

**PHMSA's**  
**Data Management & IT**  
**Modernization Roadmap**

November 4, 2009

# Table of Contents

1.	Introduction.....	3
1.1.1.	Information Technology Program Review (ITPR).....	5
1.1.2.	DOT Inspector General and the House T&I Committee reviews.....	7
1.2.	PHMSA’s IT Modernization Roadmap .....	8
2.	People.....	13
2.1.	Organizational Roles and Responsibilities .....	14
2.1.1.	Strategic Analysis and Program Evaluation (SAPE).....	15
2.1.2.	Information Management Office (IMO).....	15
2.1.3.	Program Analysis and Risk Evaluation (PARE) .....	16
3.	Process .....	18
3.1.	Business Process Reengineering.....	18
3.2.	Data Architecture .....	20
3.3.	Data Quality .....	23
3.4.	Data Collection (acquisition) .....	24
3.5.	Reporting and Dissemination.....	26
3.6.	Data Management Processes.....	26
4.	Technology .....	29
4.1.	Current Challenges.....	30
4.2.	IT Modernization .....	32
5.	Summary.....	40



# 1. Introduction

**This Data Management and IT Modernization Roadmap** addresses deficiencies associated with PHMSA’s ability to effectively, efficiently and economically leverage Information Technology (IT) to collect, process, and timely make available quality (trusted) data and information to drive agency-wide safety decisions. Specifically, this roadmap identifies People, Process, and Technology initiatives that are required to overcome its weaknesses and to establish a robust analytical capability as a core program function in PHMSA’s safety programs. It should be noted, however, the data and IT deficiencies are broader than the Special Permits and Approvals programs of Hazmat Safety. Rather, this is an agency-wide (Hazmat and Pipeline Safety and Administrative Support) challenge requiring an agency-wide approach. Anything less will exacerbate the problem and continue to result in data quality challenges, redundancies, increased IT costs, and agency-wide inefficiencies.

## PHMSA Data Management includes:

- Organizational Roles and Responsibilities
- Enterprise-Wide Data Standards
- Data Planning and Governance
- Data Standards, Ownership and Compliance
- Data Quality and Availability
- Data Attribute/ Metric Management
- Data Delivery/ Performance Management
- Data User/Access Management and Storage

Figure 1 PHMSA Data Management

It should be noted that this roadmap discusses strategies to mitigate organizational gaps in the way PHMSA manages data and leverages IT to support its safety mission. That being said, there are, however, several initiatives currently in process within PHMSA designed to either improve the quality of information, increase transparency, increase accessibility, and/or improve the collection of data. For example, the Pipeline Safety office is currently developing a Data Quality Action Plan. This plan aims to establish formal roles and responsibilities governing the collection, use, and processing of data and information. This effort will compliment the process initiatives (specifically roles and responsibilities) discussed later in this roadmap. There are also initiatives underway to integrate the various data bases across PHMSA and modal and state partners. The Hazardous Materials Intelligence Portal (HIP) and the Pipeline Data Mart (PDM) have enabled PHMSA to bring together the disparate data into one repository. This was the first time the data from various systems was visible in one location. It immediately highlighted and identified significant data collection, quality, and reporting gaps. These gaps are discussed in greater detail throughout this plan. HIP and the PDM are also being used to make PHMSA data available and accessible to modal (FRA, FMCSA, FAA, and USCG) state (pipeline) partners. While these initiatives do result in nominal data quality improvements, it does not address the root cause (i.e., absence of People (dedicated and focused on data management) and Process (common and repeatable standards governing the life-cycle of data). Without addressing the root cause, PHMSA will be unable to leverage technology strategically to supports its safety mission.

## 1.1. Background

As outlined in the PHMSA Strategic Plan, goals of PHMSA include safety, environmental stewardship, reliability, preparedness and response, and global connectivity as well as the drive



to be a high-performing organization. To reach its goals, PHMSA must continue to strengthen its capabilities to attain what lies at the heart of all high-performing organizations, both commercial and government. These organizations have robust data analysis capabilities and consider data a vital strategic asset and make accessibility, transparency, data quality, and analysis critical components of the Agency-wide business strategy (i.e., business intelligence). Characteristics of these organizations include:

- Dedicated, centralized, and specialized strategically-focused analysts that define the data needs to support performance measures, continuously analyze data, and provide recommendations to leadership to strengthen business performance;
- Clearly defined policy and repeatable processes governing the enterprise use, reporting, and access of data and information;
- Agency-wide data quality standards to ensure data is trusted;
- Making information accessible, on-demand, in a consistent fashion across the business (flattening – horizontal approach)
- IT systems that are designed based on the data and information needs and that effectively enable mission execution;
- An integration strategy to bring together data from various systems into one repository for analysis, reporting and decision support; and
- An operational/tactical analysis capability to measure performance and identify improvement opportunities at the program and system levels.

PHMSA’s existing organizational structure and business approach poses several significant challenges to attaining a data analysis capability that drives program priorities, objectives, strategies, and enabling PHMSA to base decisions on solid analytics. The Office of Hazardous Materials Safety, Office of Pipeline Safety, and the Office of Administration operate relatively autonomously. For the most part, there is an absence of dedicated and specialized staff, common standards, processes, and technology platforms supporting and servicing these offices. (It should be noted that Pipeline Safety does have dedicated and specialized staff; however, there is an absence of roles, responsibilities, and accountabilities.) This fragmentation has contributed to the data analysis and IT challenges facing PHMSA. The good news is it can and will be corrected by addressing the issues at an agency level instead of the traditional focus at the level of individual operating programs (i.e., Special Permits). For example, both Hazmat and Pipeline have similar lines of business –

<b>Line of Business</b>	<b>Applicable to Hazmat</b>	<b>Applicable to Pipeline</b>
Regulations	Y	Y
Inspection/Enforcement	Y	Y
Incident/Accident Reporting	Y	Y
Outreach & Training	Y	Y
Special Permits	Y	Y
Approvals	Y	N



Grants	Y	Y
Registration	Y	N
User Fee	N	Y

Even though there are similar lines of business, the program offices leverage diverse technology platforms to support their objectives. Understanding there are differences (regulatory) between the two programs, there is no compelling reason the two programs cannot operate on common and standard technology platforms. The following example aims to illustrate how these inconsistencies perpetuate and contribute to data quality, decision support, and inefficient and ineffective use of information technology. Within the Office of Pipeline Safety there are in excess of five systems supporting the various lines of business. As a result of IT and data being decentralized across PHMSA, each office has the authority to develop systems. These systems are generally developed in silos and do not leverage common technologies (i.e., not even the same tools that are used on other Pipeline or Hazmat systems) nor a common database with standard elements. The system will not have an integration strategy defined for how the information collected and processed will be introduced into the analysis life-cycle (how data elements from different systems are linked and related). The end result is multiple databases that are disparate and lack common terminology and data definitions; absence of a standard approach to defining data quality (who defines the trusted view); and a technology platform that costs more to develop, maintain, and operate.

### 1.1.1. Information Technology Program Review (ITPR)

PHMSA conducted a formal Information Technology Program Review (ITPR) in 2008 to identify current and future business and technology performance gaps that inhibit PHMSA's Lines of Business' ability to effectively execute mission activities. The review resulted in a Transition Plan to enable PHMSA to close those identified gaps and attain its goal to support a rapid, data driven and risk-based approach and become a "conduit" for information sharing amongst PHMSA functions and external entities

The review concentrated on Information Management and focused on the following Operating Principals and Strategies of PHMSA:

- Improve Integrity of the System and Reduce Risk;
- Be a Model Safety Agency – One our Stakeholders and the General Public will Respect and Trust;
- Engage, Lead, and Help Strengthen the Capabilities of Others Who Share in Achieving our National Goals; and
- Anticipate Future Needs for Transporting Energy Products and Other Hazardous Materials.

The ITPR identified a series of gaps and proposed project activities to mitigate those gaps across four domains of the Information Management (IM) Maturity Model:



- **Governance:** Develop common standards, processes, and procedures that are repeatable, thereby improving PHMSA's ability to prioritize strategic investments. Develop an enterprise platform, organization-wide standards and common data architecture. Eliminate IT development activities that continue to perpetuate stove-pipe systems, result in long time-to-deploy timeframes, increased costs, and missed expectations.
- **Data Architecture and Management:** Institute a data architecture and management plan to include how PHMSA will define and access data, what governance controls will be in place to control changes to data, how it identifies common services and standards for data, how it will apply risk models, how it measure data, etc. Without this, PHMSA will struggle to become a data driven organization.
- **Organization** - Develop roles and responsibilities, rules of engagement, and follow common standards and procedures.
- **Technical Architecture** - Establish data-marts for the Office of Hazardous Material Safety (PHH) and Office of Pipeline Safety (PHP) and providing Business Intelligence (BI) capability. Provide repositories of clean and structured data specifically configured for analytical processing. Provide the capability to mine data, dynamically generate reports, automate analysis of data, apply risk models, provide enterprise-wide consistency in analysis, and increases current Program Analysts efficiencies.

The ITPR determined that PHMSA's business model has shifted to focus more on analytical processing and analysis. This shift requires a significant change in the way PHMSA designs and supports its Information Technology systems. Furthermore, the review established that PHMSA should begin leveraging analytical processing systems to support the requirements needed for analytical decision support and to improve its transactional systems. Innovative technologies will be needed to augment and/or replace legacy systems requiring state-of-the-art skill sets. The use of these new technologies will create capabilities that will help PHMSA become an information-focused model agency that optimizes mission alignment and effectiveness through real-time insights into business activities. The ITPR proposed a series of projects designed to strengthen PHMSA's information architecture through consistent, enterprise-wide platforms, strong and scalable data models, data quality standards, and data access/sharing capabilities. The projects will provide the vital Information Management framework needed to enable the transformation to a risk-based, data driven agency.

The ITPR concluded that the benefits gained through a strong Information Management framework include providing PHMSA the ability to quickly make tactical and strategic decisions regarding its systems' functionality and business requirements. An Information Architecture built with the information sharing needs in mind provides multiple channels for information dissemination and ensures that all systems and tools are clearly aligned with business rules, required functionality, and PHMSA's and strategic goals. Standardization and consolidation of data assets allow PHMSA to effectively analyze and use its enterprise information to quickly and efficiently make strategic decisions, encourages proper sharing and dissemination of information and ensures data is aligned to PHMSA's business and strategic goals. This strategy will allow PHMSA to turn data



into information, information into knowledge, and use that knowledge to make sound business decisions that strengthen the safety program.

### **1.1.2. DOT Inspector General and the House T&I Committee reviews**

Recent independent reviews of PHMSA by the DOT Inspector General and the House Transportation and Infrastructure Committee supported the ITPR findings stating that PHMSA's analytical capability is inadequate and decentralized resulting in significant inefficiencies and decisions being based on fragmented data and assumptions versus facts. These reviews identified several critical vulnerabilities, of which PHMSA concurs, including:

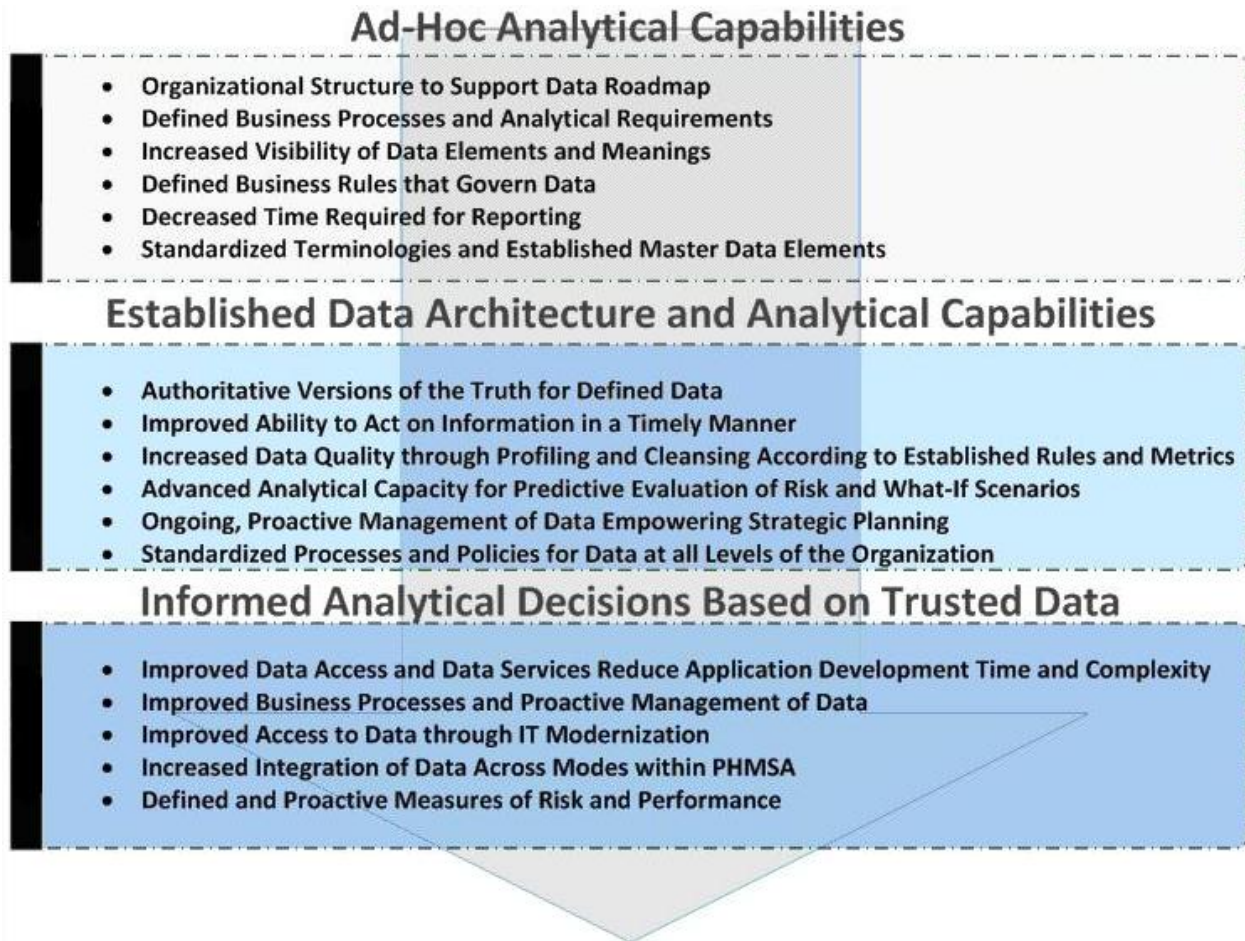
- Absence of dedicated resources for conducting strategic and operational analysis, establishing business rules and processes, applying consistent data quality standards, reporting, and executing program and performance evaluations and reviews.
- PHMSA has multiple IT Systems supporting the organization. These systems have no common/enterprise standards in place for how data elements within these systems should be defined (business owners set standards that are disparate); the data quality standards that should be applied; and policies and repeatable processes for use, accessibility, and reporting. This has resulted in data and system level stove-pipes, data integrity and quality challenges due to each system operating with a diverse and non-documented plan, etc. As a result, analysis efforts are very complex and generally are based on incomplete information. In addition, due to the ad-hoc data practices, analysts will usually reach different conclusions. This creates data trust issues and impacts PHMSA's ability to make sound safety business decisions quickly with precision and accuracy.
- Information Technology (IT) cannot be leveraged as a strategic business enabler due to the ad-hoc data practices and business processes.



## 1.2. PHMSA's IT Modernization Roadmap

The IT Modernization Roadmap details a migration strategy that incorporates people, process, and technology capabilities designed to enable PHMSA to leverage data to drive program priorities, improve our ability to detect emerging risks and target/focus our prevention activities, as well as to evaluate the effectiveness of our programs as a means of reducing risk

Throughout the migration the benefits of effectively managing data will be quickly recognized by utilizing an agile, iterative approach. Iterative organizational benefits that can be expected over time are outlined below:



The roadmap focuses on a three (3) year timeframe that allows the organization to mature and reevaluate the roadmap within each year as the organization and technology evolve.





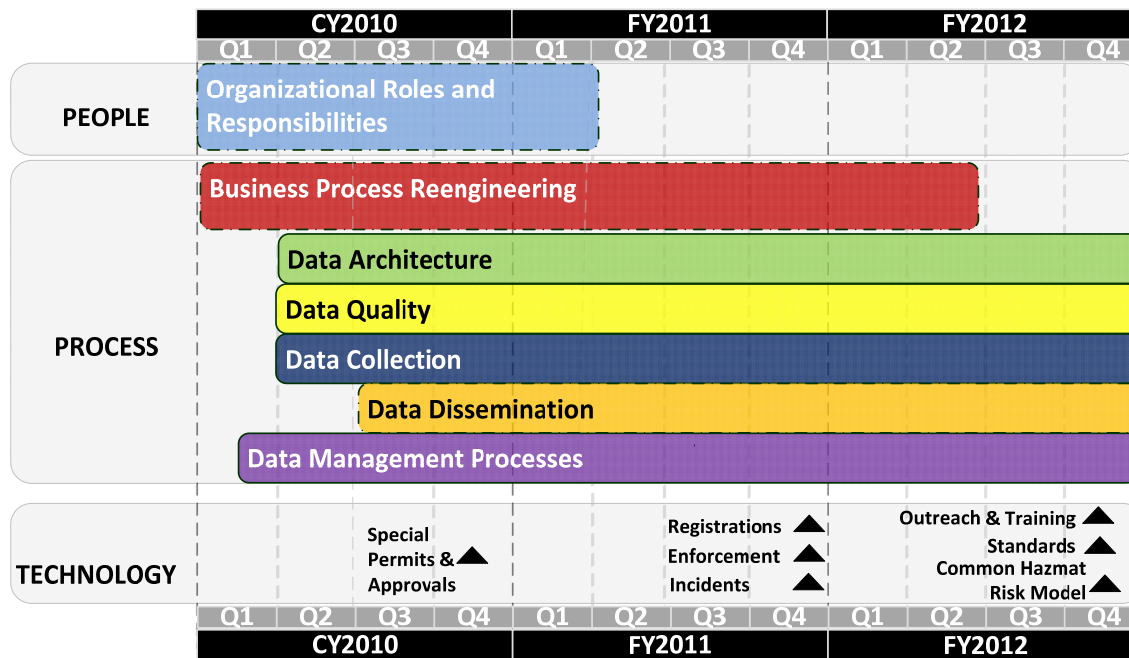


Figure 2 Data Management Roadmap Summary

The three years outlined in this roadmap provide the foundation for PHMSA’s longer term vision in years four and five. The end result is improved data quality, trust in available data, managed security and usage rights, easier access to data, and enhanced organizational performance.

Within the roadmap components, the following major initiatives are highlighted:

Component	Initiative	Organizational Capability
People	Organizational Roles and Responsibilities	<p><b>Organizational structure supporting Agency-wide Data Management:</b></p> <ul style="list-style-type: none"> <li>• Dedicated and strategically focused analysts</li> <li>• Common focus on standards, policies, and processes</li> <li>• PHMSA-wide data quality standards</li> <li>• Evaluation results that can serve as the basis for business process reengineering</li> <li>• Program offices working towards common goals</li> </ul>
Process	Business Process Reengineering and Management	<p><b>Business-driven approach defining Data Management Processes and Requirements:</b></p> <ul style="list-style-type: none"> <li>• Clearly defined, common performance measurements</li> <li>• Commonly understood and managed Business Processes Agency-wide</li> </ul>



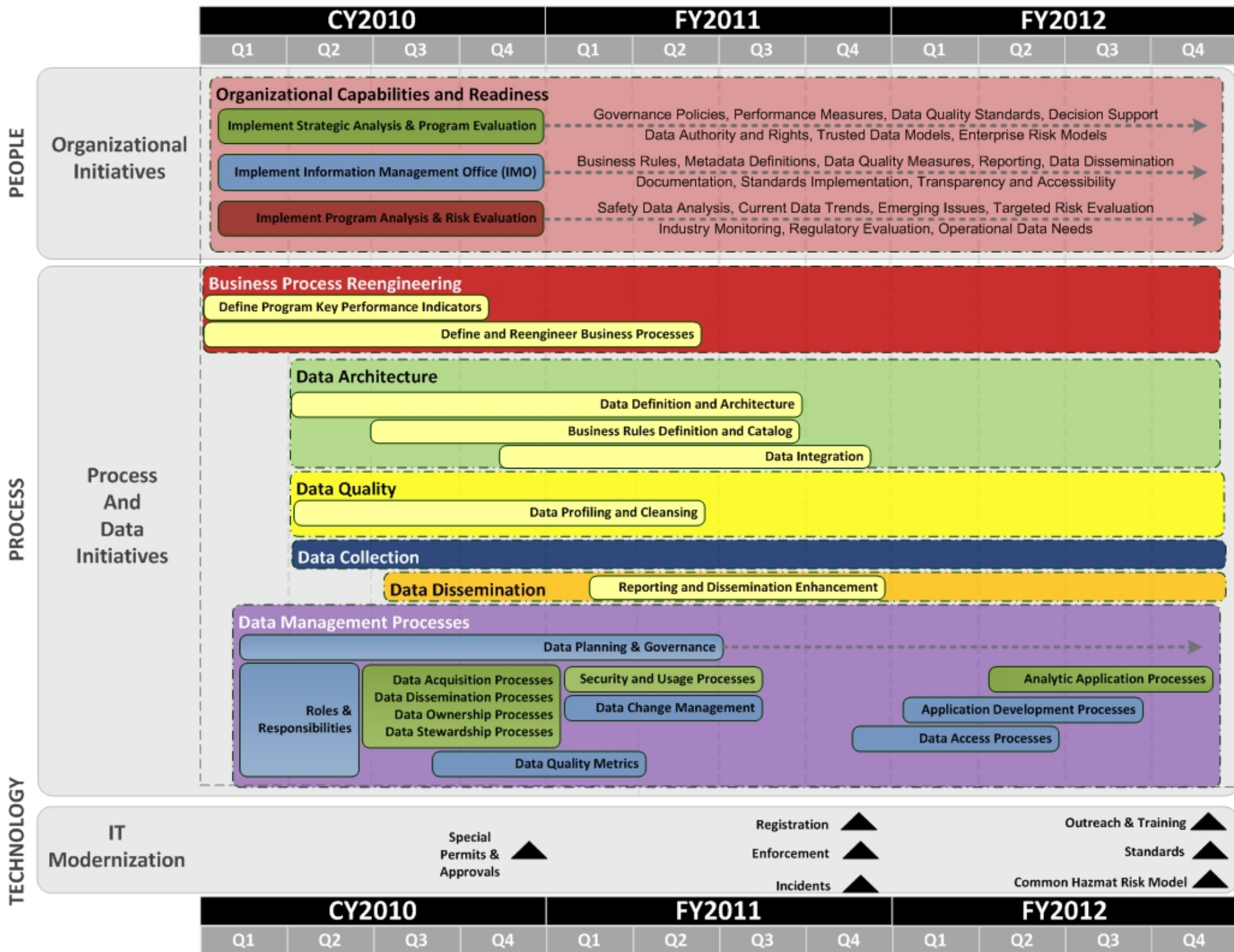
		Increased performance via Business Process optimization, reengineering processes, and lifecycle management
	<b>Data Quality and Governance Framework</b>	<b>Business and Technical Requirements enabling Data Management:</b> <ul style="list-style-type: none"> <li>• Common process and standards for data quality, transparency, security, accessibility, and distribution</li> <li>• Data quality metrics to define and ensure trusted data</li> <li>• Defined roles and responsibilities to enhance strategic analysis, program evaluation, program analysis, and risk evaluation</li> </ul>
	<b>Data Architecture</b>	<b>Processes and Governance enabling Technical Architecture:</b> <ul style="list-style-type: none"> <li>• Cataloged and searchable PHMSA data</li> <li>• Common business rules that conform data (IT) with process (business)</li> <li>• Tested, cleansed, managed, trusted data</li> <li>• Easy-to-use data structures for reporting and analysis</li> </ul>
<b>Technology</b>	<b>Technology Enablers</b>	<b>Technology-enabled Architecture:</b> <ul style="list-style-type: none"> <li>• On-demand access to trusted data &amp; information</li> <li>• Common and reusable technology platforms</li> <li>• Common data architecture &amp; transparency</li> </ul>





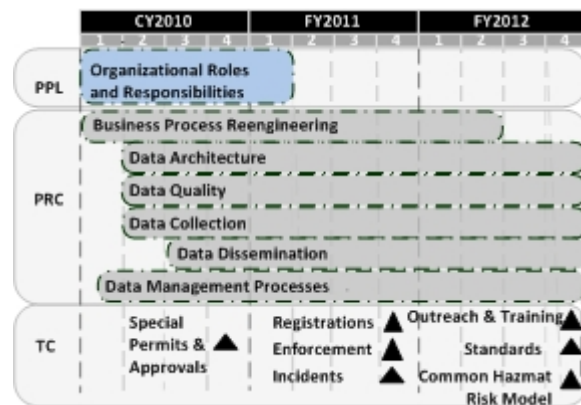
U.S. Department of Transportation

Pipeline and Hazardous Materials Safety Administration



## 2. People

PHMSA's currently suffers from a lack of resources (personnel) dedicated exclusively to strategic, operational, or tactical data analysis to drive program priorities, measure organizational performance, or effectively make safety decisions based on trusted and sound data. In the pipeline safety program, responsibility for data-related functions is fragmented, overlapping, and not clear to many in the program. In the hazardous materials safety program, the data related functions are understaffed with most analysis efforts being assigned to many different parts of the organization in an ad hoc way.



There is also an absence of resources to ensure that our various data programs can be connected with each other. In addition, there are no resources dedicated to establishing enterprise standards, processes (repeatable), or common taxonomies to manage, monitor, and ensure the quality and integrity of PHMSA's vital strategic asset: data. As mentioned, there are, however, staff that perform some of these functions as collateral duties which has resulted in ad-hoc processes further compounding PHMSA's data quality and integrity challenges. Furthermore, IT is decentralized across PHMSA which makes it difficult for the Chief Information Officer (CIO) to address data and IT challenges at an agency level.

There is no silver bullet to resolve PHMSA's data analysis and quality (inc. transparency and accessibility) vulnerabilities. To remediate the vulnerabilities, dedicated and specialized resources must be introduced to establish, adapt, and maintain an analytical capability that enables PHMSA to base vital safety decisions on trusted data (versus fragments of information) which drive organizational priorities, strategies, and program objectives. These dedicated resources will provide enterprise standards governing the identification, collection, storage, use, and reporting of data and information; consistent and repeatable business processes and application of enterprise data quality standards.

In summary, the following figure illustrates the proposed organizational functions that will be responsible and accountable for PHMSA's analytical capabilities and needs. These functions are described in greater detail in Section 2.1, Organizational Roles and Responsibilities.



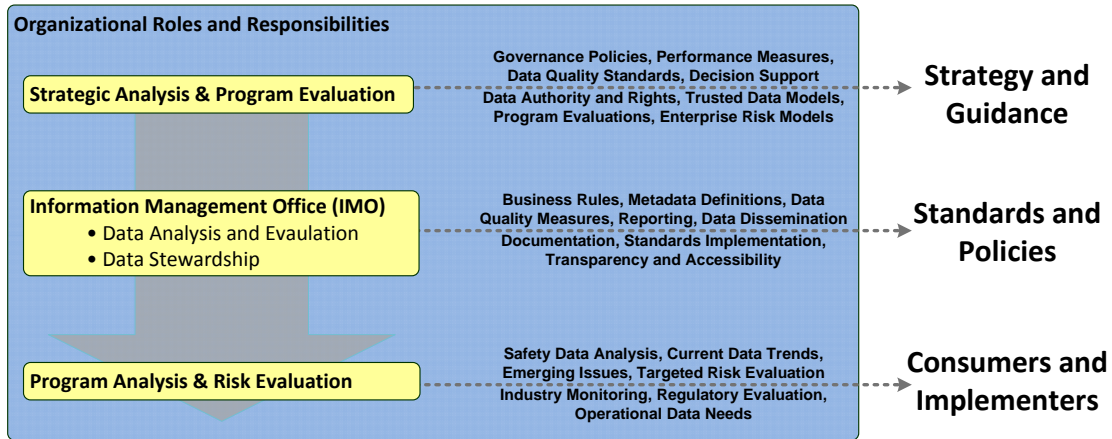


Figure 3 Proposed Data Management Organizational Structures

## 2.1. Organizational Roles and Responsibilities

Without dedicated support staff and the necessary organizational support structure, PHMSA is hindered in its ability to ensure data quality, provide access to trusted data, perform analysis and risk evaluations, support operational and executive decisions, and measure safety performance. PHMSA is currently in the process of conducting an organizational assessment to identify staffing needs and organizational structures to support and enable the data analysis capability. The target state organizational structure and resource plan will be completed on December 15. PHMSA, however, believes the functions described below are vital to support its data driven objectives. The following high level benefits can be realized with an organizational structure focused on data analysis, quality, and collection.

Capability	Benefits
<b>Dedicated and strategically focused analysts</b>	<ul style="list-style-type: none"> <li>Better analysis, performance measures, and recommendations to strengthen safety and risk evaluation across PHMSA.</li> </ul>
<b>Common focus on standards, policies and processes</b>	<ul style="list-style-type: none"> <li>Clearly define common policy and repeatable processes governing the enterprise use, reporting, and access of data and information</li> </ul>
<b>Agency wide data quality standards</b>	<ul style="list-style-type: none"> <li>Ensures trusted data, data cleansing, and remediation of poor data</li> </ul>
<b>PHMSA offices applying consistent approaches to data quality and Information Technology</b>	<ul style="list-style-type: none"> <li>Improved efficiencies and effectiveness through reusable components and lessons learned</li> </ul>



The following new functions are vital to form the foundation that will establish, adapt, and maintain PHMSA's analytical program.

### **2.1.1. Strategic Analysis and Program Evaluation (SAPE)**

Strategic Analysis and Program Evaluation (SAPE) - This team will be responsible for defining data needs on a strategic level, including:

- Governance Policies
- Performance Measures
- Decision Support and Analysis
- Data Authority and Rights
- Trusted Data Model (System of Record designations)
- Program Evaluations
- Enterprise Risk Model (working with the modal partners to identify common approaches/models to measure and assess safety risks)
- Identifying Leading Indicators

SAPE will develop and conduct program evaluations to assess the effectiveness of PHMSA programs, develop and monitor performance measures, conduct strategic assessments of industry and the operating environment, establish standards and policies for risk evaluation, establish a PHMSA-wide data governance structure, define enterprise data quality standards, provide decision support to the senior leadership team to guide agency priorities, and generally oversee the targeted risk evaluations conducted by the operating programs to ensure objectivity. SAPE will provide the core resources needed to evaluate PHMSA's operating programs, determine what works and what doesn't, and provide an analytical basis for redirecting programs. The organizational assessment will identify the actual needs and organizational placement.

### **2.1.2. Information Management Office (IMO)**

PHMSA proposes to establish a new Information Management Office (IMO) function under the office of the Chief Information Officer (CIO) that will centralize the **enabling and execution** functions across PHMSA and will provide the vital data and IT services to support strategic and program data analysis. This team will be responsible for the implementation of the standards, policies, and processes governing data throughout the agency, including:

- Application of Standardize Business Rules (repeatable)
- Common Data Architecture for use across PHMSA
- Metadata (Data Dictionaries, Diagrams, Descriptions)
- Enterprise Data Quality Standards
- Data Quality Assurance, Implementation and Measurement
- Reporting and Data Dissemination
- Documentation



- Standards Implementation
- Transparency and Accessibility
- Privacy and Security

The IMO will be composed of two groups and additional duties and functions of the Chief Information Officer (CIO) as follow: (The organizational assessment will identify the actual needs and organizational placement.)

**CIO** will be responsible for developing and overseeing PHMSA-wide data quality standards and managing PHMSA-wide data governance framework, working with SAPE to oversee integration of the data and systems architecture required for analytics, developing and implementing processes and policies for responding to changing information needs, implementing the established PHMSA data model and architecture, and providing IT capabilities to enable analytical capability.

**Data Evaluation and Analysis** will be responsible for working with the data and business owners to help define consistent and standard enterprise business rules, common and reusable data quality metrics, monitoring data quality, conducting program evaluations to estimate the extent or error or missing data in our data systems, defining and developing metadata to support analysis (i.e., consistent analytical interpretation), preparing data for PHMSA program and risk evaluation based on analytical needs, reporting and data dissemination, and working with SAPE to implement enterprise standards, policies and processes.

**Data Stewardship** will be responsible for overseeing routine data collection/acquisition processes, applying data quality standards, measuring and reporting on data quality performance as information is reported and processed, correcting data as needed, implementing PHMSA's common data architecture/model, and collaborating with analysts to resolve data quality anomalies and/or conflicts. .

### **2.1.3. Program Analysis and Risk Evaluation (PARE)**

The Program Analysis and Risk Evaluation (PARE) team will provide the basic intelligence necessary to understand hazardous materials and safety risks and to more effectively manage them. PARE will require personnel with analytical skill sets, including program analysts, data analysts, operations research analysts, statisticians, and economists and will be responsible for performing operational analysis and providing feedback to management teams and governance offices, including:

- Safety Data Analysis
- Current Data Trends
- Emerging Issues





- Targeted Risk Evaluation (i.e., lithium batteries, outside force damage of Pipelines, wet lines, etc.)
- Monitoring Safety Performance of Industry
- Risk analyses to support development and prioritization of the regulatory agenda
- Risk modeling for targeting resources (inspections, grants, etc.)
- Cost Benefit Analysis and Regulatory Evaluation
- Operational Data Needs

In addition, PARE will evaluate the current safety data, identify trends and emerging issues, conduct targeted risk evaluations of important safety issues, monitor regulated industry, oversee regulatory evaluations, identify data needs and provide input to program management. The organizational assessment will identify the actual needs and organizational placement.

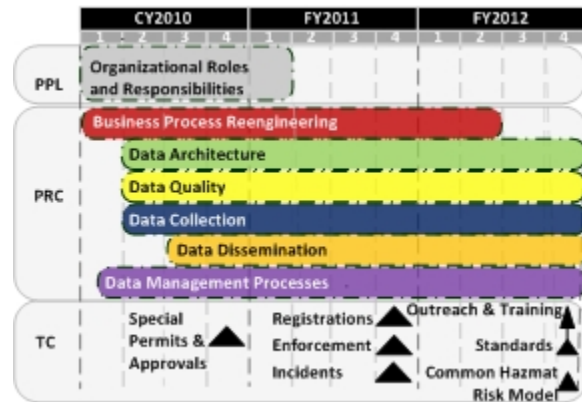
PARE will be users of the centralized and trusted data model defined by SAPE and PARE and implemented by IMO. This ensures the PARE is focused on conducting the necessary analysis to support program objectives and is not making changes to business rules, processes, and/or attempting to normalize data to form a trusted opinion.

By creating a specialized and process focused culture, PHMSA can begin to improve data quality and integrity, strengthen the linkage of program objectives to performance measures, ensure decisions are based on trusted enterprise data, and improve the performance and delivery of its IT systems. Upon approval of the December 15, 2009 target state Organizational Structure, PHMSA will update the PHMSA Organizational Manual per DOT Order 1100.74.



### 3. Process

A common process framework provides the building blocks that support the organization in carrying out its mission. Optimized, managed business and information processes are required to ensure data quality and consistency, and promote trusted data which will allow PHMSA to transition to a data driven organization where decisions are made based on trusted information.



The clearly defined and established roles and responsibilities detailed in the People Initiatives directly support developing, implementing, and maintaining the policies and processes of the Process Initiatives. SAPE, IMO, and PARE are tasked with establishing the common, repeatable processes that will ensure quality, accessible, secure, and trusted data. In turn the processes developed will support SAPE, IMO, and PARE in fulfilling their missions by providing the framework within which they conduct all data-related activities (from planning to collection, analysis, use, and eventually distribution).

#### 3.1. Business Process Reengineering

PHMSA’s current business processes have been developed as needed to support time and mission-specific goals for each Line of Business (LOB) within the Hazmat and Pipeline Safety Programs. Processes differ between each major safety Program in PHMSA but processes differ within each LOB (i.e., special permits, approvals, enforcement) within a Program. PHMSA needs to take an organizational approach to standardizing, re-engineering, and formalizing processes to improve data analysis and program evaluation, decision support, transparency, accessibility, and security at the Agency and Programmatic levels not simply within each LOB. Consistent data quality relies on the business processes that collect, manage, analyze, and distribute data. Because business processes form the fundamental basis underlying all PHMSA’s information needs, re-engineering and subsequently managing current ad hoc processes is an essential prerequisite to data quality, useful performance measures, risk evaluation, and strategic analysis; not to mention effective and economical use and delivery of IT systems and services.

At the very core of every activity requiring supporting data and analysis is an underlying business process. Suboptimal or informal processes lead to suboptimal or inconsistent data, hindering any efforts to ensure that data is accurate, timely, and trusted, which restricts management from quickly and confidently making decisions based on the data available.



By reengineering, optimizing, and managing business processes, the following benefits will be realized:

Capability	Benefits
<b>Clearly defined Key Performance Indicators</b>	<ul style="list-style-type: none"> <li>Improved progress tracking towards goals, measures of success, and risk evaluations</li> </ul>
<b>Commonly understood and managed Business Processes PHMSA-wide</b>	<ul style="list-style-type: none"> <li>Consistency in processes across PHMSA forms the foundation of standard, common business language and activities, which will also improve data consistency across the Organization</li> </ul>
<b>Business process optimization, reengineering, and lifecycle management</b>	<ul style="list-style-type: none"> <li>Replaces ad-hoc processes with more efficient optimized paths, reducing staff effort, training, and improving efficiency</li> </ul>

In order to ensure PHMSA is collecting the necessary data, PHMSA needs to review, validate, and revise its Key Performance Indicators (KPI) (e.g. what do we need to know, what data do we need in order to answer the question, do we have enough data, etc.) to strengthen PHMSA’s ability to make sound safety decisions that are based on sound data vs. fragments and assumptions. These KPIs will define business measures determining PHMSA’s success in meeting its desired business goals both internal and external, such as reducing the number of serious incidents, measuring the effectiveness of the compliance and enforcement activities, and reducing the number of special permits. The standard metrics by which performance is measured across the agency directly defines the data and information needs. These needs drive the business processes of the agency, which provide the roadmap for data collection, application of business rules to ensure data accuracy and consistency, and distribution of data to end-users, analysts, executives, partners, and others. A well-defined path from point A to point B eliminates redundant steps and duplicated effort, avoiding unnecessary costs, and reducing possibilities for error. Business Process Reengineering (BPR) seek to understand how business processes are currently implemented, compare the current ‘As-Is’ to the goal ‘To-Be’, redesign, and finally implement an optimal path. The Hazmat Program has already begun this process by reviewing the Special Permit LOB; however, to ensure maximum effectiveness other related LOBs that share or provide information need to be reviewed and re-engineered working iteratively rather than one LOB being re-engineered one at a time, so improvements are made across the Program succinctly and efficiently. For example, the Registration, Standards, and Enforcement LOBs are closely aligned with Special Permits and need to be considered when formalizing and improving the Special Permits LOB. Furthermore, the needs of similar LOBs within the Pipeline Safety Program should be reviewed to ensure the technology platform is designed with agency-wide needs in mind and to identify opportunities for synergies.

Data quality errors are often introduced by incorrect, informal, or suboptimal business processes. PHMSA’s current business processes are largely undocumented, in some cases as-hoc, and not generally understood. This lack of consistency is one of the key



factors implicated in PHMSA’s data quality issues and has led to the lack of trusted, authoritative data throughout the Agency. Without formal, re-engineered business processes and a proactive stance in managing and optimizing those processes, PHMSA will be unable to effectively identify and remedy data anomalies and errors. Defining the business processes and requirements is the first step in any IT Modernization activity. We must resist the urge to jump straight from concept to execution (i.e., begin redesigning the applications) as this will only perpetuate the problem (i.e., designing an IT system to support ad-hoc business processes ). Current issues have been exacerbated by IT systems accommodating informal and different processes to complete the same task, which has led to the data quality issues mentioned above.

PHMSA will continue the task of BPR by detailing the long term goals of both its human and technology-enabled business processes, and analyzing the maturation path required to reach the desired ‘to-be’ state. Inefficient, erroneous, or unneeded business processes will be redesigned and optimized to best support PHMSA’s mission and objectives. As previously mentioned, PHMSA is in the process of executing a BPR of Special Permits and Approvals. The deliverable is due in February 2010. PHMSA previously developed Standard Operating Procedures and documented the current “As Is” processes and workflow for Special Permits..

### 3.2. Data Architecture

PHMSA’s safety and regulatory mission – effectively its ability to proactively identify and manage risks – is currently hindered by disparate data, which leads to data quality issues, lack of authoritative data, and lack of insight and transparency into data available for analysis. In order to ensure decisions are based on trusted data, PHMSA must begin by establishing a comprehensive common architecture that ensures transparency, accessibility, and eliminates redundancies. PHMSA currently faces several data challenges such as standards for defining and collecting data are not consistent across the Agency, there is redundancy in data collection, each system maintains a diverse database with data elements redundantly but diversely defined, and data dissemination is performed by several groups using different business rules and data to disseminate information (resulting in different results depending on who runs the query). PHMSA needs to establish a common data architecture. The capabilities and benefits are listed below:

Capability	Benefits
<b>Establishes standards, attributes, and taxonomies</b>	<ul style="list-style-type: none"> <li>• Ensures data elements are defined and collected in a consistent manner across the Agency</li> </ul>
<b>Defines and catalogs all PHMSA data into a searchable repository</b>	<ul style="list-style-type: none"> <li>• Reduces time for data discovery</li> <li>• Improves understanding of available data</li> </ul>
<b>Establishes common business rules that conform with process</b>	<ul style="list-style-type: none"> <li>• Improves data conformance to standards</li> <li>• Allows data to be tested against data quality</li> </ul>



	metrics
<b>Tests data against business rules and defined quality metrics</b>	<ul style="list-style-type: none"> <li>• Improves data quality and trust</li> <li>• Marks bad data and allows for remediation</li> <li>• Increases data accuracy, timeliness, completeness, and validity</li> </ul>
<b>Integrates and conforms transactional systems for cross-functional analysis and reporting</b>	<ul style="list-style-type: none"> <li>• Simpler access to data for analysis and reporting</li> <li>• Conformed data increases quality by having one master source (one version of the truth)</li> <li>• Reporting-optimized structures speed queries</li> </ul>

The PHMSA Data Architecture will be predicated on the principal of reusability (i.e., plug-n-play). The output of the BPR activities will define the data and information needs (inputs and outputs) that are required to achieve established organizational and Departmental performance measures. PHMSA will employ common standards for the definition, use, storage, and analysis of data. This translates to establishing common data definitions (what does an incident cause code mean), business rules (how should it be collected and used throughout the process), meta-data (date of last update, assumptions or changes made), standard data elements (eliminating the redundant data elements collected and processed). By doing so, PHMSA will be able to establish a repeatable and trusted data architecture that minimizes the opportunities for redundancies. In addition, all PHMSA applications will be built leveraging this common data architecture versus the current practice of reinvention and silos. Furthermore, by establishing a data architecture approach, PHMSA can apply consistent and standard data quality measures and attributes to its IT systems. This ensures the quality of data is enforced at the time of entry and collection vs. many cycles later (i.e., the current environment) and will arm PHMSA with the vital trusted data it needs for decision support.

Within Hazmat Safety, one key critical gap is the absence of a common identifier for shippers and carriers. Each Hazmat Safety IT system maintains a separate company name database. In addition, PHMSA’s modal partners do as well. As a result of the sheer number of diverse company name databases and the non-standard process for recording company name information (i.e., UPS could be listed as U.P.S, United Parcel Services, UPS, UPS, Inc, etc. It could be listed this way in one and multiple systems. This makes relating the information very complex and inefficient.). A solution to this gap is the application of a common identifier (like the Operator ID Concept in Pipeline) for all shippers and carriers. PHMSA is exploring the possibility of leveraging its Registration program as a means to assign a common identifier. This would, however, require rulemaking to enact considering not all shippers and carriers are required to Register with PHMSA. This common identifier must be applied at the PHMSA and modal levels. This will ensure the company information is relatable across the various



modal and PHMSA level systems and will vastly improve the quality and integrity of PHMSA's data and information. PHMSA recognizes that leveraging a common identifier (i.e., Registration program) might take some time. To that end, PHMSA plans to leverage D&B numbers as an interim solution. PHMSA is currently using D&B information in its Hazardous Materials Intelligence Portal (HIP). This service, however, is very costly and is not comprehensive. As part of its modernization strategy, as defined in the Technology section, PHMSA will use a common company name database for all systems leveraging the interim D&B unique identifier. Modal systems would have the potential to plug into the common company name database as a look-up service to enforce and further standardize on the collection and recording of company name information.

Finally, a common data architecture will ensure transparency, accessibility and reliability of information. (provide the right information to the right people at the right time). In order to achieve this, data access and enterprise information integration is necessary.

PHMSA's strategic roadmap and ongoing data management initiatives have developed Target and Vision states for the overall enterprise data architecture that will support its mission, integrate with IT Modernization initiatives, and keep pace with the rapidly changing state of technology. Enterprise Information Integration (EII) provides for a defined layer of abstraction that integrates data from multiple sources within PHMSA and provides that data in a standard format for consumption by a variety of applications, end-users, and toolset. EII provides a standard, repeatable, and common way for users and developers to access and distribute (integrate) data throughout PHMSA, essentially allowing for plug-and-play functionality when data sources are updated added or retired.

The Enterprise Information Integration phase of PHMSA's data management program will unify structured, unstructured, and hierarchal data. Selection and implementation of a Data Services Platform during EII implementation addresses the following key issues:

- Data integration on demand
- Integration with multiple source systems in use within PHMSA and external parties delivering data used by PHMSA, including databases, packaged applications, messaging systems, third-party data feeds, text files, and spreadsheets
- View of aggregated enterprise information for easy access.
- Interoperability with current application interfaces data integration platforms and business intelligence and analysis tools in use at PHMSA
- Reduction in load on back-end systems for frequently used information
- An extensible layer providing flexible security, data access auditing, and logging.

PHMSA's success achieving enterprise information integration that provides meaningful information to decision makers is contingent upon its data quality processes and standards, which are discussed below.



### 3.3. Data Quality

PHMSA's current issues with data quality stem in part from the lack of adequate or consistently applied policies, processes, and standards governing data. By defining and implementing a comprehensive framework of processes that addresses data throughout its lifecycle – from its acquisition/collection to its final distribution – PHMSA can ensure consistency, accurately measure data quality, remediate data quality issues, provide transparency into current and future business and information needs, and provide better analytical capability for program, risk, and safety performance evaluation. In order to achieve this, the right personnel and organizational structure are needed to develop the standards, common taxonomy, business rules, data quality guidelines, and measures. Without the people (w/ accountability and responsibility) to define the standards and common taxonomies, only short-term surface improvements will be achieved.

There is no silver bullet to address PHMSA's data quality, transparency, and integrity issues. The data must be reviewed by analysts to determine alignment with the agency information needs and to develop the PHMSA trusted view (i.e., creating the apples to apples view of the disparate information across all data systems (i.e., Special Permits, Approvals, Enforcement, etc.)). There are two parts to data quality. 1) DQ1 – This is at a higher level and translates to having the governance, processes, and resources in place to ensure the integrity and veracity of the data. Examples include – How does PHMSA define the data and information needs, who defines the enterprise data quality plan, who defines how information should be reported/presented, who determines the single version of the truth, and how is information used to support agency decisions. 2) DQ2 – this is performed at the system level and can be thought of as edit checks and validation routines (i.e., did you enter the correct zip code format). DQ1 is a challenge for PHMSA due to the absence of a properly structured and resourced analytical capability. DQ1 is critical to supporting a data driven organization and for resolving the data quality and integrity challenges. DQ1 sets the standards, policies, and processes and adjudicates data conflicts. DQ2 can be performed; however, it will not be standardized and results will not be optimal due to the ad-hoc business processes. DQ1 challenges will be addressed by the SAPE and the PARE during BPR and definition of performance measures. Without addressing the DQ1 challenges or the resources necessary to perform the required pre-requisites, PHMSA will continue to experience data integrity and quality challenges and struggle to become an information driven organization. The IMO will address DQ2 issues by setting standards, taxonomies, and requirements for attributes and meta-data. In conjunction with the SAPE, PARE, and IMO PHMSA will soon begin establishing common standards, data quality metrics, metadata, etc with the Special Permits and Approvals LOB. Improvements to other LOBs will also be made to address overlapping processes that support Special Permits and Approvals such as Enforcement, Standards, and Registration. In order to ensure standards are being followed and anomalies are being identified and addressed before data is added to the trusted repository (data warehouse), the IMO will monitor, assess, and audit DQ performance.



Before PHMSA can assess and audit data quality, it must first define and implement a series of comprehensive data quality metrics for each data set.. Data quality is measured across a number of dimensions, not just whether data is correct. PHMSA's data quality metrics will include the following:

- **Accuracy** - The measure or degree of agreement between a data value (or set of data values) and a source presumed to be correct
- **Completeness** - The degree to which values are present in the attributes that require them
- **Consistency** - The measure of the degree to which a set of data satisfies a set of constraints
- **Timeliness** - The degree to which a data item or items is available at the time required or specified
- **Uniqueness** - The ability to establish the uniqueness of a data record and key data values
- **Validity** - The condition where data pass all edits for acceptability, producing desired results

Data profiling and cleansing is a large part of data quality. PHMSA will establish business rule definitions to test its data against those rules and defined data quality metrics. Data elements that comply will be given a high quality score and will not require cleansing. Data elements that do not comply will be candidates for cleansing. PHMSA's data quality processes will determine how to weight errors and anomalies and whether data should be rejected, automatically cleansed, or manually cleansed. Examples of data quality issues that can be cleansed automatically include capitalization and punctuation errors, non-conformance with a list of acceptable values or range of values, nulls or null-like values (blanks, N/As, etc.), duplicate records, and incomplete records. More complex issues and data anomalies will be flagged for rejection or examined and cleansed manually. Profiling and cleansing will occur in an iterative and ongoing fashion.

PHMSA is currently conducting a data quality assessment aimed to determine the quality of PHMSA's data as a sound basis for risk-based decision making. A draft report was completed on October 19, 2009 and is currently undergoing internal review. The final assessment w/ recommendations is scheduled to be completed on November 10, 2009.

### **3.4. Data Collection (acquisition)**

PHMSA collects, acquires and creates data throughout the course of its operations. Data is collected from operators, shippers, carriers, manufacturers, testers, state agencies, and others involved in pipeline and hazardous materials transportation activities. Data is acquired from other administrations within the Department, government agencies, and commercial sources. PHMSA creates data with every inspection, investigation, package test, special permit review, training activity, financial transaction, and other business actions.





Much, but not all, of this data is maintained in data bases. Data is received via handwritten forms, web-based transactional systems, automatic uploads from other systems. All transactions must conform to PHMSA's governance policies and standard operating processes in order to ensure that data quality is assured and the data can be integrated and used by others for decision making.

Each source of data should have a data collection or acquisition plan which provides sufficient detail on the goals and objectives for the use of the data, how it integrates with other data for decision making, and the standard operating procedures and processes used to collect or acquire the data. In addition, the data collection plan provides a basis for adherence to requirements in the Paperwork Reduction Act of 1980 and drives identification and development of IT support requirements. A data collection plan ensures that data is repeatable, reproducible, accurate, and stable.

- Data collection processes are *repeatable* when operators are able to reach an identical outcome, multiple times with a particular process or standard with the same or similar equipment.
- Data is *reproducible* when operators measuring the same items with the same equipment and processes reach the same outcomes.
- Data is *accurate* when an observed average measurement is the same as the associated standard value.
- Data is stable when variations of measurements over time are minimal.

Data collection governance, guidelines, policies and procedures need to be communicated to all data stewards and staff creating data as to what is to be collected and the rationale behind it. This should be followed up with some form of training or demonstration that will further enhance a common understanding of the data collection process as defined in the plan.

It address the who, what, where, when, and why of data acquisition/collection. For example, when a carrier completes an incident form on-line, data quality checks will be built throughout the entire process. Specifically, only certain characters will be allowed for specific data elements and pre-defined selections will be used when ever possible. Existing data such as registration number or carrier name will be validated against the PHMSA data warehouse (HIP/PDM) to perform the validation or possibly pre-populate certain fields to minimize data entry mistakes, yet the process will be flexible enough to allow a carrier/shipper/operator to update out dated data. These processes will also address what data needs to be shared, with which LOBs, and how. For example, any data element that needs to be used in more than one module will be obtained by using a look up function in the PHMSA data warehouse (HIP/PDM), which will ensure trusted data is being used and eliminate the creation of redundant data, which later comprises accuracy and overall data quality.

The correct application of process, technology, and roles/responsibilities applied to the data collection process is a crucial step in ensuring the quality of data that is used for decision making. For example, a common identifier is desperately needed for all Hazmat shippers and carriers to improve the ability to share data on an entity across DOT modes.



Currently PHMSA is using a company's Dunn & Bradstreet number to link data together from each mode on a single entity. The process is very expensive, not completely accurate, and cumbersome. To improve the process, PHMSA is exploring the possibility of expanding its Registration program to require all shipper and carriers to register with PHMSA. Shippers and carriers would use the Registration Number (unique identifier) to log into PHMSA systems to submit all required data (applications for special permits, approvals, registration renewal, incident reports, etc). This would allow information to be easily integrated and shared throughout DOT; however, it will require a regulation change, which will take some time before implementation can occur. Process, technology, and role-based enablers are all necessary to ensure a quality data-driven acquisition practice across PHMSA. These processes and data needs will be identified and documented. Data collection plans will be developed and refined during BPR and the development of KPIs.

### **3.5. Reporting and Dissemination**

Currently, there are several groups responsible for data dissemination throughout each Program and at the Agency level. This practice becomes challenging because each group pulls its own data from transactional systems and does not always apply the same business rules to the data to obtain the consistent results. Disseminated data in the correct forms to the correct people at the correct time is a central part of data management. Quality, trusted, and managed data provided as data services will ensure that applications and developers have access to commonly-governed, domain-specific data sets that have been modeled for specific usage yet are flexible enough to evolve with the applications' changing data needs. Data dissemination will be centralized under the CIO to ensure consistent reporting services are provided. This will be accomplished through the use of data services, which will ensure there is only one version of the truth, as defined by SAPE and PARE and enabled by IMO, and that information will not vary depending on who disseminates the data, as well as save analysts time by not having to recreate commonly used data. Business Intelligence, Reporting, and Analytic applications will consume, disseminate, and deliver actionable information to stakeholders through the PHMSA organization through on-demand, scheduled, and event-based mechanisms. These improvements will be implemented by the newly established IMO, as part of the Data and IT Modernization iterative process.

### **3.6. Data Management Processes**

Currently, PHMSA's Data Management Practices are inconsistent. There are minimal standards that are applied across the Agency or across LOBs; data is mostly decentralized and collected in silos within each LOB. If processes have been created to address quality issues they are manual and may vary from person to person or likely lost if that institutional knowledge leaves the Agency. To address these weaknesses, PHMSA will implement a Data Quality and Governance Framework, which will assist the Agency in defining, documenting, and implementing clear processes as guidelines and standard operating procedures. Implementing the Data Quality and Governance Framework will provide the following capabilities and benefits:



Capability	Benefits
<b>Common processes and standards for data management</b>	<ul style="list-style-type: none"> <li>• Integrated data management capability across PHMSA</li> <li>• Centralized views and management of data collection, usage, transformation, and distribution</li> </ul>
<b>Formalized Data Quality Metrics</b>	<ul style="list-style-type: none"> <li>• Ability to directly measure the quality of each data element and remediate poor quality data or work with modal partners to remediate issues found</li> <li>• Standard metrics allow comparisons of data quality regardless of source</li> </ul>
<b>Defined roles and responsibilities for data management teams</b>	<ul style="list-style-type: none"> <li>• Increased efficiency and reduction in duplicated effort</li> <li>• Managed, cross-functional teams collaborating across offices and operating programs will ensure decisions are made with the big picture (enterprise) in mind.</li> </ul>

The information below highlight tasks performed under the Data Management Processes:

### **Roles and Responsibilities**

Once PHMSA has the right personnel in-place (SAPE, PARE, IMO, data owners, and others) to begin addressing the data issues, it is imperative that each person understands what they are responsible and accountable for (the scope of their role), specific tasks they are expected to perform, and how their performance will be determined (critical success factors). As part of this process initiative, PHMSA will establish formal roles and responsibilities along with performance standards that will be included as part of the annual Performance Plan. These roles and responsibilities (governance framework) will define and include clear lines of delineation and rules of engagement for the definition, collection, use, storage, and dissemination of data and information.. The identification and establishment of the proper groups and ensuring each group understands their role and how and why their function is important is critical to the overall success of any transformation. For example, data owners must understand the cause and effect actions of their role relative to ensuring data quality.

### **Security and Usage**



Security and usage processes govern the controls put in place for each data element to provide protection against unauthorized access, sharing, and usage. Security processes will be put in place that govern and ensure that:

- Confidential data is kept secure, both at rest and in motion
- Systems are secured, at both a physical and software level
- Database controls are in place that restrict data access
- Standards exist for system access to data
- Control at the operating system level is limited
- Access to development and test environments is controlled
- Access to production systems is limited and traceable

Each component model in the architecture will have a responsibility in terms of security, but an overall set of standards and a complementary architecture will govern data security.

### **Data Change Management**

Data Change Management should ensure that only those with adequate rights can change data or data definitions and is essential to maintaining governance procedures. Data Owners will be responsible for setting the policies that govern data change. Data Stewards and IT together will implement those policies.

Data Change Management goes beyond simple authorization and access rights to the system(s) of record. These processes also outline in detail the steps and scenarios under which systems of record can and should be modified in support of overall data quality and usability within the PHMSA organization.

### **Data Access Processes**

Data access processes help define ways in which developers, data dissemination teams, analysts, and end-users can best use data to focus on their core tasks. Encapsulated, managed master data items become the core focus of data access processes.

Data Access Processes are a central tenet of data governance. Applications and developers will access PHMSA data through a modeled set of Data Services processes. Decision makers and knowledge workers will use business intelligence toolsets, analytic applications, and operational tracking reports in order to gain visibility into the day-to-day operations of PHMSA. Access to the system(s) of record and the actions of applications that clean, and conform data downstream from the system(s) of record should have a series of process in place to ensure that only data owners have access rights.

### **Application Development Processes**

With improved access to data and data services, application development processes can be developed that allow improved development functionality, faster turnaround, and enhanced usability. Application development teams will work in



a consistent manner across PHMSA to produce applications that have a similar look and feel, reducing training time and errors. This includes the development and adoption of a common Software Development Lifecycle (SDLC) that addresses analytical needs and is driven jointly by SAPE, IMO, and PARE. A centralized, common process of data access for application developers will be streamlined in order to reduce adding layers of complexity to the process of application development and robust in order to ensure that each application receives properly modeled data, at the right level of granularity, accuracy, and efficiency every time.

### **Analytic Application Processes**

Analytic applications allow personnel involved in heavy data analysis, statistics, risk evaluation, and data mining to work in consistent environments. Defining analysis and the applications that support analysis at this level are keys to ensuring that consistent analysis and risk points are used across PHMSA.

Once a commonly governed architecture for PHMSA data collection and dissemination is in place, analytic applications can spread rapidly throughout the organization. The Analytic Application Process begins with the need to measure a business process or its individual components. Each business process or component task/sub process can be decomposed into statements of granularity, data attributes and critical success factors (units of measure).

Once an understanding has been developed regarding the components of the business process needing measurement, targets for the critical success factors (the analytics) can be derived. Depending upon the analytic methodology embraced by PHMSA for Analytic Application Development (Balanced, Six Sigma, OAGO), target thresholds can be set. Managed goals can be incorporated into each level of the analytic process.

## **4. Technology**

The role of Information Technology (IT) is to both strategically and tactically enable the business to execute its mission in the most effective, efficient, and economical method. As PHMSA's business and mission requirements evolve so must IT to ensure consistent alignment with the mission. PHMSA recognizes IT is a strategic business partner; a partner that provides robust, secure, reliable, and low cost 21st Century agency-wide IT solutions and strategies to maximize business and safety performance. Through its modernization initiative, PHMSA will leverage IT to strengthen safety performance by providing on demand access to trusted (quality) data and information for decision and operational support. PHMSA's IT capabilities will include the provision of:

- A one stop shop to data and information that will ensure data is available horizontally in the organization vs in stove-pipes;
- A robust data collection capability that will enforce data quality standards and ensure the vital information needed to assess safety risks is collected;



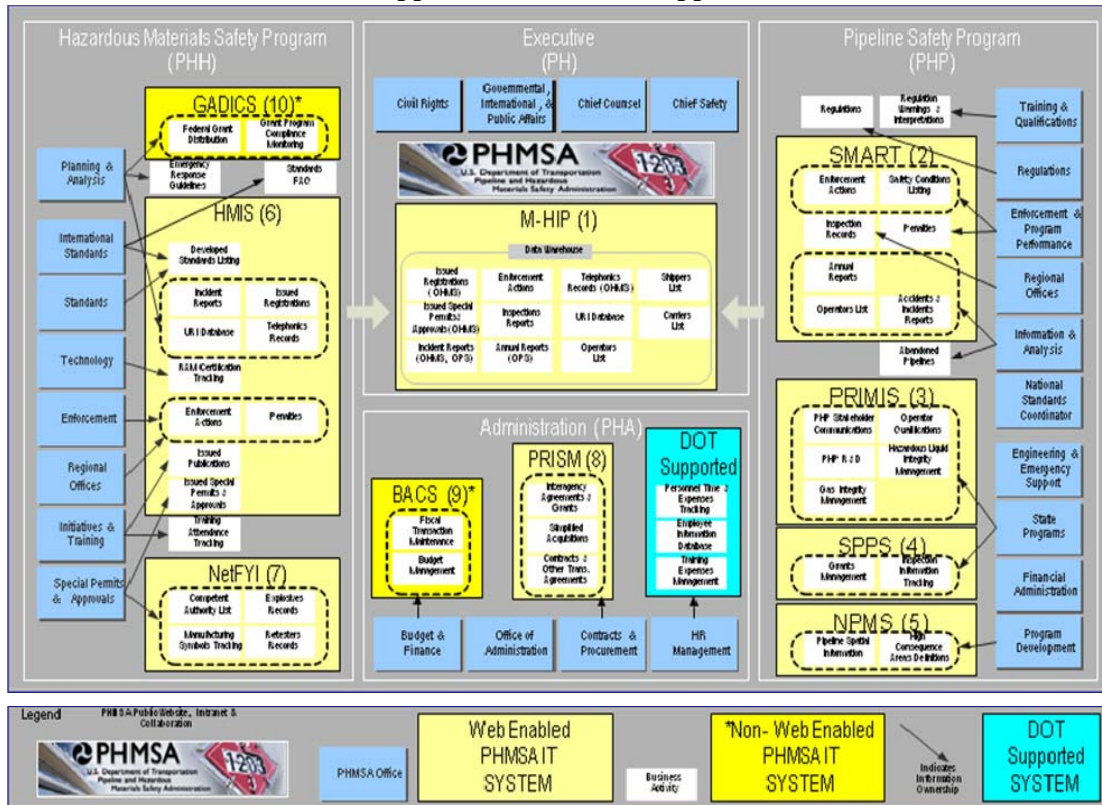
- Automated analytical services that will perform cross checks at the time of data collection and processing (i.e., did a shipper/carrier just get inspected, are there similar special permits, etc.);
- Business Intelligence to enable timely risk based and data driven decisions (i.e., trending and predicative analysis, direct limited resources toward the events posing the greatest risk, providing the analytical tools to identify training and outreach opportunities and target better rulemaking performance, etc.);
- A dynamic and reusable infrastructure to enable collaborative services and data sharing with other Federal, state, local partners;
- Web based strategies and services to communicate critical safety information to customers and stakeholders;
- Support of its diverse mobile workforce, including modal partners, to ensure the necessary tools, data, and services are available 24 x 7 to make critical risk based data driven solutions;
- Employing best practices for application development activities to ensure systems meet expectations are delivered on cost and schedule performance, and provide maximum business value.

#### **4.1. Current Challenges**

PHMSA's current decentralized and federated IT model contributes significantly to its struggles to leverage IT strategically. As depicted below, PHMSA has operated for over 20 years with 3 separate IT organizations. The results have been stove-piped applications, lack of standards, redundancy, significant data quality errors, diverse and obsolete technologies, increased costs, as well as ad-hoc business process and non-reusable technologies (i.e., redundancies). For example, the Special Permits application is not accessible or available to the Standards office. Systems used to process applications and collect data are built in silos and only address the requirements of one office and not the analytical needs of the agency (i.e., special permits information is needed by enforcement



and standards. However, the application does not support this business need.).



### PHMSA IT Current State

Within this environment, when a new IT requirement is identified by one of the program offices, the program office will typically develop the solution with no insight into whether the requirement could be supported with an existing solution within PHMSA (or the program office itself). This results in a new system, developed in a silo and supported by its own data repository. In some cases, modules within the systems themselves would be developed in this same silo scenario, resulting in redundancies within the application itself. (For instance, the HMIS application supporting the Office of Hazardous Material Safety has over six separate business modules (i.e. Special Permits, Enforcement, Incidents, Registration, Approvals, etc.) Each of these modules maintains a separate Company Name database that does not integrate with any of the other modules.)

This type of non-integrated IT planning, design and development spans across PHMSA and contributes to significant technology, data analysis and quality performance gaps inhibiting PHMSA’s ability to effectively execute its mission activities. It results in redundant and inconsistent data sources, conflicting business rules, lack of software development and data standards (diverse data elements, naming conventions, and definitions) and non-standard technology platforms. This approach is a major contributor to the existing data quality, integration and transparency issues PHMSA currently faces. By not leveraging existing resources, it drives up costs and schedule (longer time to implement due to not sharing services and data) and invariably results in solutions not



meeting expectations. Without a systems approach to IT modernization (which includes centralization of IT), the integrity and quality of PHMSA's data and information will continue to erode resulting in further data analysis inefficiencies. That is, decisions will be based on fragments of data and multiple versions of the truth while long wait times for data will remain an issue. As a current example, PHMSA enforcement personnel wish to identify high risk shippers, carriers, and operators. In order to do so, they need quick access to trusted data. Under its current model, PHMSA continues to struggle to provide data quickly and trusted.

## **4.2. IT Modernization**

During its IT Modernization, PHMSA will leverage the vital output from the Business Process Re-engineering initiative which will establish the business requirements, performance measures, data collection needs, and data quality standards. This will enable PHMSA to leverage IT strategically by ensuring the new capabilities are built based on the data and information needs (defined by the SAPE and PARE) of the agency. The IT systems will become tightly integrated and transparent (made available to partners, industry and citizens where appropriate). Industry standards and best practices will be leveraged to produce modernized systems and related data management processes resulting in the development, standardization and support of shared capabilities that are scalable, distributable, and designed to ensure technology remains in alignment with the PHMSA business models. These capabilities will be made available using state-of-the-art technologies and they will result in trusted data for decision support and better tools to direct PHMSA's limited resources toward events posing greatest risk. This modernization will result in unified risk-based and data-driven approaches to identifying high risk, high consequence companies and to drive business decisions, rule makings, and litigation activities. While the initial phase of the technology modernization is focused on mitigating the data and IT challenges within Hazmat Safety, the technology strategy being employed is a PHMSA solution. That is, as new business requirements emerge and/or at the time systems are refreshed that support PHMSA's Pipeline office, they will be supported by the common platform discussed below.

### **Target Data Architecture**

PHMSA has already begun laying the foundation for this IT Modernization. As depicted below, PHMSA has initiated the development of a data architecture model which supports both transactional (i.e., HMIS) and analytical (i.e., HIP and the Pipeline Data Mart) processing and is in the process of implementing components and systems which support each. PHMSA recognizes that its transactional databases (i.e., HMIS) must provide real time accesses to clean and trusted data while efficiently processing high volumes of transactions (applications, incident reports, rulemakings, registration, grants, etc.) while reflecting changes as they occur. It must also support analytical data produced directly from the transactional systems but which is designed for analysis and decision support. This model includes using Business Intelligence for analytics, reporting, and the development of risk models. To that end, PHMSA has established the Hazmat Intelligence Portal (HIP) and the Pipeline Data Mart (PDM) to integrate data from



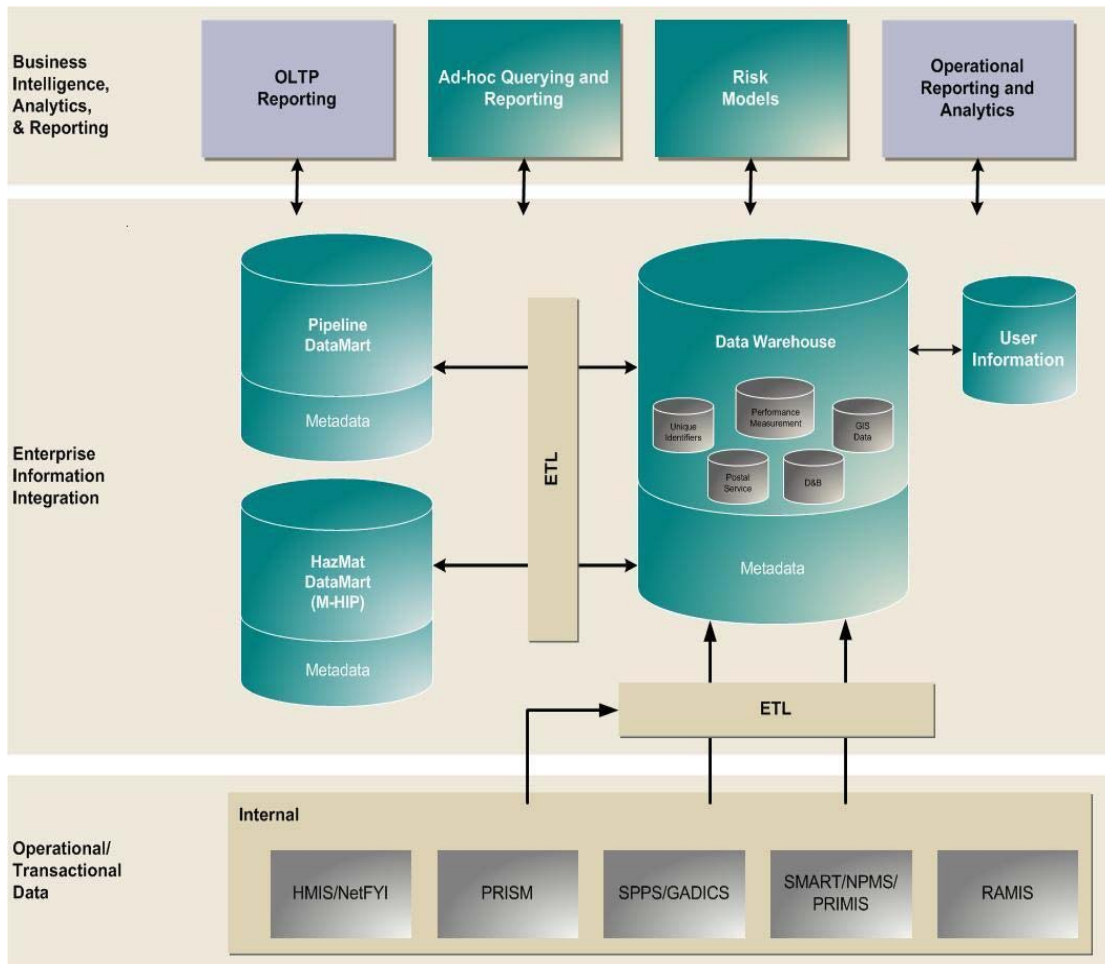


multiple, diverse, transactional data sources and provide ad-hoc reporting and querying capabilities. Continued development of this model will result in a source of PHMSA master data and a repository of integrated data to conduct performance measurement reporting and will support informational and analytical needs by providing integrated and clean, transformed enterprise-wide data for strategic, operational and tactical analysis. It should be noted that while HIP/PDM do contain most of PHMSA's data in one repository, the quality of the data is poor. The data contained in HIP/PDM is only as good as the source systems. While there are data quality improvements in HIP/PDM, the process is very inefficient and cumbersome due to the absence of common data quality standards applied at the entry and processing points during collection. As a result, data quality issues are recurring.

PHMSA is designing the common data architecture with the following guiding principals:

- Built leveraging common standards (standard naming conventions and taxonomies), data definitions (what does the data element mean), and meta-data (data about data – last update, assumptions, where it came from);
- Each system will leverage the common data architecture to eliminate diverse data elements and attributes (significant improvement in data quality)
- Data elements will have standard data quality parameters assigned and applied (identify and address dq challenges at the time of collection and processing before they become systemic);
- Data needs will be defined by the SAPE and PARE w/ common business rules defined by the IMO;
- A unique and common identifier will be applied for shippers and carriers (similar to the Operator ID concept employed by Pipeline).
  - PHMSA will leverage Dunn and Bradstreet (D&B) numbers initially to link and relate PHMSA, FMSCA, FRA, FAA, and USCG data. However, PHMSA views this as an interim solution as D&B services are very costly and are not comprehensive. PHMSA prefers to leverage its Registration program to require all shippers and carriers offering or transporting hazardous materials to register with PHMSA. This will require a rulemaking to enact.
- Provide a one stop shop to trusted data and information.
- Ensure transparency, accessibility and reliability of information. (provide the right information to the right people at the right time)





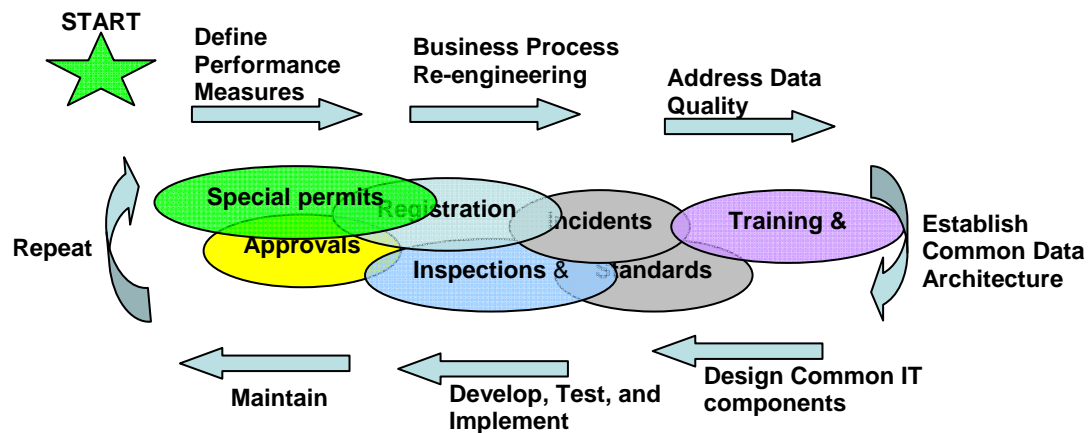
**PHMSA Target Data Architecture**

**An Iterative Approach**

PHMSA’s approach to IT Modernization will be iterative, structured and business driven resulting in critical analytical capabilities aligning business, data, and technology. An example of the process PHMSA will use to modernize its applications is depicted below in PHMSA’s Iterative Process diagram. In this example, the current HMIS application will be dissected module by module in order to address the identified performance gaps that have become evident. Performance Measures will be established and Requirements and Design will be initiated only after the business processes (current or newly defined) are identified and validated. The existing and (newly established) data requirements will be identified and associated data quality measures will be developed. The process would feed into PHMSA’s Common Data Architecture and would include defining, maintaining, and enforcing the use of standardized data element names, abbreviations and definitions, as well as the definition of data quality attributes (i.e., edit checks, pre populating data, using pick-lists). Existing flaws in the data models will be addressed and



redesigned to support Common IT components. For example, in order to eliminate the multiple Company Names in HMIS and the modal partner systems, D&B numbers will be used as the interim Common Hazmat Identifier until such time PHMSA can require the Registration of all hazmat shippers and carriers. All HMIS modules (Special Permits, Approvals, Enforcement, Incidents, etc.) will leverage the newly defined Common Hazmat Identifier to link and relate company records (i.e., creating a shipper/carrier profile or scorecard). This iterative process will continue for each module within HMIS resulting in a consolidated HMIS application. These Common data Architecture and IT Components will be leveraged (where appropriate) by other applications PHMSA-wide.



**PHMSA Iterative Process**

**Concept of Operations**

As envisioned (**PHMSA Concept of Operations**), there will be direct integration between the transactional systems (e.g. HMIS) with the analytical systems (e.g. HIP/PDM). A “One Stop Shop” portal will be realized where industry, citizens, modal and state and other business partners access PHMSA services via the Internet or Intranet from a computer or handheld. The following paragraph describes planned functionality and demonstrates the collaborative and analytical capabilities. (Please refer to the illustration below).

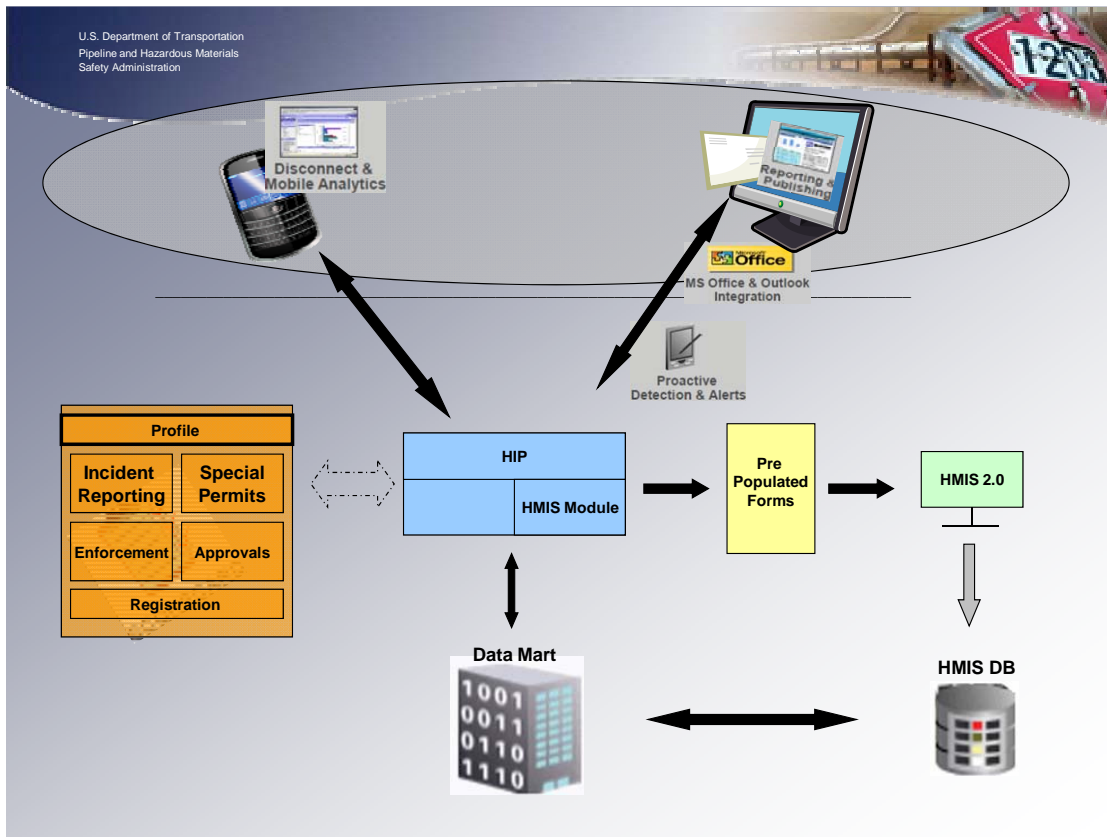
Once logged in (PHMSA staff or industry using the common identifier) , the company profile will be displayed (Yellow Block on the following illustration). All relevant information PHMSA has available regarding the entity logged on would be displayed (example: Registration information, Scorecard or Rating information, Special Permits, Approvals, Incident information, Enforcement actions (from PHMSA and the modes) assigned



to a given Shipper/Carrier/Operator, etc.). The user would have the ability to update their profile on-line and perform a variety of functions associated with their particular profile. For instance, where appropriate, they could perform such tasks as renewing a Special Permit or Approval; applying for Registration, or submitting Incident Reports. Depending on the function being performed the user would be directed to forms that are pre-populated using the existing information regarding the company and/or transaction from PHMSA's data warehouse (HIP/PDM). This eliminates the need for data entry and ensures data quality. The user would only be required to provide additional data needed for PHMSA to assess safety risk and process the application. Upon submission, all data will be evaluated against the data quality standards. In addition, the system will conduct checks to determine if there are open enforcement actions (i.e., did the company just get inspected and is now applying for an approval), in the case of special permits or approvals – are there already special permits available to support this need, if an approval – enforcement would be notified if fitness determinations are required, provide notifications to modal partners, etc. Depending on the transaction, the relevant data would be processed thru the appropriate business process transaction (including modal partner reviews where appropriate), the transactional database (HMIS) updated and appropriate updates made to the data warehouse. If appropriate, the user would be able to upload attachments to transactions, check the status of any requests, and perform a variety of other functions relevant to their profile.

This same scenario would exist for new external Industry users. That is, users would have the ability to register for on-line capabilities using a D&B number (that will be validated with D&B to confirm identify) and once registered; they could apply for permits/approvals, check statuses, review their score cards, make updates, etc.





### PHMSA concept of Operations

PHMSA recognizes this modernization strategy will take some time to complete. However, by leveraging an iterative approach, benefits will be realized in 120 – 180 day increments and will allow PHMSA to constantly re-evaluate strategies based on current and future business needs. PHMSA understands that, in the interim, business services must continue and improvements to current capabilities must be provided. To that end, PHMSA will also execute the following near term activities designed to improve current services while the modernization efforts are underway. Note: The short term strategies below are designed to provide interim limited capabilities and will not be the final solution provided during execution of the IT modernization strategy. The near term strategies will be implemented using existing resources.

- Stabilize the current HMIS platform – on-going. Efforts have been very successful thus far.
- Implement Special Permit Evaluation Forms (i.e., Party Status) – December 2009
- Automatically generate and send letter to holder if renewal application has not been received w/i 60 days of expiration – December 2009.
- Modify the HMIS workflow to facilitate modal reviews (and allow access to modal partners) – December 2009



- Provide an Itinerary Planner (IP) function in HIP (planned capability) – December 2009. The IP is an automated inspection scheduling tool that uses the current enforcement risk model, called the National Business Strategy, to identify the perceived highest risk companies. The IP requires an inspector to select the highest risk company from the list. In addition, the system forces the inspector to select other high and moderate level risk companies as the inspection week is planned. Investigators will create itineraries which will be made available to other Investigators for review. Managers will be able to review itineraries, approve them or reject them. More efficient planning will result using GIS and Calendaring capabilities to view trips and inspection overlap will be avoided enabling PHMSA Enforcement offices to direct limited resources toward the events posing greatest risk.
- Prepare a detailed statement of work for the IT modernization services. Contract award is expected to occur in February 5, 2010.
- PHMSA will lead an effort with its modal partners to identify requirements and options to the common identifier challenge. – Proposed plan by February 2010
- Finalize the Business Process Re-engineering (BPR) of Special Permits and Approvals. This effort is scheduled to be completed in February 2010.
- Establish an interim on-line application (establish forms) capability for Special Permits and Approvals to enforce data collection standards. This interim capability will be on-line in March 2010.
- As a temporary patch to the collaboration gap, leverage HIP to generate analytical reports (i.e., target fitness inspections, support the analysis of which special permits to codify, development of early warning indicators as special permits near expiration, etc.) – ongoing

In addition to the near term projects referenced above, PHMSA plans to complete the following activities in CY2010, contingent on the availability of funding:

- Review, validate and establish new performance measures for Special Permits and Approvals;
- Establish a risk model for Special Permits and Approvals;
- Establish and implement common data attributes, definitions, meta-data, and data quality rules in support Special Permits & Approvals.
- Establish and implement policies and processes governing the collection, use, storage, and dissemination/reporting of data and information.
- Implement the required infrastructure to support the modernization efforts (i.e., single sign-on, portals, new technologies and development tools, etc.)
- Implement a common database in support of both Special Permits & Approvals. This database will be the foundation that all other modules (Enforcement, Incidents, Registration, etc.) will leverage.
- Clean the current data to the extent practical.
- Using the target state Business Process Reengineering document as a baseline, a detailed systems design and requirements document will be prepared that



translates the business and data requirements into an IT blueprint that application developers will use to construct the new Special Permits & Approval module

- Develop the new Special Permits & Approvals application
  - Will include the capabilities identified under near term and those described under Concept of Operations.

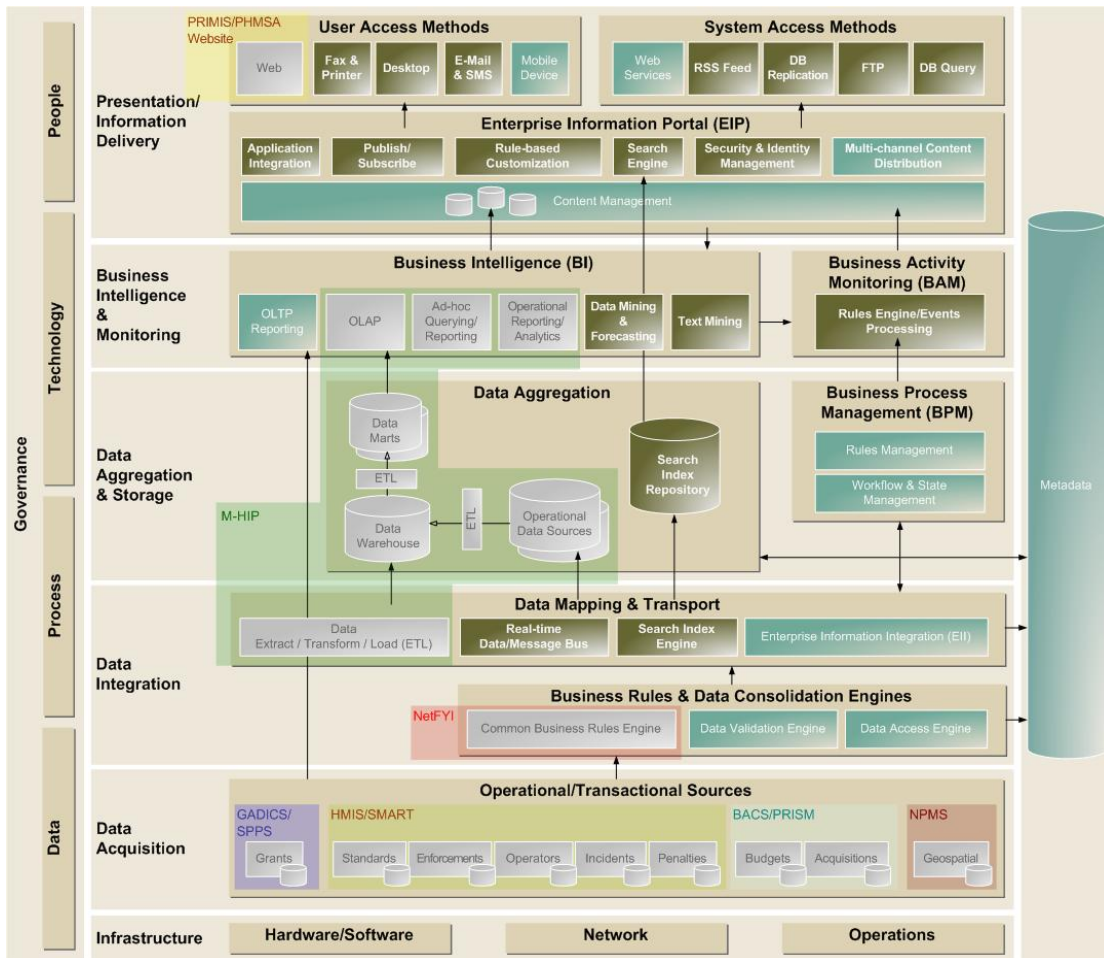
In FY11, PHMSA will repeat the steps defined under FY 2010 for Registration (common hazmat identifier), Enforcement, and Incidents. For FY12, PHMSA plans to modernize Outreach and Training, Standards, and implement a common hazmat risk model. In FY13 and 14, PHMSA plans to further mature the transactional and analytical services by establishing and providing GIS capabilities, providing the analytic support platform to support predictive analysis, developing content and information management portals, enhanced reporting and presentation of information to make analysis efforts more efficient, and to help fund operations and maintenance of the new systems.

PHMSA will leverage existing resources to complete the short term actions. For the mid term actions, PHMSA plans to reprioritize funding to begin the vital modernization activities. Longer term funding strategies will be included as part of the budget process.

### **Target Technical Architecture**

Using centralized data and standard and repeatable business processes will produce consistent, reliable and accurate reporting of data and information from a robust and sophisticated technology platform (see **PHMSA Target Architecture** diagram below). This service based approach will be leveraged to eliminate stove-pipes and redundancies, ensure reusability, and establish common standards and tools for use across PHMSA. Automation will be leveraged, providing transparency and promoting more timely collaboration and access to information for users, the regulated community, modal partners, and stakeholders. PHMSA will begin to improve data quality and integrity, strengthen the linkage of program objectives to performance measures and improve the performance and delivery of its IT systems. This technical architecture will promote data sharing with PHMSA's modal and state partners as they access PHMSA portals and HIP/PDM.





**PHMSA Target Technical Architecture**  
People, Process, Technology (and Data)

## 5. Summary

In Summary, PHMSA has developed this comprehensive 5 year plan as a strategy to provide PHMSA with the necessary people, process, and technology capabilities to address the data analysis and IT deficiencies within the organization. This plan will result in a robust data analysis capability that drives program priorities, objectives, strategies, and decision making while aligning technology with business and data needs. The plan contains a 3 year roadmap which allows the organization to reevaluate within each year as the business processes and technologies mature. The 3 year roadmap will contain sequencing objectives for each year and focus on business and IT modernization while Years 4 and 5 will focus on refining maturing, and maintaining the processes and capabilities developed in years 1 thru 3. The 3 year roadmap leverages an industry proven, iterative business based approach which is crucial to ensure the alignment of business, data, and technology. It is designed to provide immediate short term benefits which will be leveraged for all PHMSA's IT Modernization initiatives.





The first year of the roadmap specifically addresses the modernization of the information technology systems that support the Special Permits and Approvals Programs. Specifically, within Office of Special Permits and Approvals, PHMSA plans on implementing short-term and long term solutions which will result in immediate enhanced productivity, accountability, and overall management of the safety functions assigned to that office. Short term goals include the analysis and cleansing of existing data; common data definitions and data quality rules to support Special Permits & Approvals; development of a common database in support of both Special Permits & Approvals; and the development of on-line application capabilities which will be available to industry with a focus on requiring applicants to submit the data needed to make a safety determination while resulting in improved data quality thru the use of standardized (automated) forms. Longer term efforts include a more advanced and integrated on-line application which heavily leverages the components of HIP for such thing as pre-population of these forms with existing data; developing specific reporting and query capabilities; provision of early warning indicators (example: did a shipper or carrier just get inspected and is now applying for an emergency Special Permit?); sending advanced renewal notifications to shippers and carriers; and automated collaboration with enforcement and modal partners before Special Permit approval.

As this plan is executed long-term, PHMSA will become more of a learning organization with its culture more analytically focused. PHMSA's IT systems will support rapid, data driven and risk-based approaches allowing PHMSA to leverage data to drive business decisions and increase safety performance.

