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

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

Table of Contents

PROJECT OVERVIEW

1.1 Project Description

The Office of Wildland Fire (OWF) organization engaged Managed Business Solutions, LLC (MBS), a wholly owned subsidiary of Sealaska Corporation, to gather requirements and document the business processes for an Interagency Fire Data Cache (IFDC or Data Cache).

As part of our project scope, MBS will provide the IFDC team with two requirements deliverables, in addition to this document.

- IFDC Section I - Requirements Document, which outlines and defines the ten identified goals for the Data Cache
- IFDC Section II - General Requirements

This document details the Alternatives Analysis we conducted, as well as our findings. MBS structured this document to consider each of the ten goals of the Data Cache discovered during stakeholder interviews independently at first, then secondly as a consolidated solution. MBS designed this document to provide information to the National Wildfire Coodinating Groug (NWCG) Data Mangagement Committee (DMC) and Wildland Fire Information and Technology (WFIT) Program Board that will allow them to make an informed decision on the path forward for the creation and implementation of the Data Cache.

As part of this analysis, MBS conducted Joint Application Requirements (JAR) workshops over fifteen days with over 100 participants from November, 2017, to January, 2018. This multi-stakeholder approach resulted in highly engaged discovery, with representatives from Department of Interior (DOI), United States Agricultural Agency (USDA), Bureau of Land Management (BLM), Bureau of Indian Affairs (BIA), National Park Service (NPS), United States Fish and Wildlife (USFW), United States Forest Service (USFS), National Weather Service (NWS), National Association of State Foresters (NASF) and the United States Geological Survey (USGS), among others.

1.2 Background

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, meets the needs of 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. Rather than independent databases, core fire program information should be available and able to be updated from the Interagency Authoritative Data Source (IADS) by other applications. 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.

Since the late 1970s, information technology has been used by the wildland fire (WF) community in processing information in order to assist their decision making and to support operations and other business decision areas. Technology has evolved tremendously since that time, but most of the evolution occurred within the agency boundaries without giving the needed attention to operations on an interagency level. To compound this situation, numerous requirements have been instituted that have a direct impact on the use and handling of information technology (IT) systems. These requirements increased on an expediential level shortly after September 11, 2001. Most of the requirements imposed were IT security controls. Agencies attempting to implement these controls did so according to agency policy, but soon discovered that cross-platform integration of systems and data sharing was impeded.

1.3 Problem Statements

- The scattered nature of data across the wildland fire community has led to limited accessibility, duplicative data collection, incomplete datasets, and inconsistent data and quality standards. Currently, no single database exists to handle the data needs of the interagency fire community. Wildland fire organizations produce, consume, and maintain various data layers to support individual applications without a strategy for central storage or capability to integrate and effectively use data at a national scale. Creating landscape scale datasets for support of enterprise business operations or reporting can be a daunting task that requires social networks, data mining, data conversion or transformation, and extensive quality assurance efforts. These efforts often consume substantial time and resources. The need to interconnect systems has grown as well over recent years. Wildland fire communities have recognized the advantages of linking diverse systems together in order exchange data and operate in a more cohesive manner.
- The systems in use today across the wildland fire community vary in architecture, age, status, and usability. Some systems are cloud-based, spatially-enabled, modern and highly usable, while others are outdated, with limited accessibility and lifespan. All systems, however, contain either valuable data or needed functionality that is used throughout the community. Despite recent efforts to take system integration into consideration when developing new systems, dissimilar IT infrastructures among agencies, and the policies that govern the use of individual agency infrastructures, has added complexity to interconnecting systems. A common, interagency-wide strategy for interconnecting systems is lacking as is the ability to centrally host access to data in various forms of its lifecycle.
- Should these issues fail to be resolved, the following are expected ramifications/risks:

a. Fire applications will continue to develop their own data management and data hosting strategies at a higher, more unpredictable cost.
b. Innovative applications will be unable to meet goals because they must continually expend budget on stand-alone data management schemes rather than application functionality.
c. The fire community will struggle to meet required federal open data initiative standards.

1.4 Project Drivers

There are several key drivers behind the wildland fire community's need to improve, extend, and make consistent, how data is collected, accessed, used, and reported on. The community would like a solution that considers the following:

- Accurate and timely entry and retrieval of data
- Reduced repetitive data entry and retrieval of data
- Provision for data clean-up
- Ability to write once, read many
- Quicker access to data located in diverse systems
- Archive of data, documents, images and other items
- Geospatially-enabled data across systems

The current wildland fire systems that are in place today have several pain points, including:

- Some systems are built on outdated system architectures that make integrating difficult
- There is no centralized data warehouse for long-term storage of needed data sets
- There is no centralized document repository for long-term storage of needed documents, images, videos, etc.
- There is no data integration service/data broker layer for resource data, fuels treatments, or fire environment data like there is for incident data using Integrated Reporting of Wildland-Fire Information (IRWIN). It should be noted that adding these data areas to the scope of IRWIN is planned.
- There is no centralized database for storage and retrieval of commonly used geospatial and tabular reference data
- There is no quality assurance or quality control over fire-related data such that users are confident that the data found in systems is authoritative and reliable
- Not all fire-related systems have the budget and/or IT personnel to maintain their own databases and/or user interfaces
- Many reporting functions are performed manually, with crucial data and outputs being kept in spreadsheets and on shared drives
- Crucial reports can take days or weeks to assemble information, validate data, and create the report
- Data collaboration among users is limited due to system constraints
- Incident command users have limited visibility into critical data

- Redundant data is common
- No way to aggregate data for comprehensive analysis or reporting
- Many systems are in need of a refresh, for either technology, contracting, support, platform reasons

1.5 Analysis Team

The following individuals comprise the analysis team. They are responsible for the requirements, analysis and creation of the IFDC Project.

Name	Agency	Job Title	Project Role
Roshelle Pederson	DOI	Data Management Specialist	Business Lead
Cole Belongie	FS	Data Integrator	Business Lead
Chris Markle	WFIT	WFIT Enterprise Architect	Project Manager
Lisa Elenz	FS	Assistant Director, Capabilities, Development and Integration	Project Sponsor
Rhonda Toronto	BLM	Branch Chief, BLM Fire & Aviation IT	Project Sponsor
Andrew Bailey	DOI	Data Manager	Subject Matter Expert
Andy Gray	FS	Project Manager	Subject Matter Expert
Andy Kirsch	NPS	Wildland Fire Management Analyst	Subject Matter Expert
Ann McDonough	FS	Contracting Officer Rep (COR)	Subject Matter Expert
Beth Spencer	FS	Project Manager	Subject Matter Expert
Bill Fletcher	FS	Assist Center Manager	Subject Matter Expert
Bill Yohn	NPS	Equipment and Facilities Programs Manager	Subject Matter Expert
BJ Glesener	NASF	Intelligence Coordinator	Subject Matter Expert
Bob Roth	FS	Aviation Mangagement Specialist	Subject Matter Expert
Brian Henry	BLM	National Predictive Services Assistant Program Lead	Subject Matter Expert
Cameron Tongier	FWS	GIS Analyst	Subject Matter Expert
Catherine Costello	USGS	GEOMAC	Subject Matter Expert

Name	Agency	Job Title	Project Role
Christine Schuldheisz	FS	Public Affairs	Subject Matter Expert
Chuck Wamack	DOI	Operations Business Lead	Subject Matter Expert
Clint Cross	FS	Applied Fire Ecologist	Subject Matter Expert
Craig Amundson	Contractor	Project Manager, Cask LLC	Subject Matter Expert
Craig Morgan	Contractor	Sr Consultant / Program Manager	Subject Matter Expert
Craig Thompson	DOI	Geospatial Specialist	Subject Matter Expert
Crystal Stonesifer	FS	Biological Scientist - Human Dimensions	Subject Matter Expert
Dan Buckley	NPS	Branch Chief of Wildland Fire	Subject Matter Expert
Dan O'Brien	FS	Northwest Coordination Center Manager	Subject Matter Expert
Darin Crisp	FS	Enterprise Security Specialist	Subject Matter Expert
Dave Haston	FS	Branch Chief Equipment and Chemicals	Subject Matter Expert
Diane Trethewey	FS	Mathematician	Subject Matter Expert
Dianna Sampson	BLM	GIS/Data Analysis	Subject Matter Expert
Ed Delgado	BLM	Predictive Services, National Program Manager	Subject Matter Expert
Emmy Harbo	Contractor	Data Management Specialist	Subject Matter Expert
Erik Torres	NPS	Information Resource Manager	Subject Matter Expert
Evan Mosby	Contractor	IRWIN Technical Lead	Subject Matter Expert
Frankie Romero	FS	Fire Use & Fuels Management Specialist	Subject Matter Expert
GaBriella Branson	AK DNR	Intel Coordinator	Subject Matter Expert

Name	Agency	Job Title	Project Role
Genevieve Giaccardo	DOI	Communications Specialist	Subject Matter Expert
Gina Papke	FS	Data Integrator	Subject Matter Expert
Greg Peterson	BLM	Supervisor - Fire and Aviation Systems Development	Subject Matter Expert
Henry Bastian	DOI	Project Manager	Subject Matter Expert
Heraclio Jaquez	BLM	Supervisory IT Specialist	Subject Matter Expert
Isaiah Hirschfield	FS	National Fire Desk Manager	Subject Matter Expert
James Silverstone	FS	Eastern Area Coordination Center	Subject Matter Expert
Jamie Parker	FS	Assistant National Incident Business Coordinator	Subject Matter Expert
Jason Swegle	DOI	OCIO Technical Architect	Subject Matter Expert
Jeff Lanham	FS	DCS Supervisor Hosting Services Support	Subject Matter Expert
Jessica Roosevelt	FS	Budget Officer (Represented Mark L)	Subject Matter Expert
Jill Kuenzi	FS	GIS Specialist	Subject Matter Expert
Jim Menakis	FS	National Fire Ecologist	Subject Matter Expert
Jodi Riegle	USGS	Computer Scientist / Cartographer	Subject Matter Expert
John Noneman	BLM	Senior Project Manager	Subject Matter Expert
Jon Norred	BLM	GIS Specialist	Subject Matter Expert
Josh Haney	NWCG	Training Specialist	Subject Matter Expert
Kara Stringer	FS	Deputy Center Manager, Great Basin Coordination Center	Subject Matter Expert
Karen Short	FS	Research Ecologist	Subject Matter Expert

Name	Agency	Job Title	Project Role
Keith Smith	NASF	Technology Program Manager	Subject Matter Expert
Kenneth Stacey	DOI	ISSO	Subject Matter Expert
Kevin Hoffman	FS	Technical SME	Subject Matter Expert
Kim Ernstrom	NPS	Fire Application Specialist	Subject Matter Expert
Kim McCutchan	FS	Chief NIICD (NIFC Radio Program)	Subject Matter Expert
Kim Van Hemelryck	DOI	Fuels and Landscapes Program Lead	Subject Matter Expert
Korby Johnson	DOI	Support to Infrastructure Team	Subject Matter Expert
Krista Gollnick-Waid	BLM	Fuels Management Specialist	Subject Matter Expert
Lani Williams	FS	Information Technology Specialist	Subject Matter Expert
Larry Van Bussum	NOAA	National Fire Weather Operations Coordinator	Subject Matter Expert
Laura Barrett	FS	Fire Management Specialist	Subject Matter Expert
Lauren Hickey	FS	Program Analyst (Fire Cache)	Subject Matter Expert
Lin Zang	DOI	Data Mangaement Specialist	Subject Matter Expert
Lori Glaeser	BLM	Instructional Systems Specialist	Subject Matter Expert
Lori Peltz-Lewis	FS	EDW Program Manager	Subject Matter Expert
Mark Fitch	NPS	Smoke Management Specialist	Subject Matter Expert
Marley Marshall	BLM	IQCS Business Steward	Subject Matter Expert
Michael Pena	BLM	IT Specialist (Security Manager)	Subject Matter Expert
Mike Cherry	FS	Emergency Management Specialist	Subject Matter Expert

Name	Agency	Job Title	Project Role
Mike Schievebein	DOI	Program Manager	Subject Matter Expert
Mike Van Hemelryck	NPS	Management Analyst	Subject Matter Expert
Mike Vigil	FS	Enterprise Architect	Subject Matter Expert
Mike Ward	FS	Prescribed Fire and Fuels Program Manager	Subject Matter Expert
Mitch Burgard	FS	Fire Applications Specialist	Subject Matter Expert
Morgan Pence	FS	Fire Applications Specialist	Subject Matter Expert
Mun-Wai Hon	Contractor	Technology Analyst	Subject Matter Expert
Nate Benson	NPS	Fire Science and Ecology Program Leader	Subject Matter Expert
Paul Schlobohm	NWCG	NWCG Branch Coordinator	Subject Matter Expert
Pete Lahm	FS	Air Resource Specialist	Subject Matter Expert
Ray Davis	FS	Old Forest & Northern Spotted Owl Monitoring Lead	Subject Matter Expert
Richard Del Hierro	FS	FAM IT Program Manager	Subject Matter Expert
Rick Gividen	DOI	Education Program Specialist	Subject Matter Expert
Robyn Heffernan	NOAA	Fire Weather Science and Dissemination Meteorologist	Subject Matter Expert
Sam Scranton	BIA	Forester	Subject Matter Expert
Scott Swendson	FS	Rocky Mountain Coordination Center Manager	Subject Matter Expert
Sean Peterson	FS	Intelligence Coordinator, NICC	Subject Matter Expert
Sean Triplett	FS	Team Leader	Subject Matter Expert
Skip Edel	NPS	Geospatial Fire Analyst	Subject Matter Expert

Name	Agency	Job Title	Project Role
Steve Larrabee	BIA	Fire Planner	Subject Matter Expert
Steve Manthei	WFIT	WFIT Program Manager	Subject Matter Expert
Steve Maurer	NASF	Intel Coordinator	Subject Matter Expert
Steve Smith	BIA	Fire Use Specialist	Subject Matter Expert
Steven Licari	DOI	Support to Infrastructure Team	Subject Matter Expert
Susan Goodman	DOI	Fire Management Analyst	Subject Matter Expert
Susan Shirts	FS	Incident Business Automation	Subject Matter Expert
Susie Stingley	FS	National Interagency Coordination Center Manager	Subject Matter Expert
Tami Parkinson	FS	Fire Application Specialist	Subject Matter Expert
Tate Fischer	FWS	National Fuels Mgmt Specialist	Subject Matter Expert
Ted Pierce	FS	Assistant Center Manager - NWCC	Subject Matter Expert
Tim Blake	NWCG	NWCG Branch Coordinator	Subject Matter Expert
Tim Wight	DOI	AD OS OCIO	Subject Matter Expert
Tod Dabolt	DOI	Geographic Information Officer/Chief Data Officer/ Director Information Management Technology Division	Subject Matter Expert
Victoria Smith-Campbell	BLM	Fire Management Specialist - GIS	Subject Matter Expert

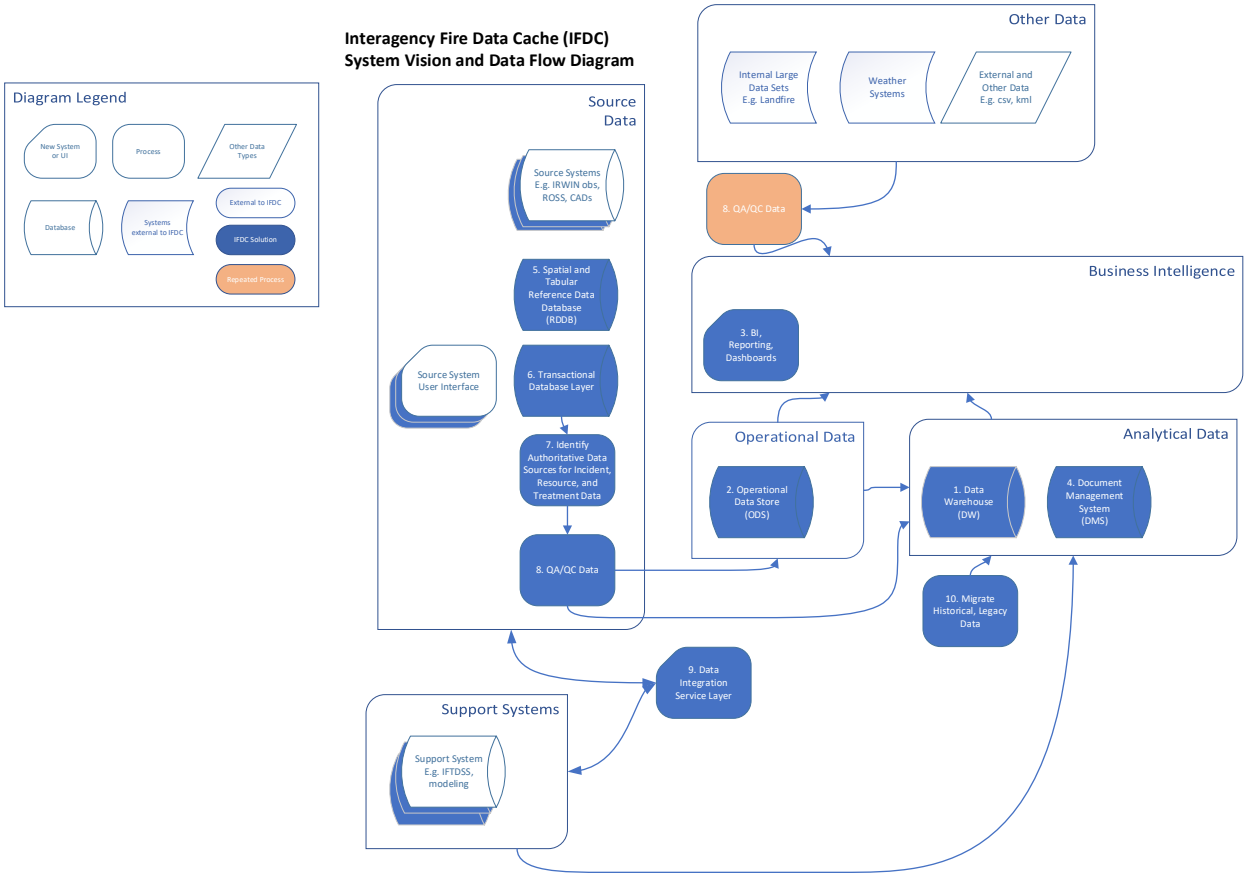
Table 1.5 - 1: IFDC Analysis Team

2.0 Data Cache Analysis

The following sections detail, where possible, the alternative options that we considered to address the business need. For ease of comparison, we chose to perform independent evaluations based on the goals for the cache. The analysis is organized as follows:

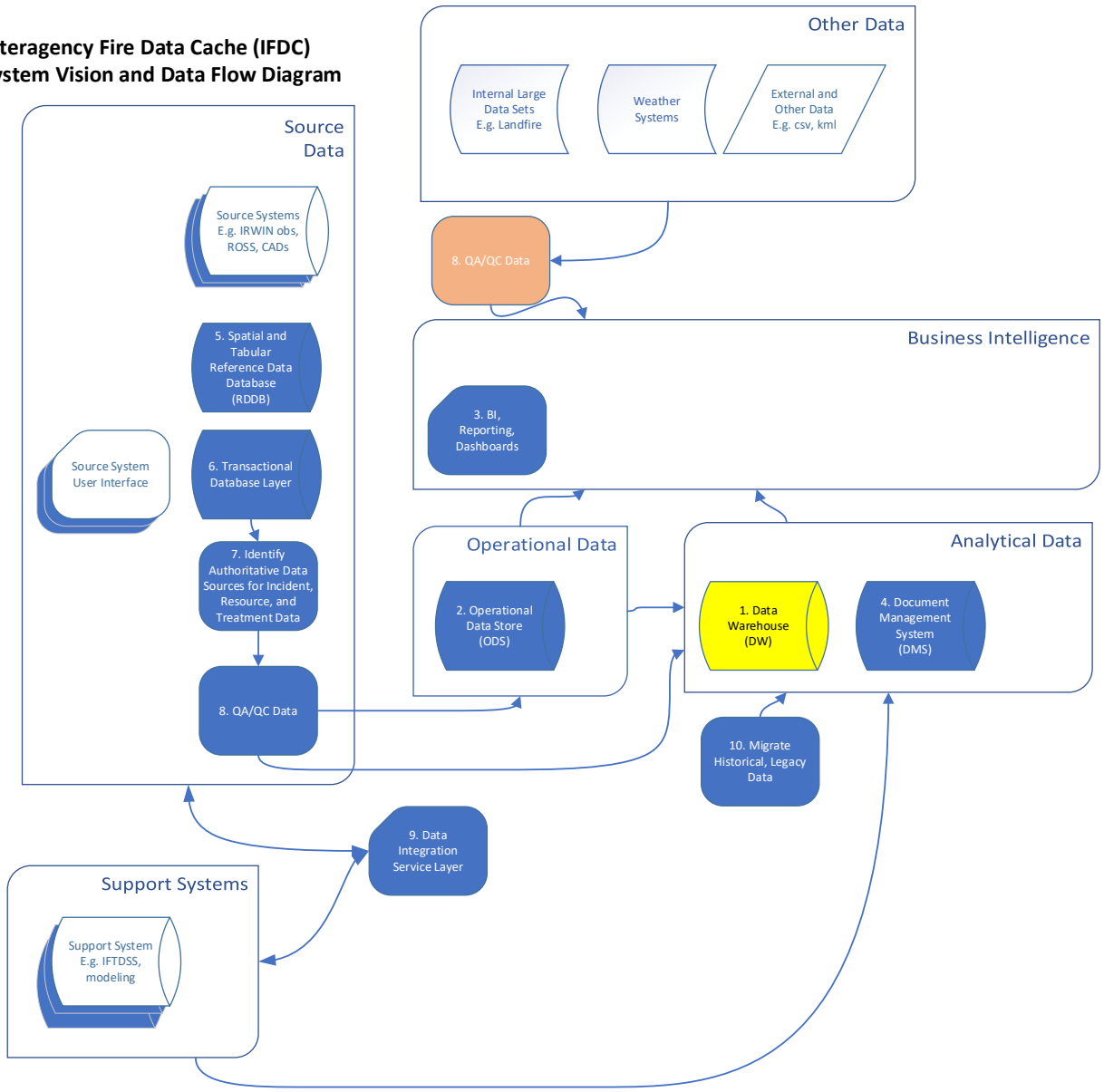
1. Display the Goal on the System Vision diagram
2. Define the business use for the Goal and the benefits
3. Where possible, identify potential technology options, as well as current systems in this role and where needs are met and lacking
4. Identify risks and dependencies for the Goal
5. Describe organizational impact for the Goal

Below is the complete diagram displaying all ten goals of the Data Cache (shown in blue), as well as the systems and processes that are not part of the Data Cache (shown in white), but are integral to the system vision as a whole.


Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram
Diagram Legend
New System or UI
Process
Other Data Types
Database
Systems external to IFDC
External to IFDC
IFDC Solution
Repeated Process
Source Data
Source Systems E.g. IRWIN obs, ROSS, CADs
5. Spatial and Tabular Reference Data Database (RDDB)
Source System User Interface
6. Transactional Database Layer
7. Identify Authoritative Data Sources for Incident, Resource, and Treatment Data
8. QA/QC Data
Other Data
Internal Large Data Sets E.g. Landfire
Weather Systems
External and Other Data E.g. csv, kml
8. QA/QC Data
Business Intelligence
3. BI, Reporting, Dashboards
Operational Data
2. Operational Data Store (ODS)
Analytical Data
1. Data Warehouse (DW)
4. Document Management System (DMS)
10. Migrate Historical, Legacy Data
9. Data Integration Service Layer
Support Systems
Support System E.g. IFTDSS, modeling

3.0 Goal 1: Data Warehouse (DW)

Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram



3.1 Business Use and Benefit

For the IFDC, the Data Warehouse (DW) will provide a "one stop shop":

- To support long term historical research and analysis of interagency data
- To integrate data from multiple sources into a single database structure and data model, enabling an enterprise-level view
- To support comparison analysis against real time data in the operational data store
- To reduce the data analysts' time to access information and reports

- To reduce the risk of data loss by providing a centralized repository that meets data management standards
- Can enable operational managers to evaluate information in order to foster more effective operational deployment of preparedness resources, for example providing comparative utilization statistics of firefighting resources such as engines, tenders, dozers and tractor plows.
- Can allow for comparative statistics for operational decision-makers of past utilization and performance for consideration for future refinement of planned resource type, location, and availability (e.g. run cards and step-up).
- Can provide statistics for senior leadership to communicate accurate, consistent intelligence with a high degree of timeliness and confidence
- To reduce/eliminate the risk of data loss
- Since all data in the DW will be available in the Business Intelligence tool, the DW will be used:
    - To create more consistent reporting
    - To allow important questions to be answered more quickly and with greater credibility than the community currently can
    - To create efficiency for business processes and decision making
    - To simplify and streamline reporting

The DW does not have inherent reporting functions; it simply enables confederated data stored in the DW to be visualized, reported on, and exported to other systems.

3.2 Possible Technology Options

Building or creating a DW is less about the technology used and more about design and implementation. Meaning, that a DW can be built on a variety of tools, such as Oracle, FME, IBM Information Server, and SQL, but the most difficult aspects of the DW are the design, extract/load/transform (ETL) process, and data management, as the DW should be based on the enterprise standards for data management for the organization.

The US Forest Service has built two Data Warehouses, one called FAMWEB and the other called Enterprise Data Warehouse (EDW). FAMWEB has fire-focused data, but lacks user community adoption and widespread use across the entire fire program and is built on an older technology. The EDW has had successful user adoption, but contains all enterprise data for the Forest Service, not just fire-related data. Both of these Data Warehouses could be considered as options for the IFDC DW. But the EDW in particular is a strong candidate for being leveraged further to meet IFDC DW needs.

3.2.1 Data Warehouse versus Data Lake

During our discussions, the concept of a Data Lake emerged as possible contender for this Data Cache goal, rather than a Data Warehouse. A Data Lake is similar to a Data Warehouse in that they both are data storage repositories. The primary difference is that data in a DW is transformed as it is loaded into the DW, where the Data Lake holds a vast amount of raw data in its native format, including structured, semi-structured, and unstructured data, then

the data structure and requirements are not defined until the data is needed. While storing data in a Data Lake can be less expensive than data in a Data Warehouse, this is not a crucial decision point for the Data Cache, as the amount of data in the DW will not be large enough to make the cost a significant factor. A Data Lake is optimized for a different purpose than the IFDC users need. Most users of the data in the DW are line of business users and not data scientists who need to reconfigure queries and data models on the fly. MBS believes that a Data Warehouse is more suitable for the IFDC than a Data Lake. However, if at some point in the future, the use case changes, the IFDC team could consider adopting the Data Lake model.

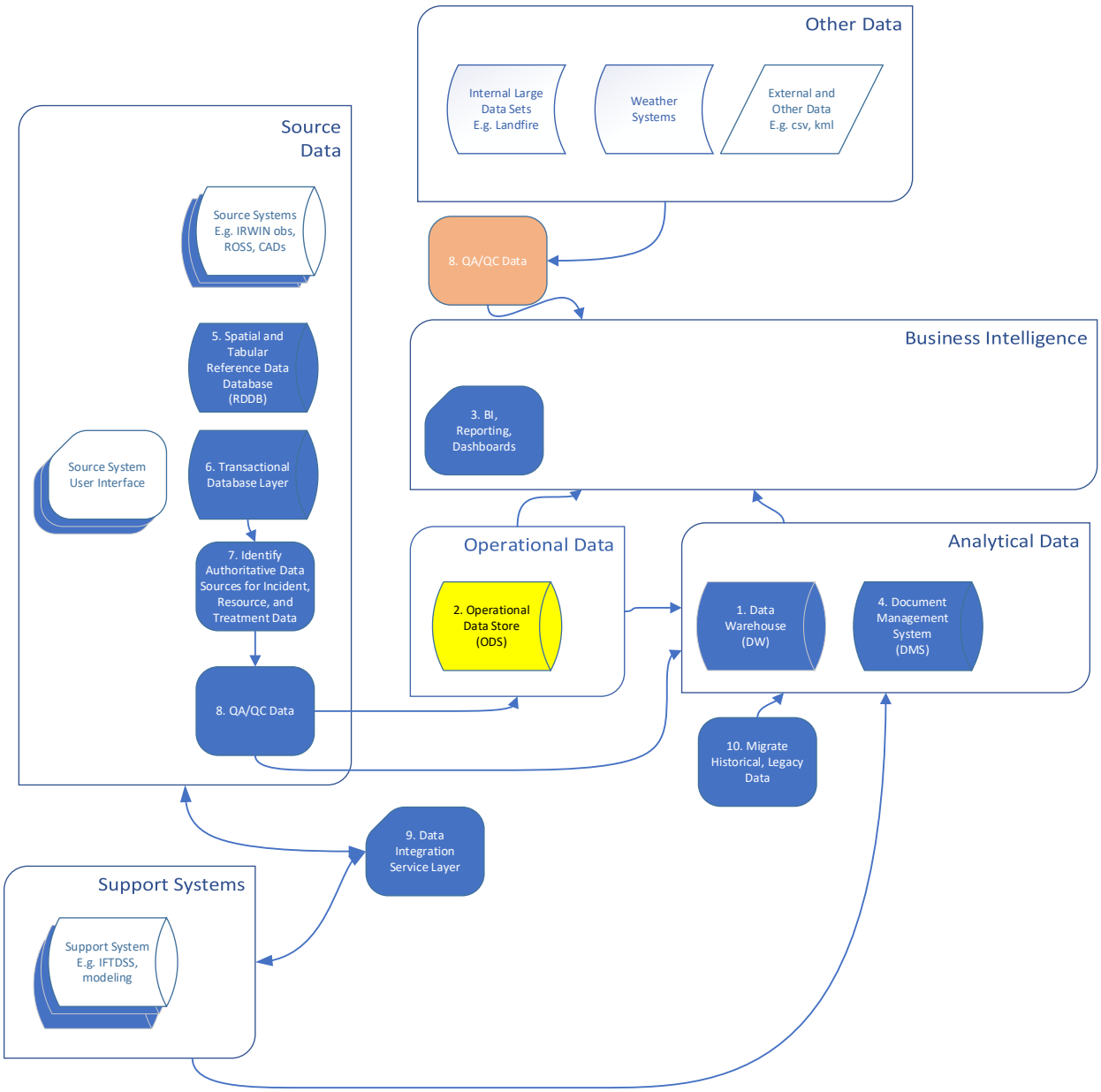
3.3 Risks and Dependencies

Many source systems that will provide data to the DW are created in older technology and aging architectures. Each system that is brought online with the DW will require careful consideration and planning. MBS had a senior solution architect evaluate the source systems that are candidates for inclusion in the DW as they exist today and made an initial determination that, due to the applications' technologies and platforms, as many as 70% of the systems will require a developer to code custom data extracts in order to access and process the data, adding time, complexity, and cost to the DW project.

3.4 Organizational Impact Considerations

Organizational Impact	
Tools	• Regardless of the tools and solution chosen to implement the Data Warehouse, existing systems and archiving tools will be changed
People	• Training employees on the new tools and their use in support of other organizational tools will be required
Process	• Multiple processes will be built to ETL the data from the many source systems in order to populate the Data Warehouse • Users should be made aware of the cadence of data being loaded into the DW, so that they know the currency of the data that is available
Budget	• Depending on solution chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for

4.0 Goal 2: Operational Data Store (ODS)



4.1 Business Use and Benefit

For the IFDC, the ODS:

- Will have the best and most recent, near real time data from source systems, such as CAD systems, IRWIN, weather systems and ROSS/IROC, 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

- Can enable deliverables for safer operational support of field level firefighters, for example, improved situational awareness by ensuring centrally created, consistent, accurate fire danger pocket cards are easily available
- Can provide information for operational decision-makers to make the best possible strategic allocation of finite resources during times of competition
- Can allow for accurate, consistent, real-time information for dispatch centers to ensure common situational awareness of and for firefighters regardless of incident jurisdiction or resources responding during initial attack operations
- Will eventually load most data into the DW, and all ODS data will be available in the BI Tool
- Will be a one-way data feed, meaning that once data is loaded into the ODS, if there are changes to the data, those changes will be made in the source systems and then loaded again into the ODS upon the next data refresh.

4.2 Possible Technology Options

Building or creating an ODS is less about the technology used and more about the design and implementation. Meaning, that an ODS can be built on a variety of tools, such as Oracle and SQL, but the most difficult aspects of the ODS are the design, ETL management, and data management, as the ODS should be based on the enterprise standards for data management for the organization.

4.2.1 The Roles of Enterprise Geospatial Portal (EGP)

Today the wildland fire community has access to the EGP for reporting and analysis work. The EGP is the default interagency authoritative data source of standardized geospatial information for the full range of wildfire activities ranging from response to planning.

The EGP leverages a central source of spatial data for mapping, decision support, business intelligence, and situational awareness through multiple tools to view and analyze wildland fire data. Users utilize the EGP for web-mapping and enterprise database components with Google and Esri technologies. The EGP is currently integrated with IRWIN (Integrated Reporting of Wildland Fire Information). In addition, EGP uses data and data services from ICS209, WFDSS (Wildland Fire Decision Support System), WIMS (Weather Information Management System), Predictive Services, Homeland Security, FAA (Federal Aviation Administration), Geomac, among others. And more data sources are being added as they become available. Users can edit some data in EGP. EGP stores some data and performs some ETL on the data.

The EGP has many characteristics of an ODS, but it also has some aspects of a Business Intelligence Tool, is an Interagency Authoritative Data Source for some data, such as incident perimeters, and provides some Data Warehouse functionality. There is definite potential for the EGP to act as the ODS for the IFDC, and MBS recommends evaluating this further.

4.3 Risks and Dependencies

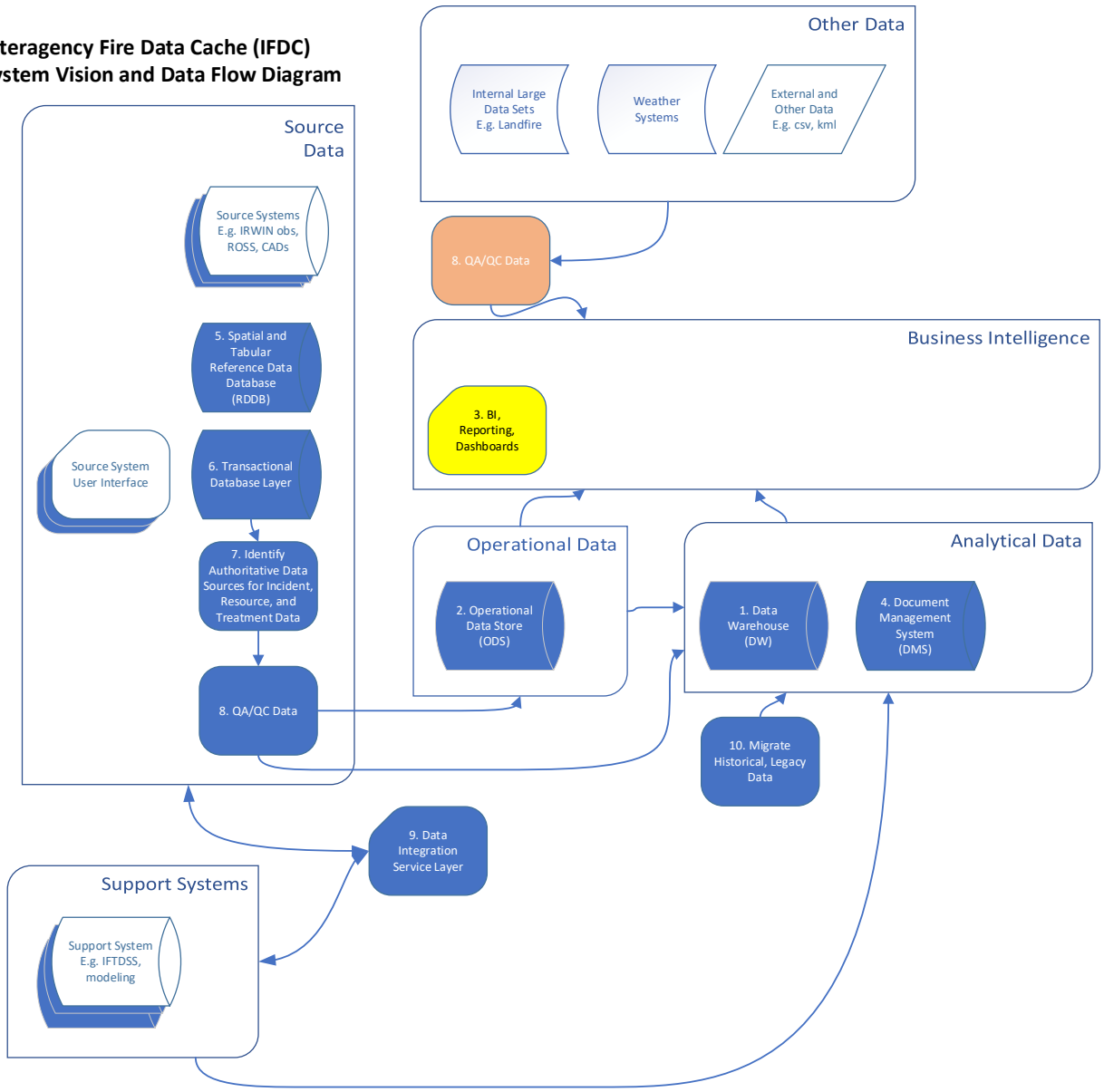
Many source systems that will provide data to the ODS are created in older technology and aging architectures. Each system that is brought online with the ODS will require careful consideration and planning. MBS had a senior solution architect evaluate the source systems that are candidates for inclusion in the ODS as they exist today and made an initial determination that, due to the applications' technologies, as many as 70% of the systems will require a developer to code custom data extracts in order to access and process the data, adding complexity, time, and cost to the project.

4.4 Organizational Impact Considerations

Organizational Impact	
Tools	• Regardless of the tools and solution chosen to implement the Operational Data Store, existing systems and archiving tools will be changed
People	• Training employees on the new tools and their use in support of other organizational tools will be required
Process	• Multiple processes will be built to extract and transform the data from the many source systems in order to populate the ODS • Users should be made aware of the cadence of data being loaded into the ODS, so that they know the currency of the data that is available
Budget	• Depending on solution chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for

5.0 Goal 3: Business Intelligence Layer (BI)

Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram



5.1 Business Use and Benefit

- 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 that is refreshed frequently. Where resource and equipment users may use a BI tool to compare resource allocations across multiple fires.
- BI tools can access data from the DW, Document Management System (DMS), ODS, and other data sources

- Will access spatial and tabular data
- The users of the system will be line of business users, and not IT users.

While the BI tool can produce reports on data in the DW, ODS, DMS, and other data sources, MBS cautions against using the BI tool to replace source system static reports, meaning reports that use data from a single source system. One potential issue is that data from the source system is only populated into the DW and ODS on the cadence established for the source system, so reports created in the BI tool may not be the most current data. The notable exception to this is INFORM, where that system is being developed on IRWIN and has no inherent reporting functionality.

5.2 Possible Technology Options

As part of the analysis, MBS had three different vendors, IBM Cognos, Tableau, and ESRI Insights, provide demonstrations of their systems to IFDC stakeholders. These business intelligence/data visualization tools are a good representation of what is available in the market today. There are, however, other tools that the IFDC team may want to consider, such as Qlik, Microsoft PowerBI, OBIEE, and Informatica.

System	Advantages	Disadvantages
IBM Cognos	• Has a presence in the interagency community today, as USFS FAMWEB and EDW are built on Cognos tools • Is well-established from a reputable vendor • Has supporting systems available to help with ETL and data deconfliction, validation, and QA.	• Formerly ranked high on both Gartner's Magic Quadrant, as well as Forester's Wave reports as a leader in business intelligence tools, but has not been in the leader category in a couple of years • Not known how internal competition with Watson Analytics will confuse the product offering or product investment • Higher learning curve • Dashboards are not as simple to build • Generally requires IT involvement for report building and dashboard creation • Uknown AGOL integration
Tableau	• Has a presence in the interagency community today, as some users are utilizing Tableau today to build reports and dashboards • Low learning curve • Better dashboard capabilities • Generally empowers line of business users to create reports and dashboards • Works with AGOL for mapping	• Has no supporting tools to help with ETL and data deconfliction, validation, and QA, meaning that the data used in Tableau reports is only as good as the data it accesses to build the reports (can use best of breed tools for these functions, however)
ESRI Insights	• Line of business users have familiarity with ESRI tools	• Relatively new to the marketplace • Works only with AGOL-enabled data

Table 5.2-1: Business Intelligence and Reporting Tools

5.3 Risks and Dependencies

The data used in the BI tool is only as good as the data that is received. The BI tool does not validate or ensure the data is reliable.
There is a decision point regarding the potential to standardize to one BI tool or allow many BI tools to point to data in the DW, ODS and DMS. Some considerations for the former are the ability to standardize training, share reports and dashboards across organizations, and the opportunity for knowledge transfer among the user community. If the decision is to use many BI tools, then users can choose their preferred tool, but the reporting efforts will be more siloed.

5.4 Organizational Impact Considerations

Organizational Impact	
Tools	• New methods for creating reports and conducting research and analysis will be available to users • Usability for line of business users should be a primary concern • IT involvement to create reports and dashboards should not be necessary
People	• If a single solution is chosen, training can be more centralized • If users are given a choice of BI tools, training will be decentralized and more difficult to manage
Process	• With training, users can create dashboards, reports and analyses autonomously
Budget	• Depending on solution or solutions chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for • Training costs should be accounted for

5.5 Related Topic

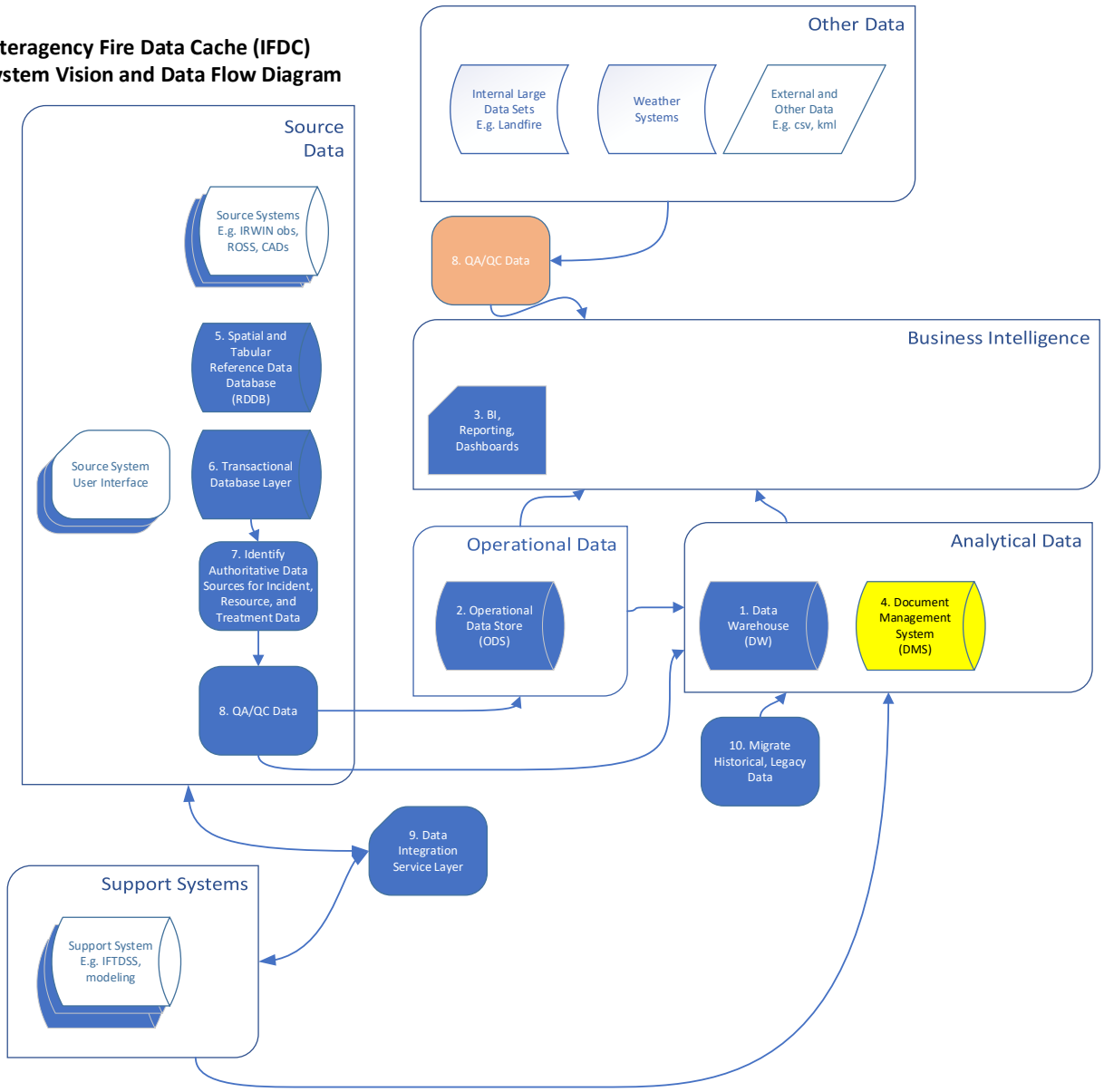
5.5.1 IBM Information Server

During the IBM Cognos demo, the vendor discussed a set of tools called Information Server that works at the front end of Cognos as part of the ETL process. Essentially, Information Server works to ensure that the data used in reporting and analysis can "be trusted but verified." After the Cognos demo, MBS requested a more in depth conversation on Information Server with the vendor. We had an in-person meeting with representatives from IBM who showed us the tools. Information Server is a family of products designed to enable the business to understand, cleanse, monitor, transform, and deliver data, as well as to collaborate to bridge the gap between business and IT.

MBS had hoped that Information Server could help the IFDC team with some of the data governance, QA/QC, and metadata management issues they face today. However, the demo we saw showed a fair amount of complexity to the product that would mean the IFDC team would need expert resources or professional services to set up the product and provide ongoing maintenance and support. Also, while the product is going through a modernization phase, it seemed to lack a contemporary user interface and user experience that we expected to see.

6.0 Goal 4: Document Management System (DMS)

Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram



6.1 Business Use and Benefit

- 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
- 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

- Possible uses for the DMS are storing IAPs, forecasts, policies, burn plans, fire management plans, user guides, training aids, job aids, etc.
- Metadata and key word tags attached to documents enables better search and discovery
- Stored documents can be accessed through BI tool
- Will ensure compliance to policies and mandates and can accommodate agencies' differing records management policies
- Could centralize and eliminate other DMS systems, saving time and money, including the FTP sites
- Storing databases from other systems could be valuable to the community and prevent data loss (eIsuite and Fire Family Plus)
- If workflow and digital signature are present in tool, there is potential to eliminate some systems or consolidate systems, such as 209
- There is potential to use a DMS to store burn plans, fire danger operating plans and related components for systems like IFTDSS

Document indexing: Document indexing is a technique that makes search and retrieval of documents and artifacts seamless. When discussing indexing, we often use the term metadata. This is essentially data that describes data, such as an abstract, key words, and summaries. Metadata is typically used to supplement and enhance the original data.

6.2 Possible Technology Options

6.2.1 IRMA's Data Store

Integrated Resource Management Applications (IRMA) is an NPS project that serves many functions. Within IRMA is a Data Store that houses a variety of documents, datasets and associated metadata. It has user administration capabilities, needed security, and wildland fire NPS related documents are stored there today. It is in active development and the administrators are open to making changes to address IFDC needs. Due to the nature of this project component, the fact that NPS is already investing in IRMA and that it is a functioning document repository today, MBS thinks that regardless of which direction IFDC takes for the other components of this project, they should explore the option of adopting IRMA's Data Store as the interagency system of record for fire document storage. IRMA Data Store meets the business requirements, has the needed security and user administration, stores needed metadata, has search and retrieval functionality and all NPS users currently can access this system. There are a couple of options for using IRMA's data store as the Document Management System for the Data Cache. One is to have NPS users grant access to external agencies, and the other is to create an IFDC instance of the IRMA data store for exclusive use.

A system called Pinyon, a US Forest Service system was discussed as a potential system to leverage for the IFDC DMS. The team should evaluate the possibility of leveraging this system.

Other options include Sharepoint and FireNet. FireNet enables interagency access for all NWCG partners and those supporting wildland fire management to a centralized and secure network of resources including documents, customized portal sites, and more.

6.3 Risks and Dependencies

There are currently several systems in use today that serve as document management systems, such as Sharepoint systems, Google solutions. There is a business benefit to consolidating the disparate DMS systems in place today and creating a centralized and standardized location for this information, whether that is SharePoint, IRMA, or something else, for reduced O&M and interface opportunities for interagency use.

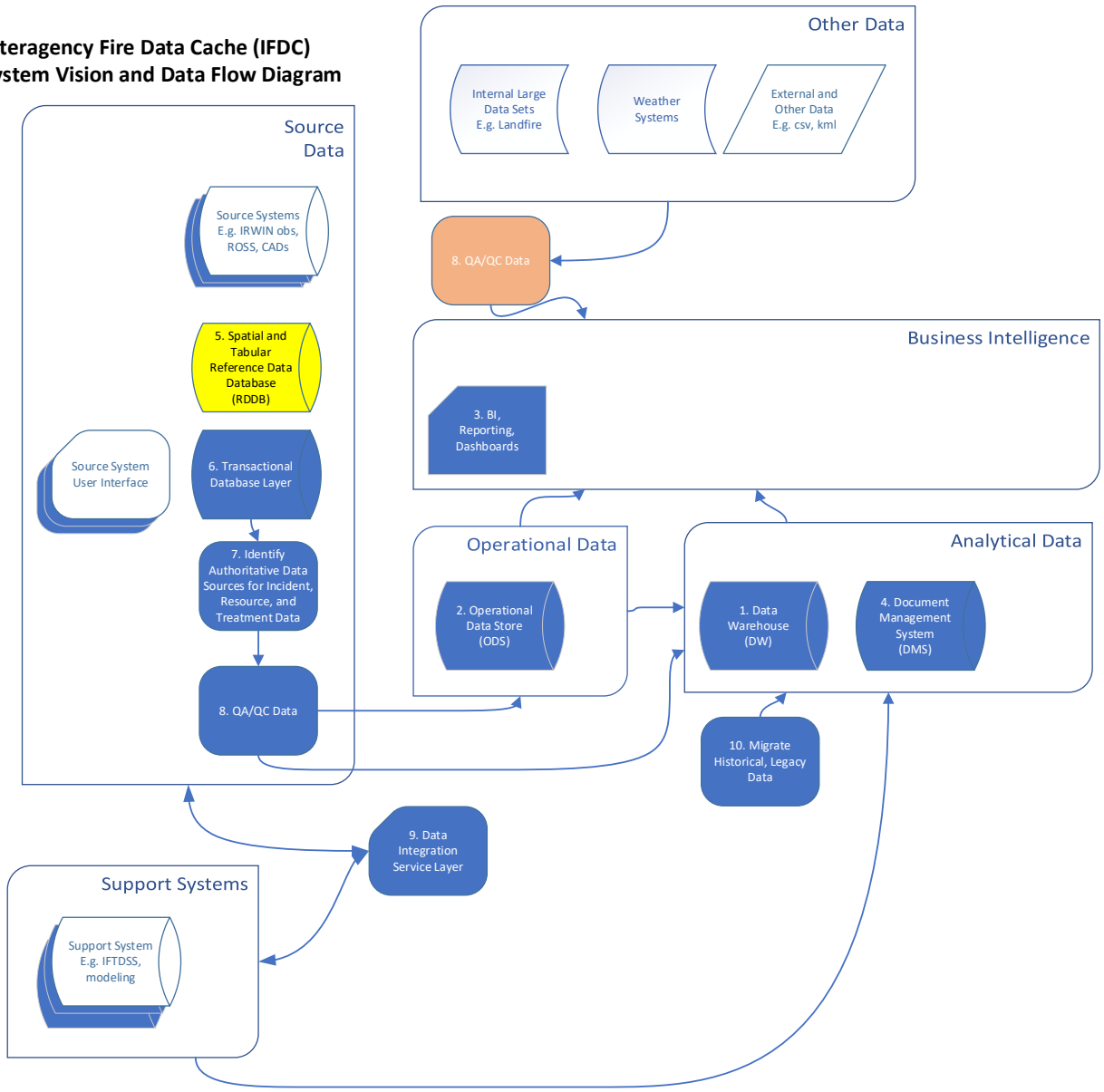
The IFDC team should consider the governance and business rules for storing documents, including security, authentication, approvals, record management guidance, etc.

6.4 Organizational Impact Considerations

Organizational Impact	
Tools	• Regardless of the solution chosen to implement the DMS, existing systems and archiving tools will be changed
People	• Training employees on the new tools and their use in support of other organizational tools will be required
Process	• Process improvements mean users can store all allowable artifacts using repeatable processes, which should shorten the learning curve for users
Budget	• Depending on solution chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for

7.0 Goal 5: Spatial and Tabular Reference Data Database (RDDB)

Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram



7.1 Business Use and Benefit

- For the IFDC, all business subject areas and applications may use the data in the RDDB for operations, reporting and analysis
- Users with the correct access and permissions will be able to create, edit, update, and delete (CRUD) the data in the RDDB
- The RDDB will be the interagency authoritative data source for data considered reference data, both spatial and tabular

- The RDDB can pre-populate data in fire reports and other reporting with authoritative reference data, creating consistency in reporting, decision-making and planning
- Time and cost savings for future development if application developers can access this data rather than duplicate data creation and maintenance across systems
- The RDDB can address some data quality assurance concerns, as the RDDB will be considered the Interagency Authoritative Data Source for the data it contains

7.2 Possible Technology Options

There are no currently existing systems that can support this function. Therefore, the RDDB would be a new project for the IFDC team.
However, Geoplatform, a DOI CIO initiative, has expressed interest in taking on responsibility for this project.

7.3 Risks and Dependencies

The risks to becoming the owner of reference data for all users are many. For one, this is a paradigm shift for the participating organizations who are accustomed to managing their own data and would require large-scale adoption in order to be successful. Another risk is the IFDC's responsibility for being stewards of reference data. Meaning that the IFDC would be responsible for storage, archival, maintenance, disaster recovery, and infrastructure of the database. Another risk is the timeline to implement.

This data stored in the RDDB should be fairly stable data, but the refresh timelines for data will vary depending on the type of data.

Providing a way to track changes to data will be critical for auditability and data integrity.

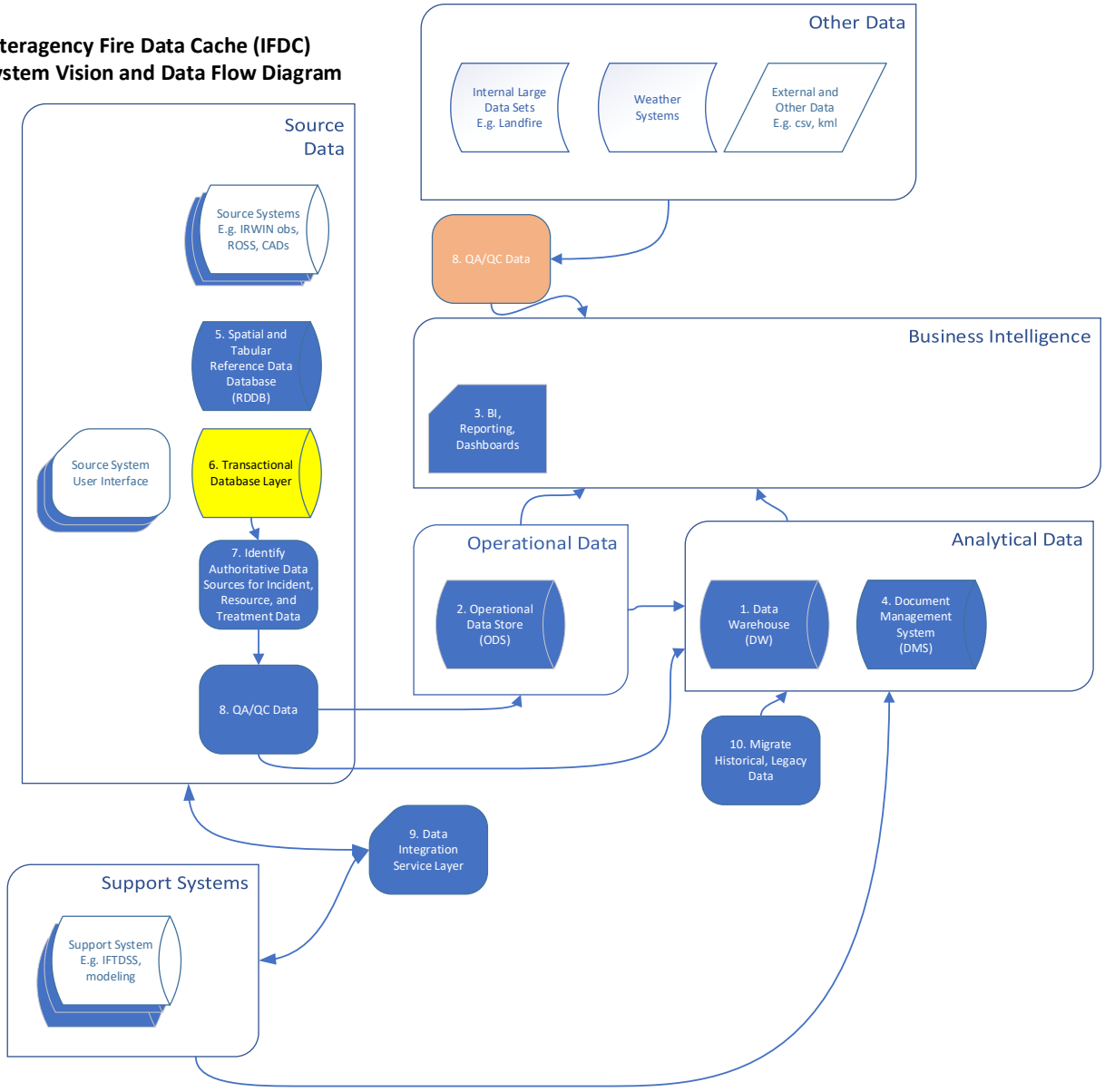
The Data Dictionary project impacts the RDDB in that the RDDB relies on information that will be uncovered during the Data Dictionary work. If the Data Dictionary project is not timely, the RDDB project will be at risk for delays.

7.4 Organizational Impact Considerations

Organizational Impact	
Tools	• Regardless of the solution chosen to implement the RDDB, existing systems will be changed
People	• Training employees on the new tools and their use in support of other organizational tools will be required • Users would need to be made aware of the RDDB to benefit from it so a robust communication plan would need to be developed and implemented
Process	• The implementation of the RDDB solution means that data considered to be reference data would be stored and managed in a centralized location. This is a departure from the processes used today and should give all users better and more reliable data
Budget	• Depending on solution chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for

8.0 Goal 6: Transactional Database Layer

Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram



8.1 Business Use and Benefit

- 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 IFDC team would have purview over ensuring that new systems are developed in an efficient way.

8.2 Risks and Dependencies

MBS would recommend having a common database layer that applications can adopt for use rather than having a single database shared by multiple applications. Sharing one database amongst multiple applications has some serious disadvantages: The more applications that use the same database, the more likely it is that you hit performance bottlenecks and that you can't easily scale the load as desired. SQL Databases don't really scale. Maintenance and development costs can increase. Development is harder if an application needs to use database structures which aren't suited for the task at hand but have to be used as they are already present. It's also likely that adjustments of one application will have side effects on other applications. Administration becomes harder. Questions such as: Which object belongs to which application?; Where do I have to look for my data?; Which user is allowed to interact with which objects?; What can I grant whom? Become difficult to answer.

Upgrading is more cumbersome. You'll need a version that is the lowest common denominator for all applications using it. That means that certain applications won't be able to use powerful features. You'll have to stick with older versions. It also increases development costs a bit. It can cause concurrency issues. What if one application modifies data that is outdated or should've been altered by another application first? What about different applications working on the same tables concurrently?

The interagency wildland fire data processing occurs across a large number of applications – on the order of 50 to 75 main applications and another 50 lower tier applications that are not directly supported by interagency wildland fire IT. These applications reside in a variety of technologies, servers and systems, making the data less accessible than if it is stored in one place. Some of the technologies are older and do not provide “open” or easy communication to other systems.

Interagency wildland fire IT has also adopted a philosophy of “Data Stewardship” rather than “Data Ownership”, which means that at least for some data, it is not clear what the best and final value is until this data is placed together in a common location and evaluated. As an example, IRWIN today acts as the common location for evaluation of data for incident data.

The combination of these two factors (number of applications and data stewardship/ownership) has led to the need for a place to store wide-ranging fire data, beyond just incidents, to ease the burden of applications trying to obtain authoritative data and to ease the burden of reporting solutions accessing data. This concept is the primary catalyst for the development of an idea called the “Data Cache”. The goal of the Data Cache would be to build a logic layer to resolve the issues of the “Eventually Consistent Database” and bring consistent and final values to a centralized database that can be used as a Source of Truth, or an Interagency Authoritative Data Source, for feeding downstream applications and reporting solutions. It is a layer between passive data collection systems and systems that must have consistent values.

The Data Cache becomes a compromise position between “Data Stewardship” and “Data Ownership” – it does not resolve accurate data values at the time of data entry and when

transactional (OLTP) data processing occurs, yet it established authoritative data values for downstream system. In this definition / context, the Data Cache –Repository is similar to a "Data Warehouse" (OLAP Database), as it collects and resolves data discrepancies from multiple sources; however, it would feed downstream source systems for further data processing making it more of a data broker, or even an OLTP-like / transactional database, and not a data sink.

Furthermore, the Data Cache has also been considered for the role of acting as the overall main database for applications. For example, if a new application is built, the application developers would not have to build a database for their application; a centralized application database would take on this role. This is a different role than what is described above. If the centralized database is acting in a pure Data Ownership role for its supported application, the centralized database is an OLTP, Source System database, or a "Transactional Database Layer". In this role, the database is directly serving the needs of a single application, not resolving data differences among multiple data sources. A term for this is a "Consolidated Data Store". There are several articles that discuss the advantages and disadvantages of a shared centralized database for OLTP applications. They are below:

https://msdn.microsoft.com/en-us/library/ff647273.aspx
https://softwareengineering.stackexchange.com/questions/105786/should-i-use-one-database-per-application-or-share-a-single-database-amongst-mul
https://stackoverflow.com/questions/3479297/multiple-application-using-one-database
https://worldclasstech.wordpress.com/2009/02/09/a-single-database-or-multiple-databases-for-a-global-company/

Similarly, a centralized database can act in the role of providing the "Data Stewardship" or passive collection of data as it arrives into an application. This is also different than the role of resolving conflicting data from multiple sources; in this role the centralized database is acting as the data collector, not the data adjudicator and not the data owner.

An interagency wildland fire data need, compounding the desire for a centralized database in one or all of these roles, is the difficulty in extracting data from systems that are not "Open". The centralized database bears the burden of communicating with the systems it needs to collect data, so that downstream applications and reporting tools have easy technological access to data.

The Data Cache has been tagged for all of these roles in different contexts depending on individual application and reporting needs. These roles and database architectures need to be separated, at least conceptually, to set clear requirements and goals for funded project initiatives and perhaps further by data area (e.g. incidents, resources, aviation, etc.).

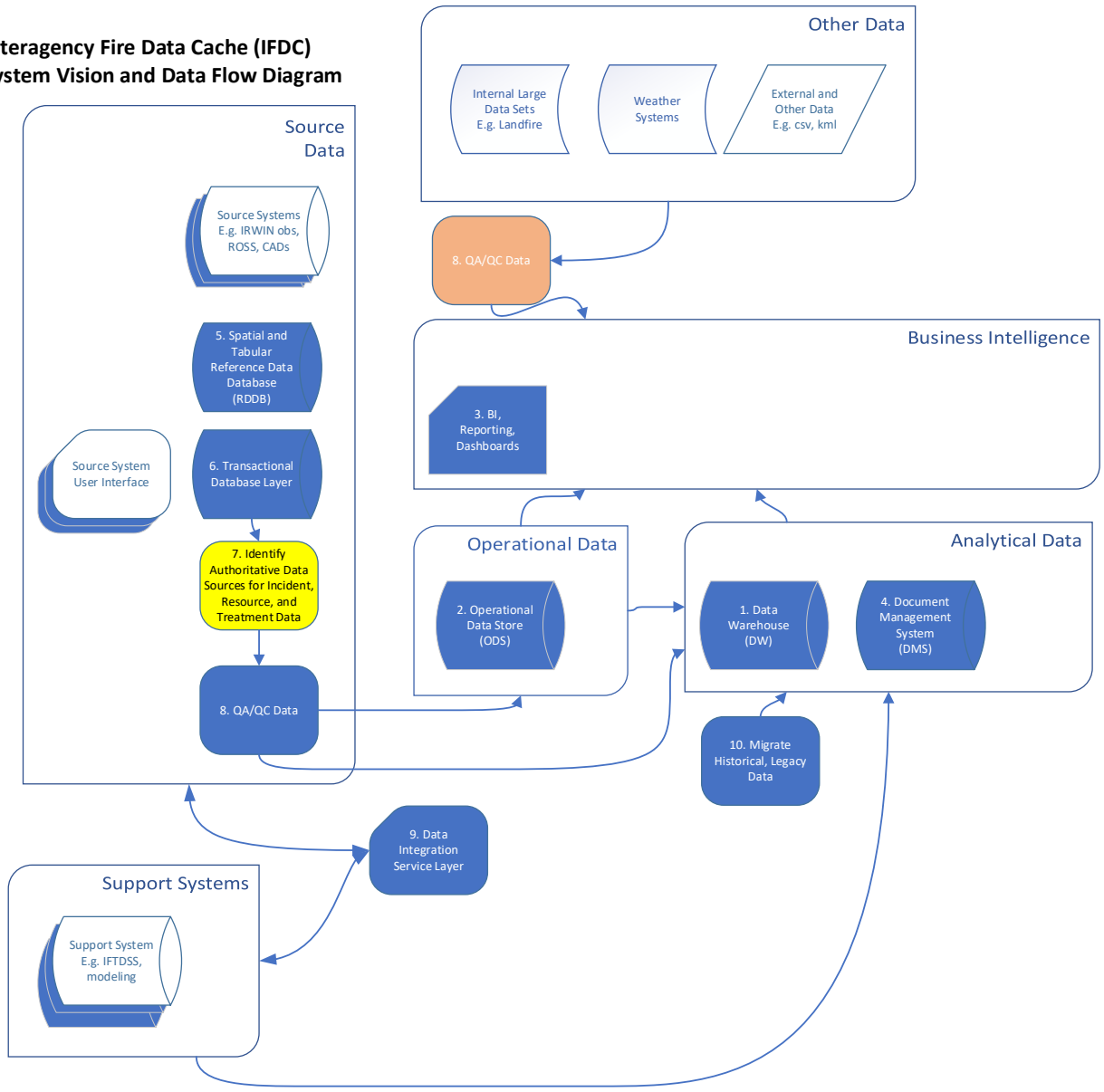
8.3 Organizational Impact Considerations

Organizational Impact	
Tools	• Implementation of the Transactional Database Layer would mean that as new systems are brought online, those systems could utilize the database and infrastructure provided

Organizational Impact	
People	• Development teams and database administrators would need to coordinate development of new systems with the IFDC team
Process	• The process for developing new applications within the enterprise would be changed; all newly developed systems will adhere and adopt the agreed-upon development tools
Budget	• Providing this service to users could be a significant cost to the IFDC team, as the IFDC team would be responsible for maintenance, support, and infrastructure costs

9.0 Goal 7: Identify Authoritative Data Sources for Incident, Resource, and Treatment Data

Interagency Fire Data Cache (IFDC) System Vision and Data Flow Diagram



9.1 Business Use

- IFDC users will have clarity about interagency authoritative data sources for Incidents, Resources, Treatments, etc. and will allow users to clearly identify which systems "own" given data (Interagency System of Record (ISOR)), as well as knowing where to correct data
- These are the locations data would be "corrected" or defined as THE data

- Users of the IFDC can be assured that the Data Cache has the most reliable data. Example of the use of this ar:, What is THE source for the initial WF location (point)? What is THE source for the "daily" (most current) perimeter? What is THE source for the final fire perimeter? What is THE source for the RAWS station data?
- This could possibly eliminate/streamline/reduce the footprint of systems (FireCode, SIT, 209, WFMI-W/WIMS, NFPORS/FACTS, Firestat/WFMI-FR, IROC/eISuite/ICBS, modeling tools)
- Downstream reporting gets much easier, more accurate and valuable

NOTE: Recently, the NWCG team has defined two terms that are helpful for this goal. As specified in Appendix A of the NWCG Data Strategy document (https://www.nwcg.gov/sites/default/files/publications/pms940.pdf :

Interagency System of Record (ISOR): Agencies and bureaus may have their own SOR for their data and an ISORs is identified by an interagency business area as the official application source of interagency data. 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. An ISOR can be an external source of data used by wildland fire.

Interagency Authoritative Data Source (IADS): 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. There may be more than one IADS and it can change depending on business process complexity and incident life cycle. An IADS may be a compilation or subset of data from other authoritative sources. The Data Lifecycle Management (DLM) process ensures sources, limitations, currency, and attributes for the IADS are documented.

9.2 Possible Technology Options

The solution for this goal is less technical and more of a business process, data governance, and data management solution. Meaning, there is not a single technical solution that can solve this goal, but once the business rules are in place, a technical solution can help administer and manage the process, much like IRWIN does today for incident data.

9.3 Risks and Dependencies

As MBS gathered Data Cache requirements from interagency wildland fire stakeholders, we discovered several gaps related to incident data:

- There is no clear way to count the number of incidents in a given year
- There is no clear way to count the number of acres involved in fire incidents in a given year
- There are multiple data sources for fire perimeter data and it is unclear which one to use for reporting
- Smaller incidents are reported in quantity and in a method different from large fire incidents

- There is no clear way to combine duplicate incidents (Current release of IRWIN starts to solve this)
- There is no clear way to identify merged fires that might have begun as separate incidents and should now be combined
- There is an application that has a single function and that is to generate an accounting code for a fire incident

All of this feedback points to the need for an Incident Interagency System of Record. The following systems act in this role today: IRWIN (incident data brokering and storage of some unique fields), EGP (geospatial data entry, such as incident perimeters), SIT (small fire quantity reporting), FIRECODE (accounting code generation), INFORM (incident data entry and incident after action reporting), Geomac (perimeter reporting and editing) and the dispatch systems, such as WILDCAD. It is unclear which of these systems should be accessed to pull an incident count and an acre count. It is unclear which system is the Interagency System of Record for fire perimeters. It is unclear how to combine the data for small fires in SIT with the fire data in IRWIN. CADs and 209 have been identified as the Interagency Autoritative Data Source for creating and updating Incident Complex data but FIRECODE, ROSS and other applications have not fully adopted the process and continue to allow users to create Incident Complexes that are not shared It is likely that the current scenario evolved due to concerns about interagency wildland fire "owning" incident data, because jurisdictional agencies responsible for managing an incident "own" the data and the interagency needs have not been fully considered.

If there existed one "Interagency System of Record" or Data Owner / OLTP System / Source System / Transactional System for receiving, tracking, performing QA / QC, storing and exporting incident data, these issues would have simplified solutions and bring clarity to data processing rules through the wildland fire application suite. The Interagency System of Record should include functions to combine or eliminate duplicate fire records, eliminate invalid fire records, choose an accounting code for funding, and store geospatial data such as initial, daily and final perimeters for the fire. (During meetings with stakeholders, the concept of this system was coined the Back Office Incident System – Electronic, or "BOISE"). Interagency wildland fire can choose an existing system and establish functional requirements for it to fill this role, or it can create a new system to fill this data ownership gap. Data ownership can be spread across the multiple systems as long as the data ownership, QA / QC ownership and function ownership is clear among the systems.

Incident data is particularly important as it serves as the basis for all other activities in the data processing systems for wildland fires. Without incidents, there would be no need for personnel, aviation, equipment, day-by-day operations or even treatments (to prevent / mitigate incidents). Incidents are at the center of the data universe for wildland fire. They are the equivalent of orders at a factory, trouble tickets / service requests for a telephone company and IT Service Management Tickets for an IT function. Clear data ownership of incident data will provide a solid foundation for the data requirements for surrounding systems.

The source (ISOR/data owner) for an incident should define which is the correct incident record (resolve duplicates). This is more of a business process than a technology solution although technology can help identify potential duplicates and document the decision for the record.

As noted above, IRWIN is performing this function partially today for incident data because it exposes potential duplicate records. To fully fulfil this function, IRWIN would have to be designated as the Interagency Authoritative Data Source for incident data and it's project scope re-defined to support the designation

Implementing an Interagency Authoritative Data Source data governance model may require some business process re-engineering andmodifications to existing systems.Current systems will have to maintain capability until the "replacement" is up.

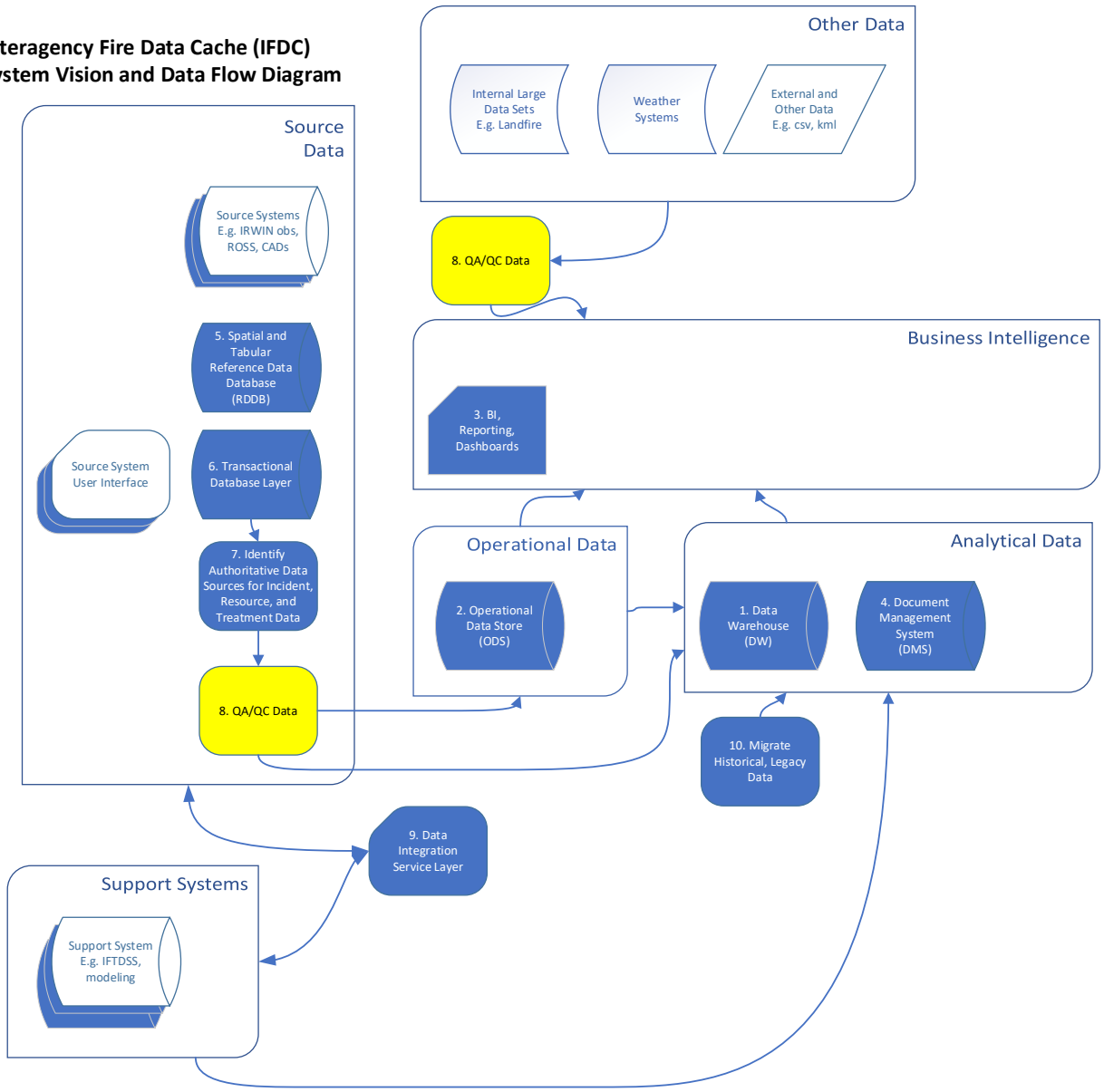
May require a cost increase before the team sees a cost decrease.
Some systems will have change from a passive system to a more active system (IRWIN, 209, etc)

9.4 Organizational Impact Considerations

Organizational Impact	
Tools	• Implementation of the IADS may or may not require new tools, depending on the direction taken. • Modifications to existing applications could occur over time based on their lifecycle stage. • Full value and benefit of the IADS approach will not be realized until the majority of applications transition
People	• Users can rely on the credibility and reliability of the data they use
Process	• Adding an IADS process for all types of fire-related data is a tremendous undertaking and many processes will need to change and be added to accomplish the project • Most processes will be simplified for users (like IRWIN has simplified dispatch workflows.)
Budget	• Depending on solution chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for

10.0 Goal 8: Quality Assurance/Quality Control Data (QA/QC)

Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram



10.1 Business Use and Benefit

- For the IFDC, all business subject areas need to be able to trust that the data in the ODS, DMS, DW, and other data used in the BI tool is reliable, valid, and accurate
- It is best if upstream systems do their own QA/QC, and downstream systems can have to trust the upstream systems' processes. However, some QA/QC can only be done once the data is aggregated. Therefore, QA/QC can happen at more than one location in the data flow.
- Will be for both spatial and tabular data

- Will allow for better data for decision making and reporting
- The DW, ODS & DMS will have increased credibility and transparency
- Is a more proactive than reactive management of data, increasing the predictability of workload management and reduces risk of bad data at critical junctures
- Can help identify gaps in data or missing data, i.e. completeness of data
- The RDDB, the Data Integration Service Layer, and the ADS goals can assume some portion of QA/QC data

10.2 Risks and Dependencies

Regarding the importance of a System of Record / Data Owner and Data QA / QC application functions:
This concept is important for identifying which applications (systems) own which data elements, bringing clarity to interface requirements between systems, data reporting, and system functional requirements. "System of Record" has a strong meaning in Federal Government IT, because when a system is identified as an SOR, it gains additional process, functional, maintenance and documentation requirements. There are synonyms for a System of Record that do not carry as strong a meaning, such as data owner, source system, transactional system or an On-Line Transaction Processing (OLTP) system.

There are many examples of when data ownership becomes important and these examples exist in interagency wildland fire. One example is that there may be multiple sources of an incident geospatial perimeter. If this is the case, if a downstream system, such as INFORM, wishes to obtain the official incident perimeter, which system owns it, and to which system should INFORM build an interface to obtain this data? (Data ownership applies to both geospatial and tabular data, as the concept of ownership is important regardless of the type of data.) As another example, there may be multiple sources of the location of an aircraft. Which system owns this data? If there is a downstream "Operational Data Store (ODS)" that is responsible for providing a near real-time picture of the allocation of resources to fires, from which application should the ODS pull the data for its operational view?

If there is clear data ownership for these pieces of data (fire perimeter, aircraft location), the interfacing and reporting questions are easy to answer. If not, applications may double-cross each other's data if both allow valid updates to the same piece of data. Reports may conflict if they draw data from different sources. Application builders and maintainers may be unclear of functional requirements if it is not known whether certain data fields are owned by that application or not. They will not be able to discern whether they should allow that data element to be created, updated, just read or deleted and what the impact of that change should be if updated data is received from some other system. Establishing data ownership resolves these issues.

The system that owns data is likely to be the system to control the Quality Assurance / Quality Control (QA / QC) processes surrounding that data. If a system owns a piece of data, it should own the review and approval processes for that data, or at least understand that it is farming that function out to some other application to return an improved value, but remains

the true repository for the data before and after the QA / QC process. Typical OLTP systems include the QA / QC processes for the data it owns.

Most definitions of a System of Record consider a System of Record as the "Authoritative Data Source" making these terms synonyms. The definition of an Authoritative Data Source might be bent to mean a downstream system that is connected to a System of Record that provides data to further downstream sources as an Authoritative Data Source. One website calls this downstream relay of data a "Source of Truth". The point is the same; however, data ownership begins with an owning System of Record / Source System / Transactional System / OLTP System / Application Owner so that downstream systems know that they are pulling the correct data values from the data owner, or they are pulling the correct data values from another downstream system that is pulling the correct data values from the data owner.

Interagency wildland fire IT is in a sensitive position regarding data ownership. It pulls data from cooperative firefighting organizations such as state and local governments and does not really "own" this data from a business standpoint. However, when Federal Government resources are involved in a major fire, Federal systems must track the operational state of fires and their outcomes and Federal Systems must report fire data and outcomes to high level leadership (e.g. the U.S. Congress), Interagency wildland fire IT must establish applications that own national, interagencydata for clear data processing flow through its systems even if this does not reflect true jurisdictionalownership of the data.

10.2.1 Master Data Management Model (MDM)

The IFDC may benefit from an analysis of implementing a Master Data Management model. A MDM for the IFDC can help address data quality and data integrity concerns. MDM is the comprehensive method used to consistently define and manage the critical data of an organization to provide a single point of reference for data that allows a set of permissible values. MDM serves data needs by removing duplicates, standardizing data (mass maintaining), and incorporating rules to eliminate incorrect data from entering the system in order to create an authoritative source of master data. At the business level, a strong data governance and data management model should be in place, then those models can use technology to implement the defined business rules and strategy.

Master data management has the objective of providing processes for collecting, aggregating, matching, consolidating, quality-assuring, persisting, and distributing such data throughout an organization to ensure consistency and control in the ongoing maintenance and application use of this information. At a basic level, master data management seeks to ensure that an organization does not use multiple (potentially inconsistent) versions of the same master data in different parts of its operations, which can occur in interagency organizations.

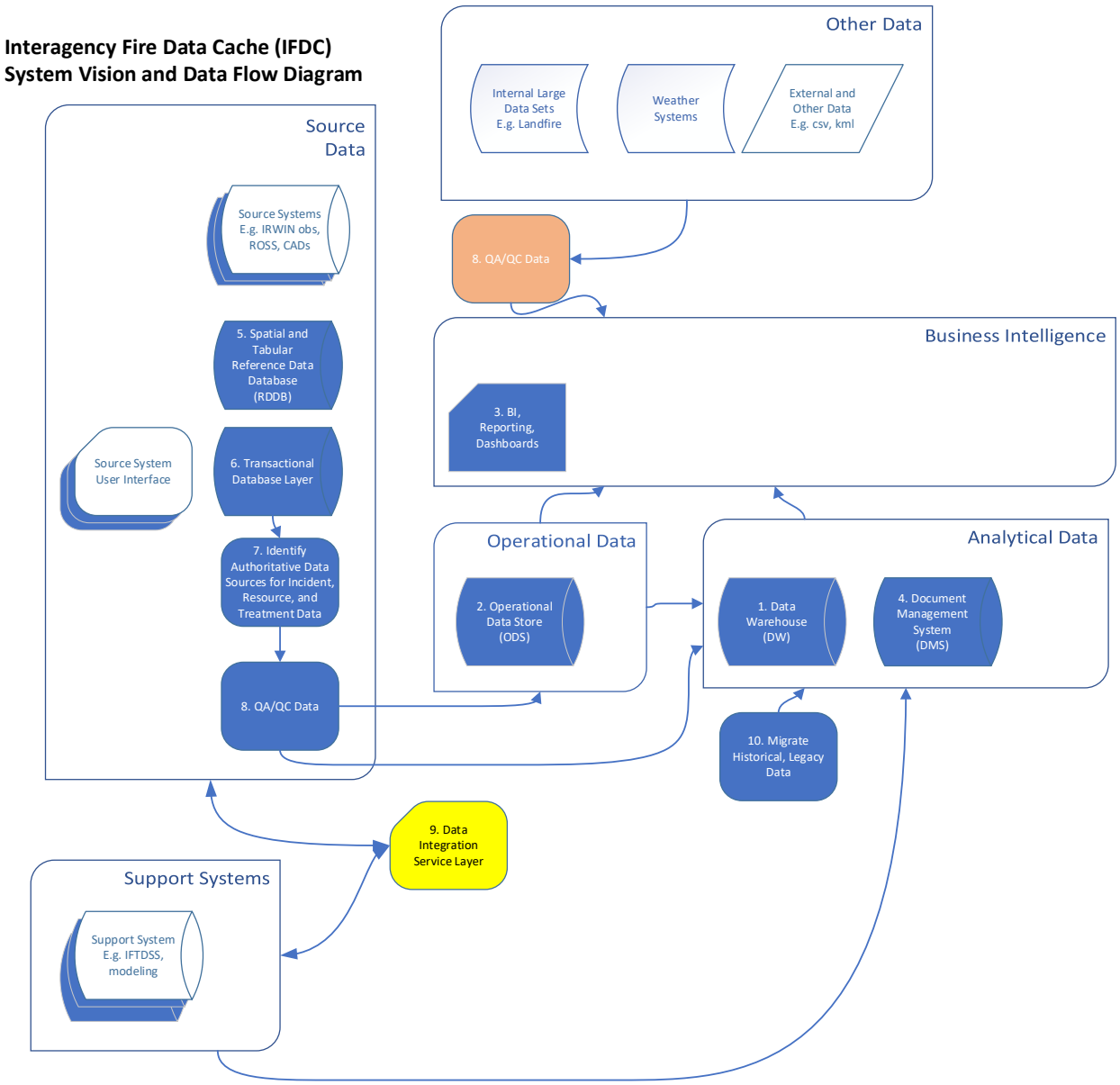
There are many tools available to implement the MDM that the IFDC team can evaluate, such as tools from vendors like Oracle, Informatica, and SAP. These tools can automate much of the business rule implementation, user interface for managing data, and integration with source systems and supporting systems.

10.3 Organizational Impact Considerations

Organizational Impact	
Tools	• Currently there is no QA/QC tool
People	• Users will require training on any new processes or tools selected
Process	• The implementation of the QA/QC Data solution is a vast departure from the way quality data is ensured today • Currently, users must rely on data coming from source systems as accurate data and there is no way to validate this • The new QA/QC process will mean that users will be able to rely on the data provided but could also experience some limitations if they are trying to create data or records that do not comply with validation criteria
Budget	• Depending on solution chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for

11.0 Goal 9: Data Integration Service Layer

Interagency Fire Data Cache (IFDC)
System Vision and Data Flow Diagram



11.1 Business Use

- Implementation of the IFDC requires a Data Integration Service Layer. This layer would broker data between transactional system, such as CAD, ROSS/IROC, WIMS, etc. This brokered data would need to connect to the current incident data exchange provided by IRWIN if it is not implemented in IRWIN.
- Currently, incident data is integrated via IRWIN and some resource data is integrated via the ROSS Service Bus. With the development of IROC, resource data integration is planned to transition to IRWIN. Fire environment and fuels treatment data remain

to be integrated. In addition, data from external systems like actual cost from agency financial systems is desired.

- The fire community can track the changes to data via metadata that is provided by the data integration service, similar to how IRWIN Observer functions today.
- The fire community can ensure data integrity by tracking the interagency authoritative data source (IADS) of the data.

11.2 Possible Technology Options

11.2.1 The Many Roles of IRWIN

The Integrated Reporting of Wildland-Fire Information (IRWIN) service is a Wildland Fire Information and Technology (WFIT) affiliated investment. IRWIN improves the consistency, accuracy, and availability of operational incident data. IRWIN is a central hub that orchestrates data among the various applications. Userscontinue to utilize existing applications but some or all of the data needed to create an incident, for example, will be pre-populated. Data is synchronized between participating applications to ensure the most current data is available. IRWIN conducts conflict detection and resolution on all new wildfire incidents to support a unique record for each ignition.

There is no single definition for IRWIN, as it is a multi-functioning tool that provides:

- **Data Integration Services:** IRWIN provides data exchange capabilities between existing applications used to manage data related to wildland fire incidents.
- **Incident Reporting:** IRWIN contains Observer, which is a primary source of metadata for incidents. It allows a user to search, filter, and display data about the exchange of incident data through IRWIN. This has become an important research tool for the wildland fire community.
- **Authoritative Data Source:** IRWIN identifies authoritative data sources for incident data, but it is also the Authoritative Data Source (ADS) for certain pieces of data, such as the IRWIN ID and duplicate incidents.
- **Transactional Database:** IRWIN is the back-end transactional database for the INFORM system that is in development.

It is noteworthy that IRWIN is also expected to be the data integration service for resource data via IROC as well as other types data.

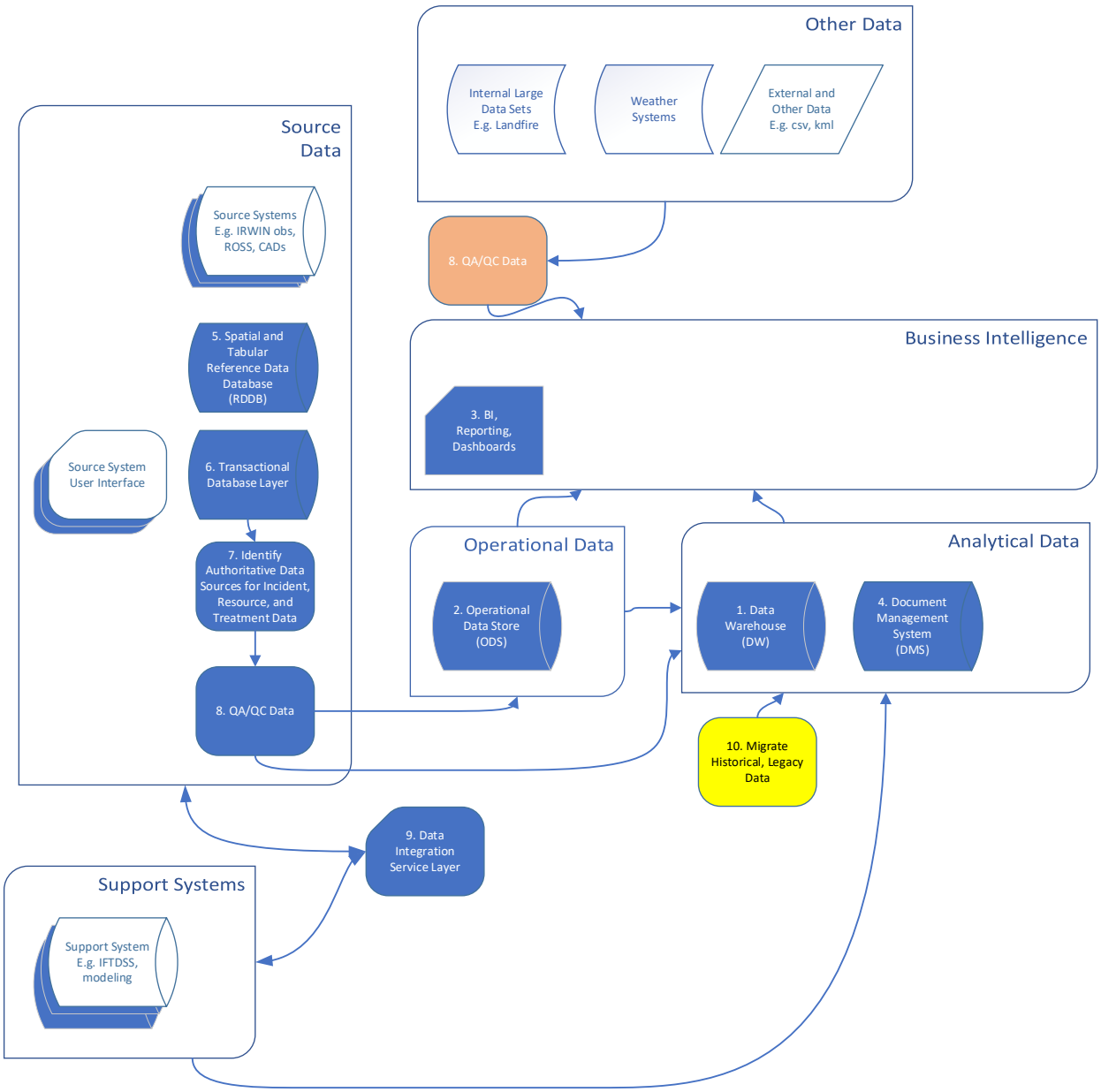
Because IRWIN serves multiple functions, it is helpful to discern and treat each function it performs separately. So, in this analysis, we will specify which function of IRWIN we mean for each area.

11.3 Organizational Impact Considerations

Organizational Impact	
Tools	• Adding the Data Broker functionality could happen within IRWIN or could be a completely separate system. In either case, the tools used and training needed should be evaluated
People	• Users will require training on any new processes or tools selected

Organizational Impact	
Process	• The implementation of the Data Broker solution for other types of data other than Incident is a significant process change and a large project with numerous process changes needed
Budget	• Depending on solution chosen, hardware, software, licensing costs may be incurred • Ongoing operations and maintenance costs should be accounted for

12.0 Goal 10: Migrate Historical and Legacy Data



12.1 Business Use and Benefit

- 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
- Would not be the repository for all historical data, but rather only data that is important and meaningful to the IFDC user community

  - There may be some exceptions for this, for systems that do not have their own historical databases for long-term storage
- If historical fire occurrence data were converted, would allow for one parent record per incident that includes the child records regarding a fire occurrence, rather than the multiple records that exist today
- All relevant data would be available for reporting and analysis

12.2 Risks and Dependencies

Migrating historical data can mean two different things. The first, incorporating an existing historical data warehouse into the Data Cache, such as weather data that is now stored at WRCC could be stored in the IFDC Data Warehouse. The second is using the Data Cache as the Interagency System of Record for historical transactional data. The second one is more complicated.

To illustrate, INFORM is intended to be the after incident report system of record. It will leverage IRWIN as its database for providing its data. There is no plan to gather historical incident data, prior to INFORM's go-live, and place it in INFORM (or IRWIN, ostensibly). So, this historical incident data needs a home.

If the Data Cache is the home for this data, the Data Cache becomes the Interagency System of Record for this historical incident data. It will need to offer the ability to edit this data through a GUI and would be subject to audits, and high user access levels as this data has already been published in various reports. These data changes would also need to be subject to some kind of QA / QC / Approval process. These functions resemble that of a System of Record, or source system, or transactional application, or OLTP application (synonyms). The validations and structure for this historical data should be similar to current and future data. This requirement /goal draws the Data Cache into the "being the System of Record", the Source System, for incident data.

A more common approach would be to store the historical transactional data with the new System of Record, in this example, INFORM (or IRWIN?). This whole conversation is confusing, because interagency wildland fire doesn't have a proper System of Record, or source system, for incident data. But if one existed, e.g. "BOISE", there would be a home for current and future incident data, and this historical incident data as well.

The concept extends beyond incidents – they may have historical transaction data for 209 (situational data by day), historical transactional data for ROSS (is this data being moved to IROC?), for aviation systems, for equipment usage systems, etc. Should source systems contain their historical data, or should this be a requirement levied on the Data Cache?

Migrated historical data would need to be flagged so users know it may not have been through the rigorous QA/QC processes that "modern" data has.
Migrating historical data requires a business decision about data that isn't collected in current systems, that may be pruning some data, holdings some data that is "orphaned" and likely not able to be edited, and recognizing that historical records may have blank fields.

For the second use case, we recommend evaluating which systems this would apply to, and if necessary, add historical data to a transactional database (7) and a User Interface to allow for editing. Then, this database becomes a source system for feeding the DW and DMS.

Another option for second use case for incident data discussed is to leverage IRWIN and INFORM to be the IADS for historical incident data.

12.3 Organizational Impact Considerations

Organizational Impact	
Tools	• This impact is TBD, as there may not be any new tools needed to perform this function.
People	• Users will require training on any new processes or tools selected
Process	• The Data Migration project would need to individually consider the process for each system that will be included in the migration
Budget	• This impact is TBD. See below for further considerations

12.4 Technology and Data Migration

Effectively migrating existing data from a legacy platform or tool to a new platform (data, documents and corresponding metadata), is a complex process. We recommend the following considerations for a successful migration:

1. What data is needed in the IFDC system? The requirements of the IFDC will help determine what data is needed, but frequently the exercise of mapping the source data can determine what data is missing from the system.
2. Where does this data need to come from? There will be multiple data sources even for the same record type. Define the ways to map the data and then determine which tables and fields contain the desired source information.
3. Does the data need to be transformed? It's an opportunity to use the data migration process to clean up dropdown options, business rules, required fields, etc. Data transformation rules need to be recorded in order to comply with the values and business rules for the new system. Some transformations can be as simple as just converting one value to another value or it could be as complicated as taking multiple,single records in the source system(s) and converting them to a single "parent" record with multiple related and associated "child" records in the new system.
4. Are there required fields in the new IFDC system that didn't exist in the source system?
5. Can the source data be filtered? Meaning, is there a logical date cutoff when the data is no longer relevant or valuable?
6. Can the data be cleansed in the source system, during transformation, or is it easier to clean it in the new IFDC system? Data cleansing is critical for a successful migration.

13.0 Organizational Mission Objectives and Key Results Areas (KRA)

The following table lists the organizational objectives and the KRAs that the IFDC project supports. KRAs answer this question, "What are the areas a system needs to address in order to meet the organizational mission?"

MBS derived the Mission Objectives from multiple sources, including the DataCache Briefing Paper_20161012_v2, Intro to the Data Cache, the Interagency Data Cache Briefing Paper DRAFT 5/15/2017, the NWCG_DMC_DataManagementStrategy_Draft_Feb1_2018, and WFIT Investment Management 5 Year Plan – Final documents, as well as stakeholder conversations.

Mission Objective	KRA
"Support the wildland fire mission with a source for interagency, national, landscape scale datasets"	Create a source for integrated interagency datasets based on NWCG data standards.
"…fundamentally improve the way we conduct information and technology to support fire business, not just refine existing silos"	Allow all stakeholders access to national, landscape scale, interagency, wildland fire data in formats that support current and future business needs.
	Access to national scale data to enable geospatial capabilities in current and future applications.
	Improve national level reporting capabilities.
" Assisting the wildland fire community to identify, define and standardize data that is reliable and accessible for planning, decision support, reporting and research."	Data is available in defined data formats.
	Data is deemed reliable and can be trusted for use in decision-making and reporting.
"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"	Data is exchanged programmatically among systems in such a way that promotes efficient and predictable processes.

Table 13.0-1: Mission Objectives and Key Results Areas

13.1 Project Key Performance Indicators (KPI)

The following table lists the key results areas and their anticipated business outcomes, represented as KPIs, in measuring the performance of a solution. Key Performance Indicators

(KPIs) are the measurable elements that accomplish the KRAs. In addition to the KPIs, MBS uses a standard set of IT Criteria that we consider KPIs and use to measure the KPI coverage of each alternative evaluated. The IT Criteria are

- Supportability (ability to find resources to support)
- Sustainability (how long will a solution last)
- Usability (how quickly will the user community adopt a solution)
- Training (how much training is required on a new solution)
- Complexity (how many systems, tools, and parts are combined to deliver a solution)

However, since we are not able to grade a solution, we have selected to view the IT criteria KPIs in these terms: How does the Data Cache help the wildland fire community with the IT Criteria KPI.

In the benefit analysis below, we used this full set of KPIs to grade each goal's benefit.

KRA	KPI
Create a source for integrated interagency datasets based on NWCG data standards.	• System enables users to link datasets across the enterprise • System makes use of integrated and deconflicted stored data so that there is one authoritative dataset • System enables user access to stored data for viewing, replicating, and reporting purposes • System enables users to access to historical records.
Allow all stakeholders access to national, landscape scale, interagency, wildland fire data in formats that support current and future business needs.	• System enables access to information from heterogeneous source systems, and can transform data into usable formats • System supports multiple business functions from pre-season planning, incident response and post-fire actions.
Access to national scale data to enable geospatial capabilities in current and future applications.	• System enables spatial analytics by utilizing services that use industry and government-approved geospatial standards and tools.
Improve national level reporting capabilities.	• System enables users to create dashboards and reports from individual or linked datasets • System enables ad hoc queries on individual or linked datasets • System enables users to efficiently measure and report on business performance measures, such as OMB and other national level reporting requirements.
Data is available in defined data formats.	• System enables users to store electronic documents/products in multiple formats • System has capability to perform QA/QC processes on the stored data • System allows access to centralized NWCG Data Standards.

KRA	KPI
Data is deemed reliable and can be trusted for use in decision-making and reporting.	• System enables deconfliction of Spatial and Tabular Data using defined business rules • System enables validation that the data being loaded complies with NWCG Data Standards • System enables services for external partner and other system access to fire data
Data is exchanged programmatically among systems in such a way that promotes efficient and predictable processes.	• System provides capability to control users/systems ability to Create, Read, Update, and Delete • System enables agency data stewards to be made aware of conflicts

Table 13.0 - 2: Key Results Areas and Key Performance Indicators

The table below is a subjectively-graded view of the relative importance of each goal when scored against the mission-derived KPIs. The goals have not been weighted by importance and so each goal is treated equally. There are two scores at the bottom of the table. The first score removes any "NA" cells from the total, meaning that if there are 22 rows and there are three NAs for a given goal, then the total score is divided by 19 instead of 22. The second score includes all NAs for each goal. We have scored them in this manner to show first how a goal stacks up using its relative strengths and weaknesses, and second to show how each goal stacks up against the mission as compared to the other goals.

KPI	Data Warehouse	Operational Data Store	Business Intelligence	Document Management System	Reference Data Database	Transactional Database Layer	Authoritative Data Source	QA/QC Data	Data Integration Service	Migrate Histo
System enables users to link datasets across the enterprise	4	4	4	4	4	4	NA	NA	4	4
System makes use of integrated and deconflicted stored data so that there is one authoritative dataset	4	4	4	4	4	4	4	4	4	4
System enables user access to stored data for viewing, replicating, and reporting purposes	4	4	4	4	4	4	NA	NA	4	4
System enables users to access to historical records	4	2	4	4	3	2	NA	NA	3	4
System enables access to information from heterogeneous source systems, and can transform data into usable formats	4	4	4	4	4	4	NA	NA	4	4
System enables spatial analytics by utilizing services that use industry and government-approved geospatial standards and tools	4	4	4	2	4	4	NA	NA	4	4
System enables users to create dashboards and reports from individual or linked datasets	4	4	4	4	4	2	NA	NA	2	4
System enables ad hoc queries on individual or linked datasets	4	4	4	4	4	2	NA	NA	2	4
System enables users to efficiently measure and report on business performance measures, such as OMB and other national level reporting requirements	4	4	4	4	4	4	4	4	4	4
System enables users to store electronic documents/products in multiple formats	NA	NA	NA	4	NA	NA	NA	NA	NA	NA
System has capability to perform QA/QC processes on the stored data	2	2	NA	2	4	4	4	4	4	4
System allows access to centralized NWCG Data Standards	NA	NA	NA	NA	4	NA	4	4	4	NA
System enables deconfliction of Spatial and Tabular Data using defined business rules	2	2	NA	NA	4	4	4	4	4	4
System enables validation that the data being loaded complies with NWCG Data Standards	2	2	NA	4	4	4	4	4	4	4
System enables services for external partner and other system access to fire data	4	4	4	4	4	NA	NA	NA	NA	NA
System provides capability to control users/systems ability to Create, Read, Update, and Delete	4	4	4	4	4	4	4	4	4	4
System enables agency data stewards to be made aware of conflicts	4	4	4	4	4	3	4	4	4	3
Supportability	4	4	4	4	4	4	4	4	4	4
Sustainability	4	4	4	4	4	4	4	4	4	4
Complexity	4	2	3	4	4	4	3	4	3	4
Usability	3	3	4	3	4	4	3	4	3	4
Training	2	2	4	3	2	2	2	2	2	2
	3.55	3.35	3.94	3.70	3.86	3.94	3.69	3.85	3.55	3.84
22 rows	71	67	67	74	81	67	48	50	71	73
	3.23	3.05	3.05	3.36	3.68	3.05	2.18	2.27	3.23	3.32

Table 13.0 - 3: Data Cache Goals Scored in Relation to KPIs

13.2 Agile Development

MBS recommends developing the goals of the IFDC using an Agile Development methodology. With Agile Development, the software development team performs short sprints (e.g. three weeks) to demonstrate software development progress frequently and allow for user and IT feedback frequently throughout the project. Agile Development is an alternative to Waterfall Development, where the development team creates extensive and complete requirements documentation, performs application design documentation and then

builds the application in its entirety to be delivered at the end of the project. Waterfall approaches can include more frequent deliveries to look more like an Agile approach; however, software deliveries would still likely occur after multiple months of effort instead of weeks. Waterfall development is more effective for building systems that have fixed or well-known requirements; agile development is more effective when requirements are likely to change or for more dynamic business needs. IFDC likely falls more into the latter category, as the science and business requirements of data collection, analysis and reporting may evolve over time.

13.2.1 Agile Work Plan

Based on MBS' experience with development projects for the DOI, MBS recommends three week Agile sprints. Four week sprints tend to be too long to receive timely feedback and two-week sprints are typically too frenetic given stakeholder workload. During these three weeks the following activities occur:

- Plan the Sprint – determine the software development tasks (backlog tasks) that will be accomplished in the upcoming sprint.
- Groom the Backlog – verify that the software development tasks in the queue for work are well defined, accurate, detailed and up-to-date.
- Develop User Stories – create requirements and design documentation for each software development task in a single document called a User Story.
- Create Test Scripts – create scripts for testing functionality in this and future sprints.
- Code – write source code for the application.
- Testing – test source code for the application.
- Write Documentation – write all documentation required to match the software being developed in this, or possibly the previous sprint.
- Demo / Retrospective – at the end of the three weeks, demonstrate the developed software, make it available for User Acceptance Testing, and gather feedback for the next three weeks of development effort.

After some sprints, the team may choose to release the software. Software release is the process of moving software from a development and/or test environment to a production environment. To accomplish this, the following tasks are required:

- Update Environments – verify that the Development, Test and Production hosting environments are ready to host the software, including the installation of any tools or programs that are required.
- Perform Security Testing – certify that the software meets IT security standards.
- Perform User Acceptance Testing (UAT) – allow the user community to test and accept the software as production ready.
- Resolve Issues – resolve all issues that arise in testing and the release process.
- Create Training Materials – create the materials required to facilitate successful training for the software.

- Create Implementation Plan – document the step-by-step actions required to move the software application(s) to the Production Environment, including data conversion, any manual steps, any database updates and any technical steps that are required.
- Finalize the Documentation Suite – finalize all documentation for the Production environment release.
- Gain an Authority to Operate (ATO) – gain the proper approvals to run the new software application(s) in the Production environment.

Prior to the first sprint, MBS recommends conducting a Validation and Planning phase. During this phase, the team meets to discuss overall application development requirements, standards and processes. This phase is also known as Sprint 0. Validation and Planning includes the following steps:

- Design User Experience – determine the branding and the standards for the software application.
- Initial Design Joint Application Design (JAD) / User Experience – determine the concepts that will lead to a successful user adoption of the application, such as creating a strategy for the placement and organization of data.
- Build User Story Backlog – Create a head start of user stories for continuous sprints.
- Start Test Scripts – Create a head start of testing scripts for continuous sprints.
- Create Database Schema – Create a physical database design.
- Support Development, UAT and Production Environment Setup – work with IT to setup application environments.
- Obtain Credentials – obtain credentials for team access.

14.0 Next Steps

Based on discussions with the team, below are the recommended next steps for the IFDC team to consider for each goal.

14.1.1 Data Warehouse

Possible technology options

- Forest Service's Enterprise Data Warehouse (EDW): Requires evaluation and discovery with USFS team to determine feasibility of use for IFDC DW.
- Forest Service's FAMWEB data warehouse: Requires evaluation and discovery with USFS team to determine feasibility of use for IFDC DW.
- Custom development, using a custom or COTS ETL tool and design
- USGS Science framework
- Explore possible other GOTS options

Business Analysis Tasks

- Determine what data is required in the DW and from which systems. The Data Dictionary effort should be collecting most of this information (e.g. element, description, who uses it). But there may be a need for a concurrent project to expedite the collection of these data elements specific to the DW.
- Determine what reports are needed and what data do they use for the reports (e.g. GPRA, OMB)
- Determine how often data should be refreshed to the DW. (e.g. weekly, monthly)

Pilot

- Pick one area of the DW and create a proof of concept or pilot for that area. (e.g. aviation data and reporting)

14.1.2 Operational Data Store

Possible technology options

- EGP: Evaluate how much of the ODS functionality the EGP is covering today,do a gap analysis of what else is required for the IDFC ODS and what the level of effort would be to add to EGP. Determine what other functions EGP is performing today outside of ODS functionality and if there are other applications that could or should take on those functions so EGP can focus on being the ODS.
- Custom development, using a custom or COTS ETL tool and design
- USGS Science framework (i.e. EROS)
- Explore possible other GOTS options

Business Analysis Tasks

- Determine what data is considered operational and needed for the ODS and which system provide that data. The Data Dictionary effort should be collecting most of this information (e.g. element, description, who uses it). But there may be a need for a concurrent project to expedite the collection of these data elements specific to the ODS.

- Determine what dashboards are needed, what reports are created and what data do users need operationally. (e.g. operational questions such as what's happening today on a fire)
- Determine the cadence for how often data should be refreshed
- Determine the cadence for how often data should be sent to DW, as well as what data should not go to DW?
- ROSS, WFDSS, InciWeb each have some functionality that acts as ODS but don't integrate data from other systems. However, if they're doing this function well, then it may be possible to leverage this for the integrated ODS
- Evaluate GeoMac and MesoWest, however GeoMac is a manual program (humans scraping data), and there is a desire to get away from this model and having conflicting datasets.

Pilot

- Pick one area and pilot it in EGP for ODS

14.1.3 Business Intelligence

Possible technology options

- EGP: Determine how much is EGP covering today for BI
- Determine which tools are already being used in the community. (e.g. Tableau, OBIEE, Cognos, PowerBI, etc)
- Evaluate the potential to leverage ESRI Insights, as IRWIN Observer is using this

Business Analysis Tasks

- Determine if one tool can be selected or if multiple tools can be used. It was discussed that some users will utilize the tool in varying levels of analysis, so one tool may not work for all users. However, the consensus was that there should be a tool that is simple enough for most users, but also allow the data to be available for users to access via other, more sophisticated tools.
- Determine the functional requirements. Consider geospatial needs
- Make a business decision as to whether transactional systems should use the BI tool for the system's static application-specific reports? (e.g. INFORM)

14.1.4 Reference Data Database

Possible technology options

- DOI Geoplatform
- AGOL
- EGP

Business Analysis Tasks

- Determine what data and which layers are fire-related and create a catalog of the data and sources
- Determine the cadence for refreshing data in the RDDB
- Identify the IADSs and ISORs
- Determine whether the data needs to be relocated from its current location to a centralized location

- Consider both tabular and geospatial data
- Determine how systems will access the data (e.g. API call to RDDB from other systems, MDM)
- Evaluate WFDSS in its use of reference data
- Determine whether Landfire data that gets changed for fire belong in the RDDB
- Catalog all the sources fire considers IADS for data so that if users need to use data they know where to go to get the data. This may not be part of the RDDB but should be considered.
- Discuss the location for Census, state, and other data fully adopted from other sources or standards
- Create a governance plan for the RDDB

14.1.5 Transactional Database Layer

Possible technology options

- If this becomes a Data Cache component, then the IFDC team should choose a technology to standardize to (e.g. SQL, Oracle, etc.)

Business Analysis Tasks

- Define the use cases for where creating this component will be useful
  - Create a list of applications that would leverage this service
- Build a business case for this option as a whole
- Create a process for each new system that is a candidate for this service to identify their business case and create a roadmap for onboarding new systems to service
- Evaluate the relationship of this component to the data management program
- Discuss IRWIN's role with INFORM as it relates to this Data Cache component and EGP's role for the applications it supports

14.1.6 Migrate Historical Data

Possible technology options

- NA

Business Analysis Tasks

- Determine which data needs to be migrated and from which source systems
- Create a cost benefit analysis for moving historical weather data that is currently at WRCC into the Data Warehouse
- Make a business decision regarding whether the DW is the archive for all data from systems or just enterprise relevant data
- Determine the data archiving strategy for source systems
- Coordinate with the states to access data if they are not using interagency tools
- Determine strategy for the IADS for incidents/source system

14.1.7 Data Integration Service (Data Broker)

Possible technology options

- IRWIN

Business Analysis Tasks

- Create a list of source systems for the different areas of data (e.g. resource, fire occurrence)
- Determine if the Program Board direction for using an integration service is sufficient
- Create the strategy for the sequencing and timing of datasets going through data integration service
- Evaluate the role of the National Incident Feature Service
    - Is it a transactional DB or data integration service?
- Consider geospatial data and if/how IRWIN can handle this

14.1.8 Authoritative Data Source

Possible technology options

- NA

Business Analysis Tasks

- Since this is primarily a data management issue, it should be considered as a separate analysis project that can run concurrently with the data dictionary project
- Assess resources necessary to complete this work as separate from data dictionary initiative
- Evaluate each business subject area and each piece of data and determine IADS for the data
- Perform gap analysis on system requirements to ensure that system is handling data effectively
- Evaluate systems that can implement the agreed-upon business rules and how they will do it
- Recognize that this is fundamental to the success of most if not all components of the Data Cache

14.1.9 QA/QC

Possible technology options

- NA

Business Analysis Tasks

- Determine at what levels this process should occur. (e.g. Source system, Data Integration Service, External data coming into DW, etc.)

14.1.10 Document Management System

Possible technology options

- FireNet.gov (Google Drive)
- IRMA data store
- Sharepoint
- Pinyon
- FRAMES
- AGOL
- FTP
- Fire Weather/Fire Behavior Archive Prototype

- Smoke Archive
- iMets
- Other GOTS

Business Analysis Tasks

- Determine what artifacts need to be stored
- Determine what metadata and tags are needed, as well as user access restrictions and user roles and permissions
- Identify and map workflow and approval processes.
- Create a user Interface
- Determine compliance requirements (e.g. 508, record management, Incident Planning SC guidance, etc)
- Identify requirements and create a process for records management and retention
- Identify requirements for disaster recovery, redundancy, and backup procedures
- Evaluate whether there is a need for a consolidated solution or if the current methods and procedures are sufficient.

15.0 Recommendation

The ten goals identified during this analysis are each important to the wildland fire community and should be considered for further analysis and future implementation. Based on the organizational mission goals and key results areas, as well as from stakeholder input, the consistent theme we heard was the need for consolidated reporting and centralized access to data, pointing to the Data Warehouse, Operational Data Store and Business Intelligence goals as providing the most business value. However, when we asked stakeholders at the end of the analysis project what they deemed most immediately valuable for them, many said the Reference Data Database is a very high priority need and others pointed to the importance of Interagency Authoritative Data Sources. We feel there are two possible explanations for the discrepancy. One is that the mission goals tend to point toward data integration, analysis and reporting. This implies a requirement to maintain the accuracy of operational data; however, MBS recommends the IFDC team should review the organizational goals and determine if there are any clarifications needed in that area.

Another possible explanation is that through the course of this analysis, many stakeholders were oriented toward the value of good data. While a technology solution such as a Business Intelligence tool can bring inherent data issues to the surface, the underlying data issues should be corrected in order for users to be able to trust that reporting and analysis is reliable and based on sound data. Meaning, the reports created are only as good as the underlying data used to create them. The goals that aid the most in this data management issue are the Reference Data Database, the Authoritative Data Source, and QA/QC. So addressing these upstream in the process will greatly enhance the downstream impacts and value the other Data Cache goals can provide.

Therefore, we recommend 1) evaluating the Reference Data Database, the Authoritative Data Source and QA/QC goals at the outset, while concurrently or shortly after, 2) evaluating the potential of existing systems to quickly meet the other goals of the Data Cache. Namely,

evaluating EGP for its Operational Data Store and Business Intelligence capabilities and IRWIN for its Data Integration Service and even Authoritative Data Source capabilities.

See section 14.1.4 for the list of potential next steps for the Reference Data Database, section 14.1.8 for Authoritative Data Source, and section 14.1.9 for QA/QC, and section 10.2.1 Master Data Management for information on how implementing a MDM can benefit the Data Cache.

There is tremendous potential in expanding the existing scope of some projects in order to meet the goals of the Data Cache, while also being mindful of the impacts of scope increases on existing projects. Expanding functionality of an existing system is recommended over adding functionality to an existing system, even if it is complementary functionality. An example of expanding scope would be to add integration of resource data to IRWIN. An example of adding scope would be to add FireCode algorithms to IRWIN.

Also, we recommend evaluating the possibility of using IRMA's Data Store for the Document Management System. There is potential for NPS to grant access to non-NPS users, at a minimum, or perhaps the Data Cache could have an instance of the Data Store created for their exclusive use.

16.0 Appendix I – Acronyms

ADS	Authoritative Data Source
AGOL	ArcGIS Online
API	Application Program Interface
ATO	Authority to Operate
ATO	Authority to Operate
BI	Business Intelligence
BIA	Bureau of Indian Affairs
BLM	Bureau of Land Management
CAD	Computer Aided Dispatch
CIO	Chief Information Officer
COR	Contracting Officers Representative
COTS	Commercial Off-the-Shelf
CRUD	Create, Read, Update, Delete
DLM	Data Lifecycle Management
DMS	Document Management System
DOI	Department of Interior
DW	Data Warehouse
EDW	Enterprise Data Warehouse
EGP	Enterprise Geospatial Portal
ETL	Extract Transform Load
FAA	Federal Aviation Administration
FACTS	Forest Activity Tracking System
FAMWEB	Fire and Aviation Management Web Application
FBMS	Financial and Business Management System
FEDRAMP	Federal Risk and Authorization Management Program
FMPC	Fire Management Program Center

FTP	File Transfer Protocol
GeoMAC	Geospatial Multi-Agency Coordination
GIS	Geographic Information Systems
GOTS	Government off-the-shelf
GSA	Government Services Administration
GUI	Graphical User Interface
IADS	Interagency Authoritative Data Source
ICBS	Interagency Cache Business System
IFDC	Interagency Fire Data Cache
IFTDSS	Interagency Fire Decision Support System
INFORM	Interagency Fire Occurrence Reporting Modules
IQCS	Incident Qualifications and Certification System
IQS	Incident Qualification System
IRMA	Integrated Resource Management Applications
IROC	Interagency Resource Ordering Capability
IRWIN	Integrated Reporting of Wildland-Fire Information
ISOR	Interagency System of Record
IT	Information Technology
JAD	Joint Application Design
JAR	Joint Application Requirements
KPI	Key Performance Indicators
KRA	Key Results Areas
LOE	level of effort
MBS	Managed Business Solutions

MDM	Master Data Management
MS	Microsoft
NFPORS	National Fire Plan Operations and Reporting System
NIFC	National Interagency Fire Center
NPS	National Park Service
NWCG	National Wildfire Coordinating Group
NWCG	National Wildfire Coordinating Group
NWS	National Weather Service
O&M	Operations and Maintenance
OCIO	Office of the Chief Information Officer
ODS	Operational Data Store
OLAP	Online Analytical Processing
OLTP	Online Transactional Processing
OOTB	Out of the Box
OWF	Office of Wildland Fire
OWF	Office of Wildland Fire
PaaS	Platform as a Service
PAID	Process Actor Interaction Diagram
QA	Quality Assurance
QC	Quality Control
RAD	Rapid Application Development
RAWS	Remote Access Weather Station
RDDB	Reference Data Database
ROM	Rough Order of Magnitude
ROSS	Resource Ordering and Status System
SaaS	Software as a Service
SIT	Situation Report

SMB	Small or midsize businesses
SME	Subject Matter Expert
SOR	System of Record
SQL	Structured Query Language
TBD	To Be Determined
TCO	Total Cost of Ownership
UAT	User Acceptance Testing
UI	User Interface
USDA	United States Department of Agriculture
USFS	United States Forest Service
USFW	United States Fish and Wildlife
USGS	United States Geological Survey
VM	Virtual Machines
VPN	Virtual Private Network
WFDSS	Wildland Fire Decision Support System
WFIT	Wildland Fire Information Technology
WFMI	Wildland Fire Management Information
WIMS	Weather Information Management System
WRCC	Western Region Climate Center
WYSIWYG	What-you-see-is-what-you-get