do not have to resolve data conflicts. • This concept creates data ownership in the current systems. • As application data owners are identified, duplicate systems may be identified and targeted for retirement, eliminating maintenance costs for the duplicate systems. • Existing systems may be able to decrease an expanded scope and/or reduce the footprint of the system, allowing a system to focus on its intended purpose and role in the enterprise's environment.

8. Quality Control/Quality Assurance of Data (QA/QC)	
Inclusion Reason	The Data Cache team expressed a need to ensure data integrity as data is passed to the DW and ODS and visualized using the BI tool.
As Is State	Currently, each system and/or agency is responsible for ensuring that their data is valid. There is no overarching method for quality control and assurance as data is being used for mission critical reporting and communication.
Technical Definition	• A QA/QC process will ensure data completeness and data consistency, by validating data against expected results. • A QA/QC process may include data reviews, approvals and workflow via a user interface. • A QC/QC process may include automated validations and checks. • The QA/QC plan should be well documented and remain stable over time.
Business Use	• For the IFDC, all business subject areas need to be able to trust that the data in the ODS and DW is reliable, valid, and accurate. • Upstream systems will have to do their own QA/QC, and downstream systems will have to trust the upstream systems' processes.

9. Data Integration Service Layer	
Inclusion Reason	The Data Cache team expressed a desire for a system to handle and manage two-way integration of systems for all fire -related data, much the same as IRWIN does today for incident data.
As Is State	Today, the access to incident data for most systems is handled by IRWIN as the data integration service. Incident data, such as a fire code, is brokered among systems, where business rules and the authoritative data source as data is changed is managed by IRWIN. However, there is no similar data brokering that happens for other fire-related data, such as resource and weather data.
Technical Definition	• A data integration service layer uses service-oriented architecture to create services that have protocols to pass and parse data using metadata. The data is then available to other systems via the services. • A data integration service layer processes information that originates from various sources then passes that information back out to other systems. • The data integration service layer eliminates dozens of system to system interfaces and individualized interface requirements for individual systems.
Business Use	• Data other than incident data, such as weather and resource data can be brokered among applications that need the data, including transactional systems, such as ROSS and WIMS, as well as other systems, such as IRWIN, EGP and the DW and ODS. • The fire community can track the changes to data via metadata that is provided by the broker. • The fire community can ensure data integrity by tracking the interagency authoritative data source (IADS) of the data.

10. Migrate Historical and Legacy Data	
Inclusion Reason	The Data Cache team expressed a need for historical and legacy data that is currently held in disparate systems to be migrated to a centralized location for archival, reporting, and analysis. If all relevant data were stored and managed by the IFDC, the community would have better access and better management capabilities.
As Is State	Today, each system and/or agency manages and operates their own archiving. However, not all users have access to the historical data. Also, there is historical data, such as RAWS weather data, on third-

10. Migrate Historical and Legacy Data	
	party servers where the wildland fire community contracts for this storage for a fee.
Technical Definition	• Historical and legacy data migration is usually handled on a case-by-case basis and is automated where possible. • Data migration occurs for a variety of reasons, including server or storage equipment replacements, maintenance or upgrades, or application migration. • Data from the old system is mapped to the new system via a data extraction and then data loading, where data is then validated and verified before becoming accessible to users.
Business Use	• As historical data is identified for inclusion in the DW or the DMS, the responsible parties will coordinate the migration of the data. • The IFDC team would have purview over ensuring that migrated data meets standards.

Application and System Data Cache Use

Before the previously discussed course correction, MBS had started an analysis of systems identified in the Application Inventory project as used and needed by the interagency fire community. This analysis endeavored to categorize for each Tier One and Tier Two application, whether that system potentially would:

- Contribute data to the Data Warehouse
- Contribute data to the Operational Data Store
- Directly consume data from the Data Warehouse (export from Data Warehouse)
- Directly consume data from the Operational Data Store (export from Operational Data Store)
- Contribute items to the Document Management System
- Be an external input to a Business Intelligence tool
- Their users would utilize the Business Intelligence tool
- Contribute data to the Reference Data Database
- Directly consume data from the Reference Data Database

Although this analysis was not completed, MBS believes that when the IFDC team is ready to begin work on the IFDC, they will find this information useful as a starting point for further analysis. Therefore, we are including the below incomplete table as part of this deliverable. We will also include the original Excel spreadsheet for the team's use. Lines in yellow indicate that a system's potential use still needs to be discussed with a subject matter expert.

System Acronym	System Name	Integration Requirement	ODS Data Input	DW Data Input	ODS Data Consumer	DW Data Consumer	DMS Input	External BI Input	BI User	Reference Data Input	Reference Data Consumer
AFF	Automated Flight Following	AFF is hosted on the DOI IAS LAN. The AFF system is a ISS front end and SQL backend with load balancers and ssl excelleraters.	X								
AWIPS II/FXCave	Advanced Weather Information Processing System/FXCave	Workstations, Virtual Server, Speed of data downloads is dependent on the network speed.			X	X			X		
BehavePlus	BehavePlus Fire Modeling System	Windows Server 2003	X	X			X		X		
BLM IT Program Costs	BLM IT Program Costs		NA								
Data Cache	Common Data Cache available to Wildland Fire Information Systems		NA								
EGP	Enterprise Geospatial Portal		X		X	X		X			
e-Isuite	Electronic Incident Suite		X	X					X	X	
ESRI Collector	Collector for ArcGIS		NA								
FAMWEB Data Warehouse	FAMWEB Data Warehouse			X					X		
FEIS	Fire Effects Information System						X				
FEOS	Fire Equipment Ordering System		NA								
FEPMIS	Federal Excess Property Management Information System		NA								
FF+	FireFamilyPlus			X		X	X		X	X	
FFI	FEAT-FIREMON Integration			X		X	X		X	X	
FGX (part of EGP)	Fire Geospatial Exchange		NA								
FIMT (end of life)	Fire Incident Management Tool		NA								
Fire GIS Services	Fire GIS Services		NA		X				X		
Fire NESS-Infrastructure	National Enterprise Support Service		NA								

System Acronym	System Name	Integration Requirement	ODS Data Input	DW Data Input	ODS Data Consumer	DW Data Consumer	DMS Input	External BI Input	BI User	Reference Data Input	Reference Data Consumer
FireCode (handled by IRWIN)	FireCode		NA								
Firefly (concept)	Fire Flight Usage Reporting		X								
Firenet.gov	Firenet.gov		NA								
FIRESTAT now INFORM	Fire Statistics System		NA								
FlamMap	FlamMap				X	X	X				
FLIGHT	FLIGHT		X						X		
FMIS	Fire Management Information System			X							
INFORM	Interagency Fire Occurrence Reporting Modules		X	X					X		
FPA	Fire Program Analysis		NA								
FS IT Program Costs	FS IT Program Costs		NA								
FTEM	Fuels Treatment Effectiveness Monitoring			X		X	X		X		
FUSE (concept)	Fire Equipment Use Application		NA								
ICBS	Interagency Cache Business System		X	X	X				X		
IFTDSS	Interagency Fuels Treatment Decision Support System			X		X	X		X		
InciWeb	Incident Information Web			X	X	X	X		X		
Interagency Incident Helpdesk	Interagency Incident Helpdesk		NA								
Interagency IT Projects	Interagency IT Projects		NA								
IQCS	Incident Qualifications and Certification System								X		
IROC	Interagency Resource Ordering Capability		X	X					X	X	X
IRWIN	Integrated Reporting of Wildland Fire Information		X						X	X	X

System Acronym	System Name	Integration Requirement	ODS Data Input	DW Data Input	ODS Data Consumer	DW Data Consumer	DMS Input	External BI Input	BI User	Reference Data Input	Reference Data Consumer
LANDFIRE	Landscape Fire and Resource Management Planning Tool							X		X	
LCMS	Learning Content Management System								X		
LDDS	Lightning Detection Data Services		NA								
Lessons Learned Site	Lessons Learned Site						X				
MODOC	Modification Documentation						X		X		
MTBS	Monitoring Trends in Burn Severity			X			X		X		
NAMS	NIFC Asset Management System									X	
NFMD	National Fuel Moisture Database		X								
NFPORS	National Fire Plan Operations and Reporting System			X		X			X		X
NIFC e-NET Infrastructure	NIFC e-NET Infrastructure		NA								
NWCG IT Projects	NWCG IT Projects		NA								
OLMS	Operational Load Management System										
PSGP	Predictive Services Geospatial Web Portal		NA								
RAMS	RAMS			X					X		
ROSS (now IROC)	Resource Ordering and Status System		NA								
SAFECOM	Aviation Safety Communique		NA								
SAFENET	SAFENET						X		X		
SIT	National Situation Report								X		
209	ICS-209		X						X		
TIPP	Transforming Incident Procurement Processes		NA								
WebCAN	Web based Computer		X								

System Acronym	System Name	Integration Requirement	ODS Data Input	DW Data Input	ODS Data Consumer	DW Data Consumer	DMS Input	External BI Input	BI User	Reference Data Input	Reference Data Consumer
	Aided Navigation										
WFAS	Wildland Fire Assessment System					X	X		X		
WFDSS	Wildland Fire Decision Support System		X	X	X	X	X		X		X
WFIPS	Wildland Fire Investment Planning System				X	X			X		X
WFIT Staff & Services	WFIT Staff & Services		NA								
WFMI	Wildland Fire Management Information - All Modules		NA								
WFMI-Lightning	Wildland Fire Management Information - Lightning Module		X						X		
WFMI-Unit ID	Wildland Fire Management Information - NWCG									X	
WFMI-Weather	Wildland Fire Management Information - Weather Module		X						X	X	
WildCAD	Wildland Computer Aided Dispatch		X	X					X		X
WIMS	Weather Information Management System		X						X	X	
VIPR (tier 3)	Virtual Incident Procurement		X	X							
Any other tier 3 or 4 apps											
GeoMAC			X		X						
BOISE											

Data Cache Data Flow Diagram

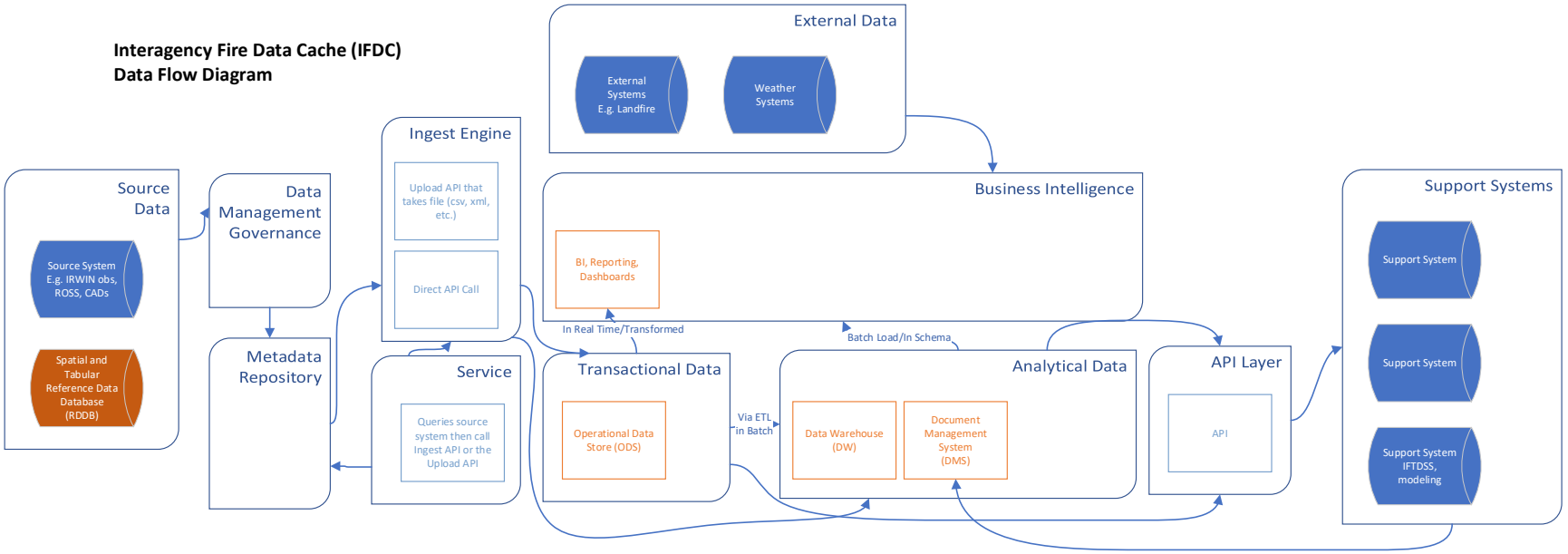
As part of this analysis, MBS initially created a data flow diagram that depicts how data might flow through the system if the Data Warehouse, Operational Data Store, Document Management System, Business Intelligence Layer, and the Reference Data Database goals of the data cache were implemented, as originally these goals were listed as priorities for the Data Cache. We are including the diagram here as a reference for future analysis, should the IFDC team decide to move forward with these goals.

- Starting on the left of the diagram, there are source systems, meaning fire-related applications that will provide data for downstream use. These systems include for example, CAD systems, ROSS (IROC), and IRWIN Observer data.
- Also on the left of the diagram is the Spatial and Tabular Reference Data Database (RDDB) that is the Interagency Authoritative Data Source (IADS) for fire-related data. This RDDB will also be a source system to provide data for downstream consumption and use.
- The source systems should be subject to data management standards and governance before the data is consumed to ensure reliability of data.
- After the data management process is complete, pre-defined metadata can be stored about the data. The metadata repository can include business data such as meaning and content, policies that govern the data, business rules, technical attributes, specifications that transform the data, and programs that manipulate the data.
- Then, depending on the source system that is providing data, one of three ingest methods can be used:
    - The simplest method is a direct API call to get data. This method can be used for modern systems and architectures.
    - The next method is to create an upload API to get data in specific formats, such as .csv, or xml. This method is helpful for systems that can export in those formats but are not able to utilize a direct API call.
    - The most difficult method is to write a service that queries the source system, then calls the direct API or upload API to ingest the data. This method is recommended for systems in older architectures and platforms where the first two methods are not feasible.
- Once data is available for ingest, some data may flow directly to the Operational Data Store (ODS) for immediate use and other data may be loaded to the Data Warehouse (DW) in a batch load.
    - Data that first goes into the ODS may be sent to the DW in a batch load once it is no longer operational data.
- The Document Management System will be populated with data, images, documents, and other data sets as determined.
- There is also valuable data in other systems that should be available for visualization and analysis in the Business Intelligence tool. Some systems are owned and managed by interagency fire organizations, such as LandFire, and some systems are external to fire, such as census data. However, regardless of system ownership, the reason this data is considered external data is generally due to the size of the data sets and the desire to not

store duplicate sets of data that could cause data integrity issues. There is no business value in storing this data in the Data Cache DW as well ain n the orginal system.

- Data from the DW, the ODS, and from the DMS, along with other data from external systems will be available for reporting and analysis by using a Business Intelligence (BI) tool.
- The API layer on the right side of the diagram is used for extracting information from the ODS, DW, and DMS for direct use in support systems, such as IFTDSS. This is primarily for data that needs to be used directly in the support system. However, Maybe: However, if there were a data integration service layer, such as IRWIN, it is more likely the data from source systems would be brokered directly from source system to support system rather than an API layer as depicted on this diagram.

Interagency Fire Data Cache (IFDC)
Data Flow Diagram



EPICS AND USER STORIES

Before this project had a course correction, MBS had begun writing epics and high level user stories for the originally prioritized goals of the data cache. Because this effort was paused in the course of the project, the table below is in an incomplete status. However, we are including the table below for future review if the IFDC team decides to pursue development.

The Epic column is generally based on the goal of the Data Cache. The original requirements came from the original performance work statement for the project.

Epic	Original Requirement	User Story	Notes/AC	Out of Scope
Business Intelligence	Users can create individual and group use dashboards and reports from individual or linked datasets Capability to serve stored data for viewing, editing, replicating, and reporting purposes Capability to access historical records	As a BI user, I want to create dashboards for use individually or by a group from individual or grouped data sets and feature services from the DW, ODS, RDDB, and external sources, so that I can visualize the data in a meaningful way.		
		As a BI user, I want to create reports for use individually or by a group from individual or grouped data sets and feature services from the DW, ODS, RDDB, and external sources, so that I can report on the data in a meaningful way.	Before the fact research, after the fact research, operational research, store model outputs	
Business Intelligence	BI System allows for ad hoc queries	As a BI user, I want to create ad hoc queries for use individually or by a group from individual or grouped data sets and feature services, so that I can visualize the data in a meaningful way.		
Business Intelligence	BI tool needs to allow for spatial analytics by utilizing feature services from ESRI/AGOL	As a BI user, I want to utilize feature services from ESRI/AGOL, so that my analysis can include spatial information from these feature services.		

Epic	Original Requirement	User Story	Notes/AC	Out of Scope
Data Warehouse	Capability to link datasets within the solution Capability to integrate stored data so that there is one feature class or equivalent per dataset	As an IFDC DW user, I want to link desired data sets in the Data Warehouse, so that I can have access to all related data for comparison and reporting.		
Data Warehouse	Capability to serve stored data for viewing, editing, replicating, and reporting purposes Capability to access historical records	As the DW, I want to batch load data via ETL from source systems, so that users can report on the data using the BI tool.	editing?	
		As the DW, I want to batch load data via ETL from the ODS once data is no longer operational, so that users can still report on the data.		
		As the DW, I want to load cached data from external systems on a pre-determined cadence, so that users can report on the data using the BI tool.		
Data Warehouse	Capability to download tabular and spatial data	As an IFDC DW user, I want an API available for exporting tabular and spatial data in a usable format, so that I can process the data in external systems.	DW data into Support Systems in batch	
Data Warehouse	Solution will have backup/redundancy capabilities	As an IFDC DW user, I want all data in the DW backed up on a fully redundant server on a prescribed schedule, so that I can be sure the data is never lost.		
Data Warehouse	Capability to track database transactions	As an IFDC DW user, I want the system to keep an audit record of all database transactions, so I can have traceability and ensured data integrity.		
Data Warehouse	DW can extract information from heterogeneous source systems, tranform data into	As the DW, I want to perform ETL on data from designated source systems on a prescribed cadence, so		

Epic	Original Requirement	User Story	Notes/AC	Out of Scope
	warehousable format, and load into DW.	that data is available in the DW as anticipated.		
Data Warehouse	For discrete datasets (ex. Fire report), store one record per event regardless of agency			
Document Management System	DMS artifacts into Support Systems in batch			
Document Management System	document repository needs to have access/roles/permissions	As an IFDC DMS admin, I want to add users to the systems, assign roles, and designate permissions for system access, so I can ensure that all users have appropriate access to the documents		
Operational Data Store	Capability to link datasets within the solution Capability to integrate stored data so that there is one feature class or equivalent per dataset	As an IFDC ODS user, I want to link desired data sets from the Data Warehouse and Operational Data Store, so that I can have access to all related data for comparison and reporting.		
Operational Data Store	Capability to serve stored data for viewing, editing, replicating, and reporting purposes Capability to access historical records	As the ODS, I want to load data from source systems on a near real time cadence, so that users can visualize and report on the data using the BI tool.		
Operational Data Store	Capability to download tabular and spatial data	As an IFDC ODS user, I want an API available for exporting tabular and spatial data in a usable format, so that I can process the data in external systems.	ODS data into Support Systems in real time	

Epic	Original Requirement	User Story	Notes/AC	Out of Scope
Operational Data Store	Solution will have backup/redundancy capabilities	As an IFDC ODS user, I want all data in the ODS backed up on a fully redundant server on a prescribed schedule, so that I can be sure the data is never lost.	Question: What is the requirement for ODS backup, since data is near real time. I think the better question is how often is data moved to DW?	
Operational Data Store	Capability to track database transactions	As an IFDC ODS user, I want the system to keep an audit record of all database transactions, so I can have traceability and ensured data integrity.		
Operational Data Store	ODS can extract information from heterogeneous source systems, tranform data into ODS format, and load into ODS	As the ODS I want to perform ETL on data from designated source systems on a real time cadence, so that data is avaliable in the ODS as anticipated.		
Out of Scope	Users can upload artifacts in defined formats to DMS	As a DMS user, I want to upload supporting artifacts and assign defining metadata, so that I can ensure the artifacts are matched with corresponding data sets		
Reference Data Database		As		
Source System Integration	Capability to store spatial and tabular data	As a source system user, I want to extract my analytical data from my source system in a spatial format, so that users can access geospatial information as it relates to my data		

Epic	Original Requirement	User Story	Notes/AC	Out of Scope
		As a source system user, I want to extract my transactional data from my source system in a spatial format, so that users can access geospatial information as it relates to my data		
		As a source system user, I want to extract my analytical data from my source system in tabular format, so that users can access tabular information about my data.		
		As a source system user, I want to extract my transactional data from my source system in tabular format, so that users can access tabular information about my data.		
Source System Integration	Capability to perform QA/QC processes of the stored data	As a source system, I want data that is sent to the Data Warehouse goes through quality checks, so that DW users are confident the data is valid and accurate.	is this the same as cleaning?	OOS
		As a source system, I want data that is sent to the Operational Data Store goes through quality checks, so that ODS users are confident the data is valid and accurate.	QC function is a function of the data management committee. Do we need to to add this to the cache requirements?	OOS
Source System Integration	Capability to validate that the data is being loaded in NWCG standards	As a source system, I want appropriate business rules applied to my data that is sent to the Data Warehouse, so that I can ensure the data meets NWCG data standards.		OOS

Epic	Original Requirement	User Story	Notes/AC	Out of Scope
		As a source system, I want appropriate business rules applied to my data that is sent to the Operational Data Store, so that I can ensure the data meets NWCG data standards.		OOS
Source System Integration	Capability to do deconfliction of Spatial and Tabular Data using defined business rules	As an IFDC DW or ODS user, I want to my spatial data to be cleaned (deconflicted) according to business rules applied, so that the geospatial data I use is accurate and valid.		OOS
		As an IFDC DW or ODS user, I want to my tabular data to be cleaned (deconflicted) according to business rules applied, so that the tabular data I use is accurate and valid.		OOS
Source System Integration	Capability to upload tabular and spatial data	Question: Should a user be able to upload directly to DW? Is that a requirement?		OOS
Source System Integration	Capability to track and report edits made to spatial and tabular data	Question: Should a user be able to edit data in the DW or ODS?		OOS
System	Capability to control users/systems ability to Create, Read, Update, and Delete Capability to control users/systems access to services	As an IFDC DW, ODS, RDDB or BI admin, I want to add users to the systems, assign roles, and designate permissions for system access, so I can ensure that all users have appropriate access to the data.		
System	Capability to notify agency data stewards of conflicts	As an IFDC data steward, I want to be notified of data conflicts, so that I can ensure that data changes are made correctly.		
System	Solution will have a SLA to match business needs	As an IFDC DW, ODS, RDDB or BI admin, I want to define SLAs for performance, so that I can ensure the		

Epic	Original Requirement	User Story	Notes/AC	Out of Scope
		systems meet the user's needs for data availability.		

GENERAL REQUIREMENTS

This Appendix provides insights regarding the overall Information Technology (IT) strategy provided by the OCIO offices at the Department of Agriculture / Forest Service and the Department of Interior as well as other stakeholders.

Attendees

The following is a list of attendees who participated in the IT Strategy discussion:

[placeholder for attendee names and titles]

DEVELOPMENT ENVIRONMENT / TOOLS

There are no set standards for development environments and tools for application development. The Wildland Fire Program does not wish to add new tools to the technology stack and may seek to draw down the list of tools in use. The Wildland Fire Program hopes to find a balance between allowing vendors to offer solutions in a variety of tools with a need to minimize the number of tools in use to keep Operations & Maintenance costs in check.

Also, the OCIO offices would like to leverage Software as a Service (Saas) and whenever possible to leverage Custom Off-the-Shelf software and reduce custom development solutions and maintenance costs.

Java and .NET are acceptable development tools in the current environment.

The following Business Intelligence tools are in use or have been considered by the Wildland Fire Program:

- *Tableau*
- *ESRI*
- *Power BI*
- *Cognos BI*
- *OBIEE (Oracle)*

The following Extract, Transform, Load and Statistical tools are in use or have been considered by the Wildland Fire Program:

- *SAS*
- *FME*
- *R*

Fire applications need to be able to run in a Windows 7 and a Windows 10 environment. Windows 8 is not prevalent. Fire applications should be able to run on the following web browsers, particularly in the case of supporting non-Federal users:

- Chrome,
- Edge,
- Firefox,
- Safari and
- Internet Explorer.

Mobile application operating system requirements should include:

- iOS and
- Android,
- But do not need to include Windows.

Low code platforms provide the ability to create applications through a Graphical User Interface and are intended to decrease the amount of technical support required to build and maintain applications. For example, the IROC system is being built in ServiceNow and Volunteer.gov for the National Park Service (NPS) is being built in Salesforce. Salesforce is also being considered for other fire applications. The following low code platforms have been considered most frequently for enterprise application development.

- Salesforce
- ServiceNow

SECURITY

One of the major considerations for Wildland Fire IT systems is the requirement for Federal and non-Federal agencies to access these systems. There is a user authentication service called login.gov which has become available. Login.gov can be contacted by an application for user authentication, during which login.gov takes care of the user authentication screen and actions and then returns control to the calling application. Login.gov is intended to allow for both Federal and non-Federal user access, similar to pay.gov. Login.gov has tremendous potential to break down the barriers of wide ranging access required of Wildland Fire Systems. More information about login.gov can be found at https://login.gov.

Security Authentication Markup Language (SAML) is another user authentication service used in Federal Systems. This service works well for eliminating user authentication barriers between agencies in the Federal Government, but does not extend to non-Federal users.

It is expected that any systems connecting to other systems will leverage Internet Protocol Version 6 (IPv6) for modern standards for IP addressing. IPv8 is also being

rolled out and should be considered for system connectivity. This is described at this website:

https://www.gsa.gov/technology/technology-products-services/cybersecurity-products-services/internet-protocol-version-6-ipv6.

It is expected that all fire systems will adhere to FISMA requirements, including the handling of Privacy, Person Identification Information and Standards compliance. All security requirements are subject to a review with a Privacy Officer.

USABILITY

Wildland Fire IT would like to establish a common branding for all wildland fire applications. Collaborative discussion will be required to accomplish this, and systems will need to evolve to the common branding.

There are several government tools and references for usability guidelines, such as:

https://methods.18f.gov/

and

https://www.usability.gov/how-to-and-tools/index.html.

508 compliance is required for all systems; the DOI communication team oversees the standards for DOI-managed applications. These standards shall be applied at the time of development -- reference

https://www.hhs.gov/web/section-508/index.html for more information.

Any new and existing records management policies should be reviewed and followed during development of new systems.

SCALABILITY

It is the general strategy of the OCIO offices to move toward Container-Based Platforms for future application architecture. Containers allow sections of functionality to be combined in such a way that it can be moved between servers easily, without breaking its connection to other sections of functionality (integration of functions as opposed to assimilation of functions).

Many of the existing applications have three tier architectures (user interface, business logic and database layer).

The goal is "Cloud First", or to evaluate moving data processing to the Cloud. The OCIO offices would like to leverage existing cloud services, such as Microsoft Azure, Amazon Web Services, or a government provided cloud if it acts a cloud solution. The goal is for government agencies to reduce server, operating system and platform tools maintenance and allow cloud providers to provide these services at lower cost.

Leveraging Platform as a Service Solutions is also highly desired by the OCIO offices, as these offerings allow flexible pricing, scaling and performance for application needs.

RELIABILITY

- Systems supporting fire operations directly need to be available 7 x 24 x 365 and have very high reliability during fire season.
- Fire systems not used for fire operations should be fully operational between 8:00 AM and 8:00PM Eastern Time (ET).
- Systems should operate with 99.99% reliability.
- Further specifics and maintenance windows will be dictated by individual Service Level Agreements (SLA's).
- All issues will be documented, reviewed and assessed for severity.

MAINTAINABILITY

To allow for future maintainability, the documents listed below shall be provided with any of the Data Cache goals during development. Current documentation standards have a set of development documentation deliverables that are more appropriate for a waterfall development approach. It is the IFDC team's desire to develop using an Agile Methodology; cases where an equivalent Agile deliverable replaces a current standard deliverables are indicated to ensure coverage.

- User documentation: user documentation may be wiki-based or on-line and maintainable by IFDC staff. It shall be sufficient to train new users on the tool and business process.
- Product Backlog consisting of User Stories: this replaces the current DOI standard of a Systems Requirements Specification (SRS), Interface Requirements Specification (IRS) and Software Design Document (SDD)
- Source Code, Unit Tests, and QA Test Cases
- Entity Relationship Diagram (ERD): this replaced the current DOI standard of a Data Management Plan (DMP)
- Requirements Traceability Matrix (RTM)
- Application Administrator's Manual
- Development Modernization and Enhancements (DME)

Any custom developed source code will be owned by the sponsoring agency.