Project Summary

U.S. Department of the Interior
National Park Service
Natural Resource Program Center
Inventory and Monitoring Division

Integration of Resource Management Applications (IRMA) and the Natural Resource Information Portal (NRInfo)

Updated June 7, 2010

1. PROJECT OVERVIEW

Natural resource data and information are a critical resource, as evidenced by the numerous management and planning decisions, information products, and services that result from its analysis and synthesis. Managers, planners, interpreters, scientists, and other staff need to be able to efficiently search for, retrieve, share, and disseminate data and information in order to effectively manage public lands, and the data and information obtained through various studies and management actions need to be shared with the general public and various constituencies, “for it is the broader public that will decide the fate of the resources.”

The Natural Resource Program Center (NRPC) is transforming the way it manages and delivers natural resource information to parks, partners, and the public. IRMA, the Integration of Resource Management Applications, is the name given to the project that is guiding this transformation. Modern information technologies, and in particular, service-oriented architecture (SOA) that IRMA is based on, are providing tools, procedures, and protocols that allow multiple data systems, within and among various programs and agencies, to share data and information.

The IRMA system and Natural Resource Information Portal (http://nrinfo.nps.gov) are the initial steps towards “one-stop shopping” for data and information that has been requested by national park managers and planners. The design and functionality of the NRInfo portal is based on interviews and surveys with hundreds of users from a number of stakeholder groups (see below) who were asked “what does the integrated data system need to do to more effectively help you do your work”?

The initial work on developing an integrated SOA data system has been led by the NPS Inventory and Monitoring (I&M) Program (http://science.nature.nps.gov/im/). As part of the National Park Service's effort to "improve park management through greater reliance on scientific knowledge," a primary role of the I&M Program is to collect, organize, and make available natural resource data and to contribute to the Service's institutional knowledge by facilitating the transformation of data into information through analysis, synthesis, and modeling. Two of the goals of the I&M Program, namely “Integrate natural resource inventory and monitoring information into National Park Service planning, management, and decision making,” and “Share National Park Service accomplishments and information with other natural resource organizations and form partnerships for attaining common goals and objectives,” can only be achieved through the development of a modern information management infrastructure (e.g., staffing, hardware, software) and procedures to ensure that relevant natural resource data collected by NPS staff, cooperators, researchers and

others are entered, quality-checked, analyzed, reported, archived, documented, cataloged, and made available to others for management decision-making, research, and education.

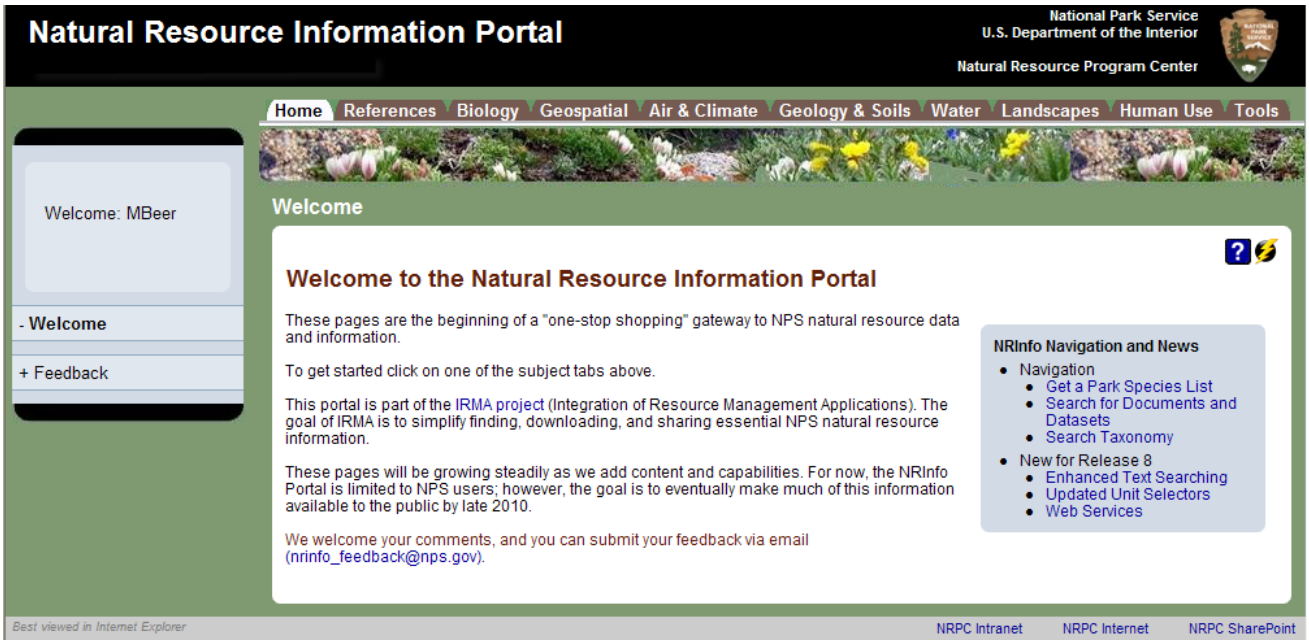

Figure 1. Home page of the Natural Resource Information Portal (http://nrinfo.nps.gov) as of June, 2010. The portal is organized by broad topic areas, such as References (metadata), Biology, Geospatial, Air & Climate, and Water. Implicit in the portal is automated authentication (i.e., no login or password requirements) and common navigation patterns regardless of the data being accessed.

The NPS Natural Resource Program Center (NRPC) is transitioning to SOA for two primary reasons: (1) to better serve users; and (2) to improve management of its current and future IT investments. Service-oriented Architecture (SOA) holds the promise of better cost sharing, rapid adaptability to changing mission needs, and improved interoperability of automated data processing systems. With ever-increasing amounts of data and continually-changing user and technical requirements, adopting an information management framework with maximum flexibility makes sense. SOA puts us in the best position to meet our users' needs over the long term, and to benefit from sharing information both internally and externally.

During 2004-2006, the NPS funded several studies, convened various workgroups, and conducted hundreds of interviews with NPS staff and collaborators to determine user needs and core requirements for modernizing natural resource data systems so that the NPS can more efficiently and effectively meet its mission. In October 2006, the new Director of the NPS Natural Resource Program Center (NRPC), George Dickison, issued a policy directive that NRPC will create an integrated data system with a central web portal, single sign-on system, and common user interface for all natural resource applications. The I&M Program has taken the lead to develop IRMA, which will eventually integrate all natural resource applications (starting with NatureBib, NPSpecies, and the NPS Data Store), eliminate redundant data storage, and streamline standard functions across applications (e.g., data entry, editing, searching and data retrieval).

1.1 Service-oriented Architecture: The DOI and Industry Standard

The Department of Interior has identified service-oriented architecture as a standard and "best practice" that will allow data exchange and integration among different data systems within and external to DOI agencies (see Figures 2-4 below). Service-oriented architecture is widely used in the online banking, travel, and shopping industries, and SOA is integral to the DOI's conceptual and enterprise architecture principles (see http://www.doi.gov/ocio/architecture/). A basic premise of SOA is that components are structured into concise, reusable, and sharable "services." These services then become flexible data building blocks that, using standardized tools, can be assembled or shared in a variety of ways depending on the information need. Advantages of this modern architecture as determined by DOI include the following:

- SOA maximizes information system investments via flexible, reusable services
- SOA increases the sharing of systems and information across agencies and organizations
- SOA reduces system risk due to reusable and shared services that have been designed for interoperability

An illustration of SOA-type information-sharing capabilities is the development of third-party travel websites such as Orbitz or Travelocity. When booking an airline flight, hotel, and rental car from one of these websites, data are being exchanged among systems developed by different companies. An airline, hotel, or car rental company is able to expose portions of its data to external users (such as Travelocity) for the purpose of data assembling and sharing. A single, centralized, data system is not necessary; rather, a common set of tools, standards, and governance procedures (inherent in SOA) makes it possible to share data among multiple, disparate systems. Figure 2 provides a simplified illustration of this concept.

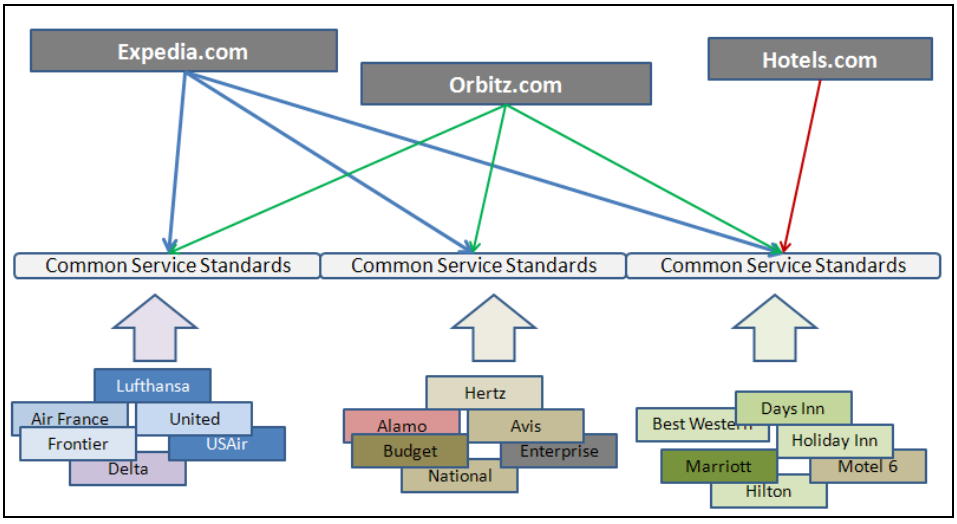

Figure 2. Use of a common set of tools, standards, and governance procedures (inherent in SOA) makes it possible for private companies or State and Federal agencies to share data among multiple, disparate systems. It is not necessary to have a single, centralized system, nor is it necessary for everyone to get together in advance and agree on common data structures or to design or develop the various components of the system all at once.

In the public sector we can put similar concepts to work (see Figure 3). Exposing portions of data systems or information sources as services and establishing common messaging specifications

among them can result in sharing information across agencies or bureaus. Access to shared information could be through agency-specific or topic-specific interfaces (e.g., NPS NRInfo Portal, a portal focused specifically on climate change, or existing portals such as Data.gov).

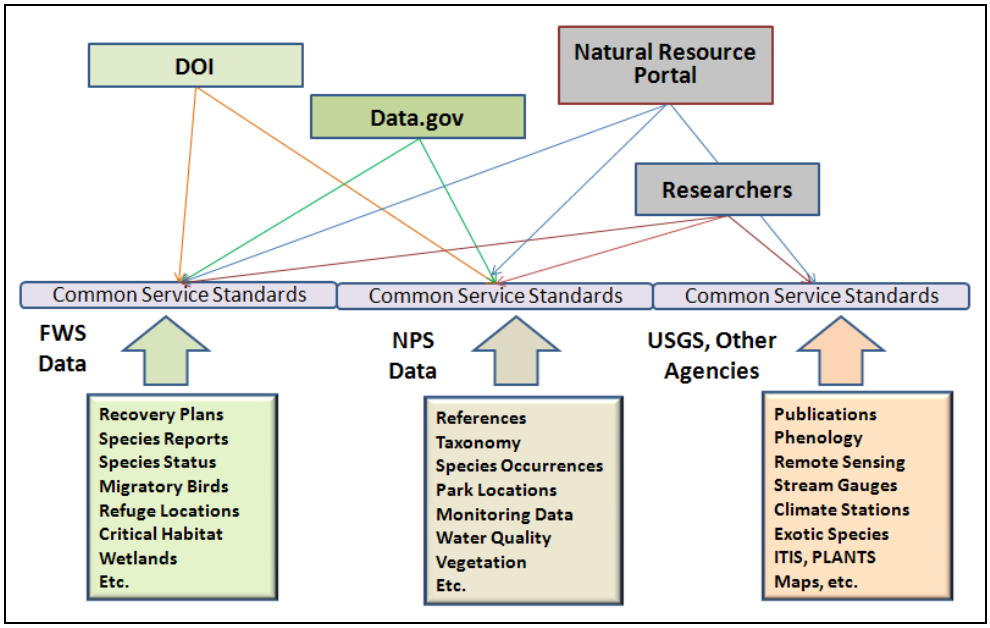

Figure 3. SOA provides the framework for sharing information among different government agencies and bureaus by using a common data service exchange standards.

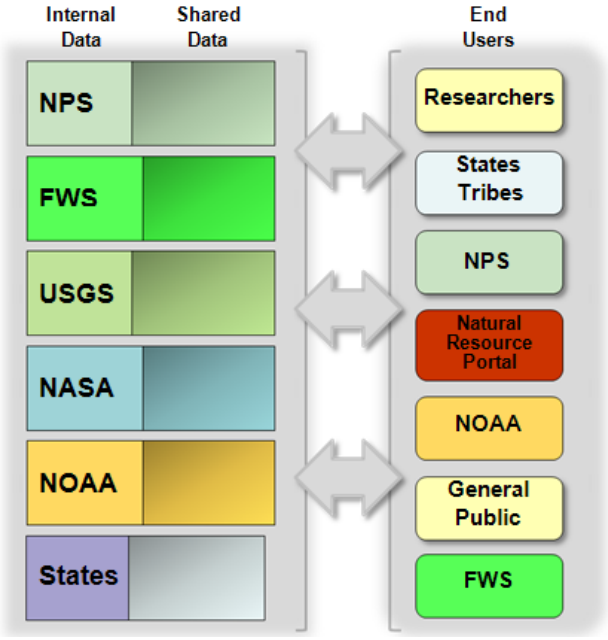

Figure 4. Conceptual diagram illustrating how service-oriented architecture, using industry and DOI standards, allows agencies to maintain and protect their internal data sets, yet make portions of their data available for sharing within and external to the DOI. A few examples of agencies within and external to the DOI are shown. SOA structures components into concise, reusable and sharable “services” that become building blocks for applications. The technology is widely used in the travel, banking, and other industries, and several DOI bureaus have already begun working on components of their data systems.

2. PROJECT HISTORY

The National Park Service, like many federal agencies, faces multiple challenges related to natural resource data collection, storage, dissemination, synthesis, and use for scientific and management purposes and by oversight entities. The natural resource information collected and managed by parks is voluminous, distributed, and complex.

In the 1990s, these challenges were addressed by developing separate computer applications and data structures to house different types of data and information. Three NPS web-based systems developed by the I&M Program serve as good examples:

- NatureBib was created to store bibliographic information such as research publications, technical reports, and a variety of memos and other documents.
- NPSpecies was created to manage information on the occurrence and status of plant and animal species in each park, along with associated attributes and evidence records.
- The NPS Data Store was developed to store primarily geospatial data sets and their associated metadata.

These systems, built with the best technology at the time, captured hundreds of thousands of records and provided tools for searching for and storing information. They were, however, "stovepipes": each had its own login, navigation, and search logic; systems did not interact well; users had to enter data in multiple locations and search different systems to answer common questions. The information, when isolated and placed in separate systems, lost value because the relationships among the data became weaker or disappeared altogether.

After a thorough evaluation of the issues (see below), NRPC determined that a new information management system was indeed needed and it would have several basic requirements: a single sign-on to eliminate the numerous and changing passwords; a central web-based portal where users could access multiple applications; and streamlined workflow for entering, editing, searching, sharing, and retrieving data. A service-oriented architecture (SOA) framework was selected as having the potential to solve the most difficult problems that had been identified.

2.1 Recommendations from the PRIDE Project and Systems Board

In 2004, the I&M Program funded a "user needs analysis" to evaluate data and information needs by parks and other levels of the NPS organization, and to determine whether the I&M goal of integrating natural resource data and information into park planning and decision-making was being met. This "user needs analysis" project was known as PRIDE (Protecting Resources through Informed Decision-Making and Education). PRIDE incorporated the concepts of enterprise architecture and used the DOI Enterprise Architecture "Methodology for Business Transformation" model to ensure that business functions would, indeed, drive information system development. PRIDE was supported by the Office of the Chief Information Officer of both the NPS and the DOI.

As part of the PRIDE project, more than 60 interviews were conducted that encompassed NPS managers, planners, and other staff at the park, region, and national levels; representatives from I&M networks; research scientists; and staff from other federal and state agencies and partner organizations. The findings from these interviews were recorded and documented, and were used as the basis of a full SWOT (strength, weaknesses, opportunities, threats) analysis.

The PRIDE team determined that improvements were needed to make natural resource data and information more accessible, to increase their availability for analysis and synthesis, and to make data and information more relevant and usable for a wide variety of stakeholders. Natural resource data and information were not being successfully integrated into park and Service-wide planning, management, and decision-making.

An additional outcome of PRIDE was the clarification of three enabling goals of NPS natural resource programs:

- Understand the state of parks' natural resources in order to make better decisions
- Expand informed and supportive constituencies for natural resource programs
- Create a more adaptive, flexible, and responsive natural resource directorate that can understand and respond to drivers of changes (e.g., climate, disturbance, demographics, technology)

Complementary to PRIDE, in 2005 the NPS Inventory & Monitoring Advisory Council (IMAC) recommended that the I&M Program evaluate the NatureBib, NPSpecies and NPS Data Store applications from a user's perspective and determine what changes should be made to eliminate redundancies and to develop a streamlined process for users. A four-member "Systems Board" chaired by Steve Fancy was formed at NRPC to carry out the evaluation. In early 2006, the board contracted part of the evaluation to a specialist who conducted 90 interviews with system users over the course of five months.

The PRIDE project and the Systems Board both confirmed that systems were not meeting user needs, and recommended rigorous streamlining of systems to produce an optimal workflow for users. Specific recommendations included eliminating the "stovepipe" application model; a transition to SOA; and the hiring of a project architect/manager to lead the transition. To implement the recommendations made by the PRIDE project and the Systems Board, and to improve the accessibility, use, and transfer of natural resource information within NPS, NRPC Center Director George Dickison issued a new policy on NRPC Natural Resource information technology (IT) systems in October 2006. The policy directed NRPC to begin the transition to a more efficient and integrated approach to natural resource information management, beginning with NatureBib, NPSpecies, and the NPS Data Store, and to transition to an SOA framework. IRMA was officially started.

3. PROJECT IMPLEMENTATION

The IRMA project scope statement was developed from key points in the new NRPC policy; in particular:

- NRPC systems will transition to Service-oriented Architecture (SOA). This transition will strategically position NRPC to better deal with future changes in technology. NatureBib, NPSpecies, and NPS Data Store applications will begin planning for this transition immediately. All other NRPC applications will initiate planning for their transition.
- NRPC will create a central web portal, a single sign-on system, and develop a common user interface for all natural resource applications.
- NRPC will integrate all applications (starting with NatureBib, NPSpecies, and the NPS Data Store), eliminate redundant data storage, and streamline standard functions across applications (e.g., data entry, editing, searching and data retrieval) while maintaining the separate missions of each application.

- NRPC will hire a SOA Project Manager to help transition to SOA. Initially this will be a contracted position that will report to the NRPC Center Director.

3.1 IRMA Demonstration Project

IRMA project staff determined that an initial "data portal" demonstration project would answer fundamental questions that needed to be asked before embarking on this large-scale project; in particular: could our data, currently in independent systems, be effectively merged into a single user interface? When users had a chance to try this prototype, was it something they found useful? And, could IRMA staff successfully make the significant internal cultural change that now emphasized collaboration as opposed to stovepipe-based independence?

The IRMA demonstration project (IRMA Demo) focused on the three priority NRPC systems, NatureBib, NPSpecies, and the NPS Data Store, single sign-on, and an integrated user interface.

Basic functional requirements were informally gathered relating to:

- How users would access a data portal
- How users enter search terms
- How search terms are applied to the source databases
- How the results were presented
- How the results were downloaded

After four months of successful staff collaboration, the IRMA Demo was presented at the annual Inventory and Monitoring Program data management conference in February 2007. Several comprehensive feedback sessions accompanied the demo. The review gave users the opportunity to comment on the preliminary model, suggest changes or improvements, and confirmed to IRMA project staff that the demo was headed in the right direction to meeting user needs.

3.2 SOA Architect/Project Manager

In April 2007, NRPC contracted with Phase One Consulting to assist with the IRMA effort. Phase One hired Randy Leonard, a project architect/project manager who had extensive private-sector experience building software development organizations and large-scale service-oriented architectures. Randy also had strong project management abilities, technical and programming skills, and demonstrated past success developing and leading large and complex teams.

3.3 Staff Reorganization

The first steps towards a transition to service-oriented architecture required fundamental changes to our organizational structure, hardware, software, and overall software development practices. In the pre-IRMA environment, each information system had its own system manager and software developer. Developers were not shared across projects and were “jacks-of-all-trades,” responsible for database design, business logic, user interface, testing, and overall system administration. The reliance on one person presented significant system risk in the event of staff turnover, and did not allow in-depth knowledge in any particular area of application development.

In 2007 the entire software development group was reorganized into three primary groups: functional analysts, developers, and quality-assurance testers.

- Functional analysts were responsible for identifying application stakeholders, interviewing users, documenting user needs, and translating these needs into the specific functional requirements a system needed to meet.
- Developers focused on building the software applications and graphical user interfaces (GUIs) as determined by the functional analysts.
- Testers developed specific test cases and independently and objectively tested software iterations, reporting bugs or other problems needing correction.

User requirements were now the drivers of the development process. Team members had responsibilities in their areas of concentration and worked across all applications instead of just one. Staff were able to broaden their skills within their areas of expertise. Collaboration and teamwork were essential and becoming the norm.

3.4 User-driven development

Before initiation of the IRMA Project, software development was typically overseen and directed by one or more resource specialists who determined overall system functionality. The approach was one of "if we build it, they will use it," which meant systems did not take into account user needs and often resulted in overall user dissatisfaction.

To ensure that redesigned systems delivered a useful end-product, user requirements needed to be solicited, understood, and integrated at every stage of development. The first step of user-driven development was to identify key stakeholders, which would include end users and any others who would be affected by or had influence on the development of an integrated natural resource information system. Input from stakeholders in each of the following groups was obtained through conference calls and interviews:

- Park natural resource specialists
- Park planners
- Park superintendents
- Park interpreters and public information officers
- Researchers and academicians
- Inventory and Monitoring Network staff
- Regional and WASO natural resource staff
- Fire Management staff
- Research Learning Centers
- Other federal agencies
- Partner organizations
- State natural resource agencies

3.5 Technical Standardization

In the pre-IRMA environment, information systems were based on a variety of software languages and tools. Technical skill sets ranged from MS Access, VB scripting, and Cold Fusion, to J2EE or .NET platforms, with both Oracle and SQL Server licenses being required. No consistent, standardized programming language or practices were in place.

An analysis was conducted evaluating major platforms and programming languages and the decision was made to standardize on the .NET development environment, primarily due to a clear cost advantage and the federal commitment to Microsoft products. Developers were then trained in using the new software tools and environment. Hardware resources were also evaluated regarding speed and security. The decision was made to move to a blade server configuration, with distinct environments established for development, QA, pre-production, integration, and production.

Essential team collaboration tools were also acquired and integrated into work processes. Microsoft SharePoint was established as the primary repository and work site for all IRMA-related project products, including documents, schedules, project plans, and announcements. Other collaboration tools such as Project Server, Project Web Access, Team Foundation Server, and HP Quality Center were evaluated and acquired.

Each distinct group within the IRMA team works with specific software tools. Functional analysts use Visio for documenting workflows and database table structures. Developers work in the .NET Visual Studio environment using C# as the programming language. QA testers use HP/Mercury Service Test as their key testing tool, and HP/Mercury Quality Center with Test Management as their communications tool. SQL Server 2005 or 2008 is used for the development and production databases.

As IRMA has progressed, a series of extensions and plug-ins have been added to enhance either development or presentation, including Ajax.net, ExtJS (a javascript library) and XML Spy. The GUI developers create the user interface using Visual Studio, ASP.Net pages, javascript and Model-View-Controller (MVC). The web server is IIS 6.0 using the ASP.NET 2.0 ISAPI filter.

Geospatial functions use SQL Server 2008 spatial data types and web mapping services for data storage and GeoJSON and OGC standard well-known text (WKT) for data interchange. GUI clients are based on OpenLayers and ExtJS extensions. These GUI components are integrated into the overall MVC architecture of the NRInfo Portal.

All development is now done in a 3-tiered environment:

- The presentation layer is the user interface: it provides the information display for the user. There is no logic or functionality embedded within a display page; it is simply for viewing the information presented by the business logic layer.
- The business logic layer contains the rules and calculations of a service and controls its functionality, separate from both its presentation and from access to the database.
- The database layer consists of the servers and databases where information is stored and retrieved. Data are independent from application and business logic layers, which improves security and performance.

3.6 Deconstruction

The IRMA team initially thought that a transition to SOA could be accomplished by building web services on top of existing databases, and modifying applications but maintaining the existing data structures. As work progressed it became evident that the existing structures were unwieldy and inefficient. Full redesign, while time-consuming and more difficult, would result in a data structure that was far more modular, efficient (faster response times), and able to serve us well into the future.

Each of the three primary systems (NatureBib, NPSpecies, and the NPS Data Store) was deconstructed, with both distinct and shared functions identified. Shared functions were those used by multiple applications; for example, establishing user identity, or selecting a park unit. Distinct functions were specific to one application only (Figure 5).

NatureBib	
Units	User Identity
Attach files	Taxonomy
Metadata	Audit
Project	Notification

NPSpecies	
Units	User Identity
Attach files	Match lists
Metadata	Observations
Taxonomy	Vouchers
Status	Notification

NPS Data Store	
Units	User Identity
Attach files	Taxonomy
Metadata	Audit
Project	Notification

Figure 5. Examples of shared functions (shaded cells) and distinct functions (white cells) within three applications.

Once these core functions were identified, basic or "primitive" services could be built to deliver them. Primitive services provide a very focused function, but can be combined in a number of ways as "composed" services to either reconstruct legacy applications or to rapidly build new applications. For example, the Unit Service, which contains information on designated NPS geographic and administrative units (e.g., parks, monuments, I&M networks, regional offices) is a primitive service that has been built once but that can be used many times by various applications that need it. The primitive services can also be handed over to others in DOI or elsewhere, to directly communicate with services hosted in the NRPC environment or to be installed locally for other uses.

A brief summary of the initial primitive services that have been identified as part of the IRMA project are as follows:

- Metadata Service: Information about documents, data sets, articles, images; links to full-text documents and files; will allow searching all content, including PDFs. This is the primary service that is called when a user makes a request to "search for data or documents."
- Identity Management Service: Eliminates user logins and passwords. System determines who you are and your specific roles and permissions for data sets and subsets (e.g., who is allowed to edit or delete certain records, upload data, or view sensitive data for a park or group of parks).
- Unit Service: Information on designated NPS geographic or administrative areas such as parks, offices, networks, or regions. Includes ability to link and cross-link units (e.g., one unit may comprise many units).
- Species Service: Park species lists and associated data on the occurrence and status of species in each park.
- Taxonomy Service: Provides crosswalks for scientific names and common names among multiple taxonomic systems, such as the Integrated Taxonomic Information System, USDA Plants data system, and various State taxonomic classification systems.
- Notification Service: Allows users to request and receive customized notifications (via email or other means) when specific events occur in an information system (e.g., "send me an email if a new species or report is added for my park").

- Feedback Service: Provides a mechanism for users to submit bugs or suggestions directly to information system managers; resolution of progress is visible. Engages users directly in improving information system and helps ensure system functionality meets their needs.

Data within these services that are appropriate for sharing can also be exposed to external "consumers," who can then integrate the information into their own systems or services. In this way, SOA provides the key to data integration and coordination among DOI bureaus, other agencies, and external partners. Information provided as a service can be shared and integrated by multiple end users, even if their respective systems are substantially different. Fundamental to this ability is XML (Extensible Markup Language), SOAP (Simple Object Access Protocol) messaging and REST (Representational State Transfer), all of which are web standards in widespread use.

3.7 Reconstruction: Composed Services

Once information has been deconstructed into focused, primitive services, it can be reassembled in a variety of ways to meet multiple purposes. Composed services replace the deconstructed silo applications, relying on the primitive services as components (Figure 6).

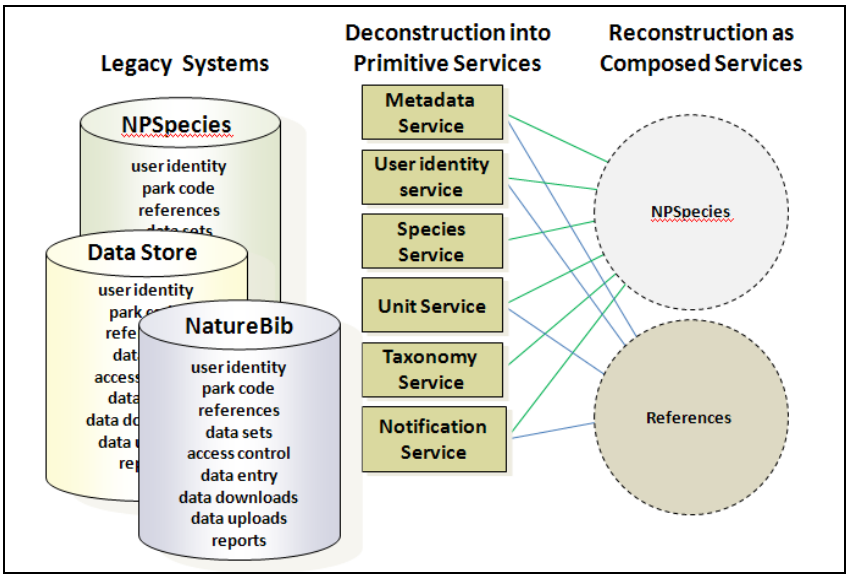

Figure 6. Illustration of the deconstruction of legacy system information into primitive services, which are then shared and recomposed to form new information systems.

Services can be composed in multiple ways for both internal use and external data sharing. For example, an information system for managing information on species observations could use many of the primitive services illustrated in Figure 6. Any primitive services developed specifically for observations may prove useful for other applications in the future.

In a similar way, external users, via SOAP or REST protocols (both internal and external to NPS) can make calls to IRMA's primitive or already-composed services, or assemble their own combination of services to meet their needs.

3.8 PROJECT STATUS as of June, 2010

NRPC has completed building the essential infrastructure of a modern, information management system (Figure 7 illustrates some of the components of this infrastructure), and information is being reliably delivered to NPS users through the NRInfo Portal. The June, 2010 release of the portal provides all NPS staff the ability to create and edit metadata records corresponding to documents, reports, data sets, and other information resources, and upload the associated data products. As a result, two legacy applications, the NPS Data Store and NatureBib, have been retired.

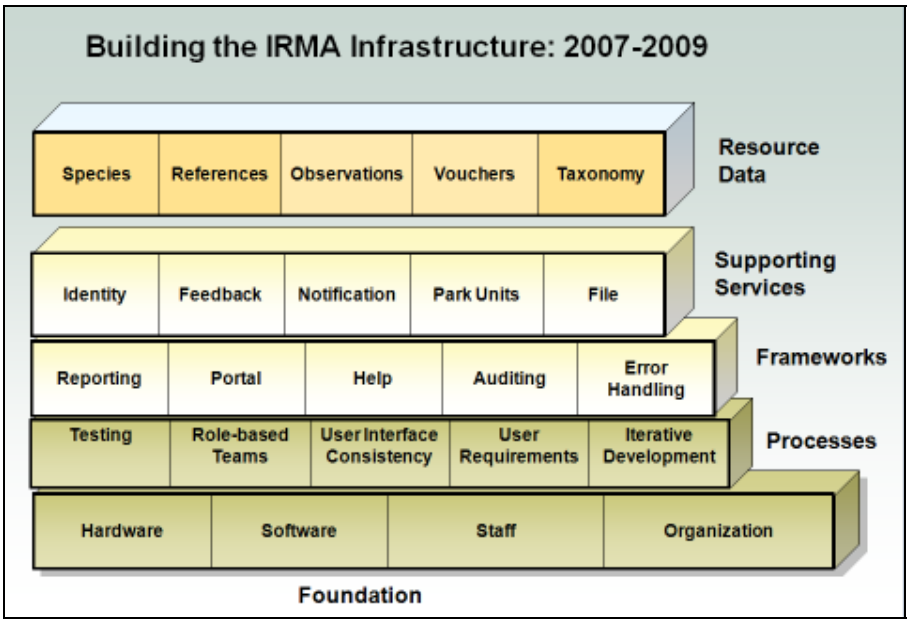

Figure 7. The initial stages of the IRMA project involved building a strong foundation and developing standard processes and frameworks that are used by all components of the integrated data system. Not all services and processes are shown.

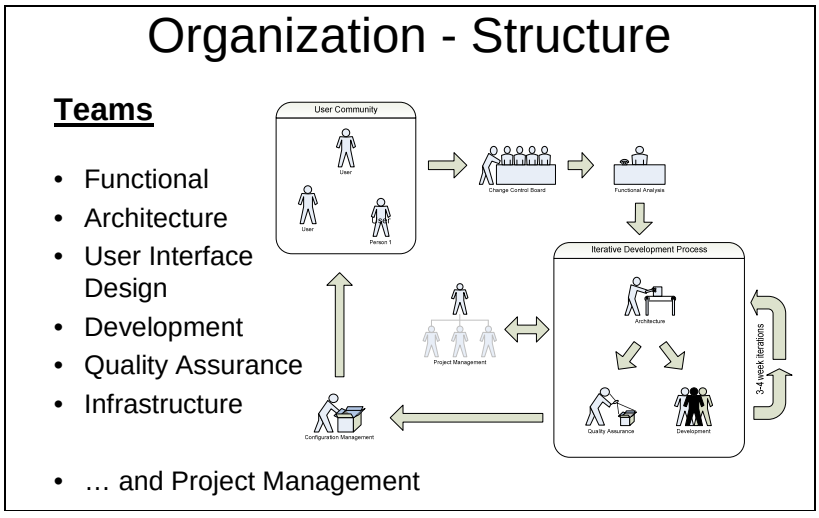

Figure 8. Fundamental changes to the NRPC organizational structure were made as part of the IRMA project. Federal staff and contractors have been organized into functional teams that work on all applications, as opposed to the old model of "one developer working on one application."

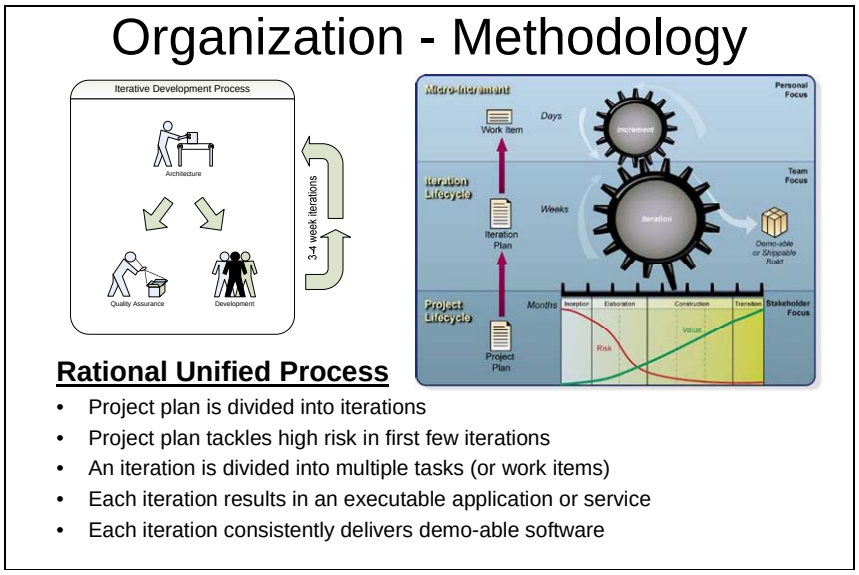

Figure 9. Software development follows the Rational Unified Process methodology, in which additional components and functionality are added with each version or iteration, and each iteration or "release" produces an executable application.

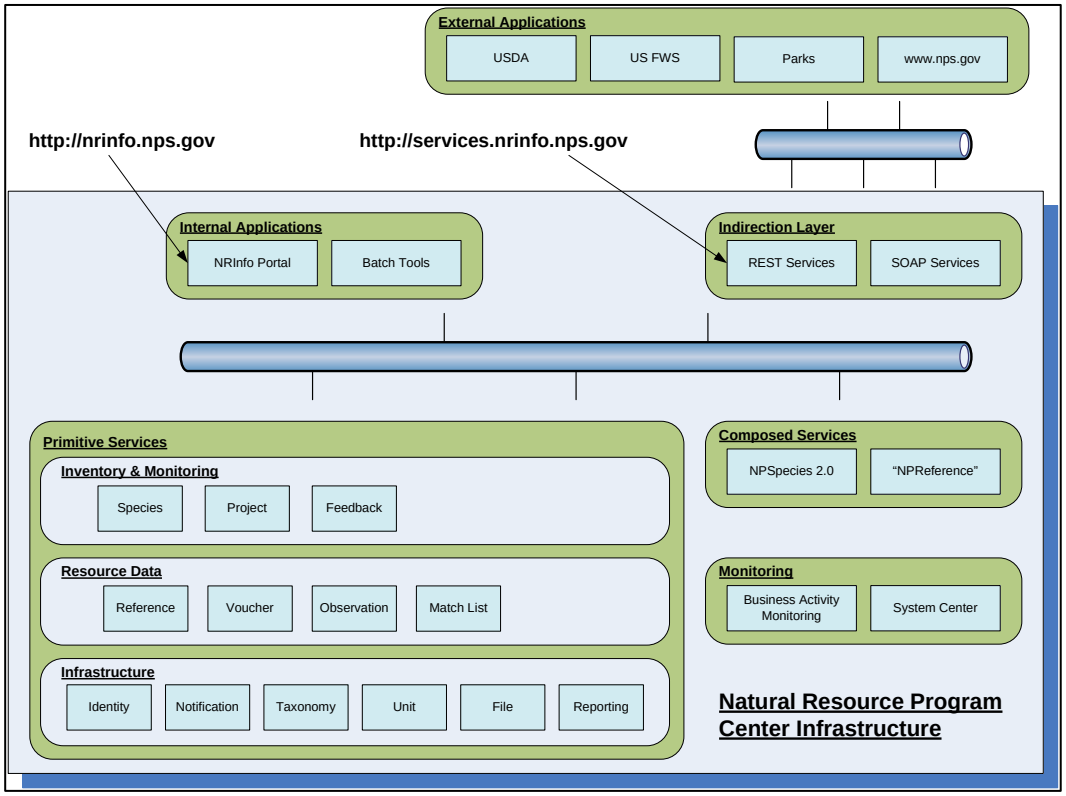

Figure 10. Overall architecture/topology of the integrated data system being developed through the IRMA project. NPS users can access data, information, and tools either through the NRInfo portal (http://nrinfo.nps.gov) or by calling web services available at http://services.nrinfo.nps.gov. Once the identity management service is fully implemented in late 2010, external users and partners will have similar capabilities.

4. Potential Integration with Other Programs and Agencies to Address High-Priority Issues

Rapid climate change is potentially the most far-reaching and consequential challenge to meeting the mission of the DOI and its bureaus. The potential loss of species and impacts to communities and ecosystems is without precedent. To meet the significant challenges posed by rapid climate change, the Department of Interior is promoting significant interagency collaboration including the establishment of Regional Climate Change Response Centers and a series of interagency Landscape Conservation Cooperatives to provide scientific data and technical support to inform land management and planning. As part of the DOI Climate Change Task Force, an interagency workgroup addressing "Climate Change Impacts Data Collection, Integration, and Management" recently made the following recommendation to DOI leadership:

> *"To meet the significant challenges posed by rapid climate change, we need a significant DOI response to develop effective and efficient means of finding, retrieving, using, and sharing the best available data and information. This will require both a top-down and bottom-up approach to addressing Department-wide data management issues. Much of the data and information needed to address climate change challenges already exists from more than 1000 sources, most of it from outside of DOI. Modern information technologies, and in particular, Service-Oriented Architecture (SOA), are providing tools, procedures, and protocols that allow multiple data systems hosting differing data structures to share information. Therefore, it is not necessary (nor practical) for data to be collected and stored in a consistent data structure or format for it to be used in data analysis, synthesis, and modeling to address management issues.* ***The key to data sharing and integration is the development of data systems that use DOI and industry standards to allow data exchange across multiple data systems."***

The IRMA team is involved in several demonstration projects with other programs and agencies to develop a proof-of-concept that illustrates the potential for web-based data sharing among multiple systems, and demonstrate how SOA and web services can be used to seamlessly share and integrate data across agency data systems. Demonstration projects that are currently underway include the following:

- The U.S. Fish and Wildlife Service and the IRMA team have already demonstrated how species data from the existing USFWS ECOS database (Environmental Conservation Online System) could be retrieved by IRMA via web services and integrated into the NPS IRMA framework. The USFWS and NPS are actively discussing how to work together on modernizing their respective data systems.
- The U.S. Geological Survey and NPS IRMA team have successfully demonstrated how it is possible to discover and retrieve documents from two different agency data systems from a single request using web service technology.
- The IRMA team has been working with staff from the U.S. Dept. of Agriculture's Plants database to develop a web service that will allow plant taxonomy data to be integrated into the IRMA system.
- Preliminary data exchanges are taking place with the Integrated Taxonomic Information System (a multi-agency partnership) as a means of populating a taxonomy service that supports NPS species applications.
- The Greater Yellowstone Science and Learning Center partnership website

(http://www.greateryellowstonescience.org) plans to dynamically incorporate into their website content pulled directly from IRMA services; in particular, species lists and subject-specific bibliographies. Successes here could then be applied to other NPS and partner websites (e.g., park web pages), allowing seamless access to continually-updated information.

- The Resource Information Services Division (RISD) in the NPS Office of the Chief Information Officer is also collaborating with the IRMA team, with the goal of providing simple, integrated access to information in IRMA services as well as the NPS Focus and NPS Voyager systems.

Mechanisms are being built that will transform any of IRMA's internal changes to a stable set of specifications that partners can reliably design against. External partners will be able to access NPS data using SOAP, REST/XML, and REST/JSON protocols. REST (Representational State Transfer) style web services allow users to fetch resources directly by using a URL composed with specific patterns, with the results returned as XML, CSV, or excel files.

For example, the URL:
http://services.nrinfo.nps.gov/npspecies/species/list/Certified/YELL/vascular plant?format=excel
sent to IRMA would fetch a list of certified species list for Yellowstone National Park plants, including park status, formatted in Excel. Figure 11 illustrates an example of how this service might be used on a park website: embedding a REST-style link to NRInfo data would eliminate the need to manually code and maintain a species list. Similar capabilities exist for fetching publication lists or bibliographies.

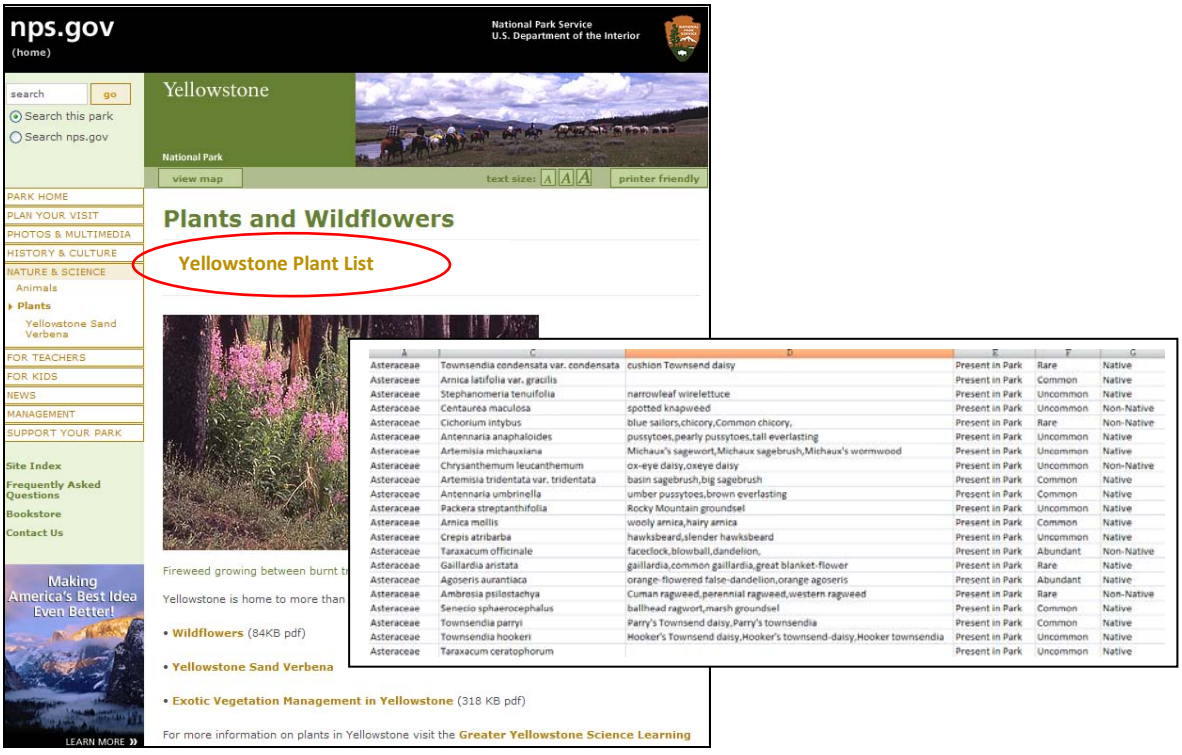

Figure 11. Data in the NRInfo Portal can be accessed in multiple ways. For example, a REST-style command could be embedded in a website link (circled in red), resulting in data retrieved "on the fly" from NRInfo Portal databases (inset). (Illustration purposes only; function not currently active on Yellowstone website)

An NRInfo "One-button Installer" or "Software Development Kit" has also been developed (Figure 12), which allows sharing of NRPC's SOA code with other partners. The installer components include basic and composed services, web plug-ins, sample data, and all of the necessary documentation.

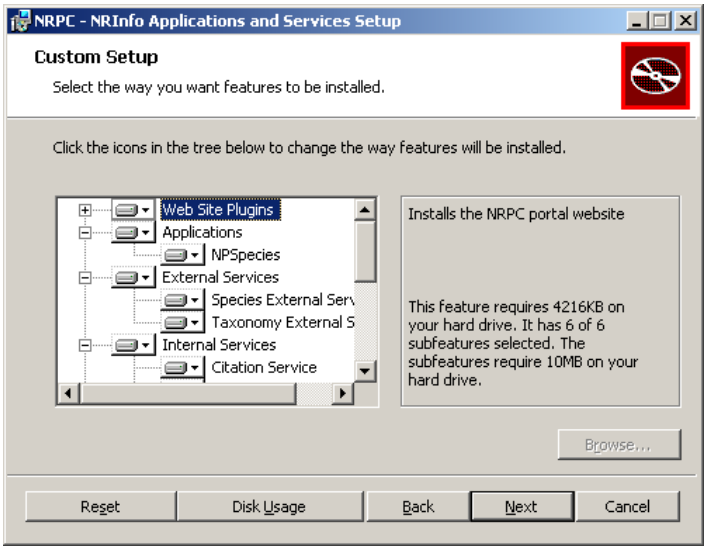

Figure 12. Screen capture of the One-Button Installer tool that will allow partners to access the NPS integrated data system.

5. IRMA ROADMAP – 2010 AND BEYOND

High-priority services and tasks for the second half of calendar year 2010 and into 2011 are:

Retirement of an additional legacy system, NPSpecies: Continued development of NPSpecies functionality in the portal, including the ability to create and update records, will allow the retirement of NPSpecies in October, 2010. (Two other legacy data systems, the NPS Data Store and NatureBib, were retired in June, 2010.)

Public access to non-sensitive data and information
Because of the high demand for information contained within the NRInfo Portal, the IRMA team is accelerating development that will allow partners and the public to access non-sensitive portal data before the end of calendar year 2010. This requires a two-pronged approach: ensuring that park staff have appropriately reviewed records that will be publicly accessible; and ensuring that the software and hardware controlling portal access is reliable and information resources are secure (see next item, below).

Identity Management Service: The portal software will automatically recognize users and grant fine-grained, pre-determined permissions (e.g., who is allowed to edit or delete certain data, create records, view sensitive data for a particular park or group of parks). This service must be robust, efficient to administer, and aligned with NPS and DOI security regulations and requirements. Identity Management must also allow authentication of outside cooperators and partners, as well as recognition of general users with no specific permissions.

Search Service: Many records in the Reference Application have associated and downloadable documents or data sets. A comprehensive search service will provide an efficient means of indexing all content (including documents, PDFs, zip files, data sets) for fast searching and retrieval. Also implicit in a search service is the development of "relevance" tools, so that those resources most closely matching search criteria are presented first in search results.

Spatial visualization tools: Spatial search and display capabilities will be expanded and refined in each portal release. By the end of 2010, the portal will provide visualization tools related to landscape dynamics monitoring (e.g., land cover, roads, population, conservation status). A subsequent priority in 2011 will be interactive access to and visualization of NPS natural resource inventories such as soils, geology, and vegetation.

Inclusion of other natural resource-related applications: Initial analysis of other natural resource-related applications will be evaluated in the second half of 2010 for inclusion in the NRInfo Portal. Key candidates are the NPS Research Permitting and Reporting System, and Park Visitor Use Statistics.

Extension of the NRInfo Portal framework: Further development of the portal framework will allow users to customize their portal pages, set preferences, save queries, and receive automated updates. This development will take place in 2011, and will be driven by feedback and priorities established by the user community.

6. KEY ELEMENTS FOR THE EARLY SUCCESSES

The IRMA project is a sophisticated undertaking that has created, within NRPC and NPS, a professional software development organization, solid business processes, and services to support these processes. The initial investment by the NRPC is expected to provide a solid return in terms of continually improving staff efficiencies, ensuring the longevity and accessibility of vital data, and supporting better-informed and substantiated park management decisions.

Several elements have been essential to the IRMA project's forward progress:

- First and foremost has been strong support from management, including the OCIO, the NRSS directorate, the NRPC Center Director, and the I&M Program Manager. Undertakings such as this are not without risk. Steadfast support and recognition of the project as a priority need to be maintained through project difficulties as well as successes.
- The commitment from management was backed up by adequate project funding.
- The project benefited from unequivocal user consensus that change was needed, and a clear direction on what users required and wanted. The IRMA Demo prototype further reinforced this consensus and gave team members focus and confidence.
- In a turnaround project such as this, cultural change can be as great a challenge as technical change. Developing a team structure in which all members have a chance to improve their skills, be in the forefront of new technology, and have a stake in project success had the effect of revitalizing staff members. Cultural change was less of an obstacle than initially envisioned.
- Outside expertise was required to launch and lead the project. Bringing in a project architect/manager from the private sector provided a fresh perspective, cutting-edge skills, and an aggressive, results-oriented focus. These traits were key to the project's success.

7. INFORMATION RESOURCES

7.1 Publicly-accessible links

- IRMA fact sheet
  http://science.nature.nps.gov/im/datamgmt/docs/IRMA_ProjectBrief_v1.2.pdf
- Portal fact sheet
  http://science.nature.nps.gov/im/datamgmt/docs/NRInfo_ProjectBrief_v1.4.pdf
- IRMA website
  http://science.nature.nps.gov/im/datamgmt/IRMA.cfm

7.2 NPS internal links

- IRMA Intranet Site: includes links to 2006 NRPC policy on SOA, additional links and resources
  http://www1.nrintra.nps.gov/nrpc_soa/index.cfm
- Natural Resource Information Portal
  http://nrinfo.nps.gov
- IRMA SharePoint Site: general project information, including presentations, executive summaries, etc.
  http://nrpcsharepoint/irma/default.aspx

Reference Application Executive Summary

Introduction

The National Park Service generates and uses a wide range and large volume of documents, reports, publications, data sets, and other valuable information related to park resources. Park managers, staff, researchers, and cooperators rely on this information in order to understand and effectively manage parks. All parks need a means of organizing and documenting this information in a way that it can be found and retrieved quickly and easily, regardless of its location or format.

As a component of the Natural Resource Information Portal, the Reference Application is developing the tools and procedures to allow parks to manage the information they produce and use. This application integrates two legacy NPS information systems: NatureBib; and the NPS Data Store; and also expands their capabilities.

The application provides essential information about documents, articles, datasets, images, and other information resources, including a description of their content, when they were created, who created them, where they are located, and the means to view or download them, if possible.

Current Features

Search Both Documents and Datasets

The Reference Application now searches simultaneously for documents and datasets, as opposed to requiring users to log in to two different systems. Users can search for information based on NPS unit, keywords, or geographically using an interactive map tool, and results can be downloaded to the user's local computer.

Create, Upload, and Edit Records

Any NPS employee can create and upload records, and edit records for which they have permission. In addition, users can now cross-link affiliated references (for example, link a report with an associated data set).

Download Attachments

Many references have an associated attachment (for example, a PDF of a document, or a GIS data set). Users can easily view or download the attachments, even very large files, for local use.

No Login or Password Needed

All NPS users can now search the Reference Application from any NPS computer or via VPN. No logins or passwords need to be requested or remembered. Sensitive records are, for now, protected and not included in search results.

What's Planned

The number-one priority now is to develop tools that allow partners and the public access to non-sensitive information resources. This means ensuring that data are appropriately flagged as sensitive (or draft), and that system tools that establish user identity and associated permissions are dependable, trusted, and error-free.

Additional planned features include incorporating other data sets and sources within the Reference Application. For example, the USGS Publications Warehouse could be searched and records retrieved, all seamlessly from within the NRInfo Portal. Conversely, we want to be able to share appropriate data with our partners. .

Conclusion

Information is the common currency among the activities and staff involved in managing NPS natural resources. The goal of the Reference Application is to provide an easy-to-use tool for cataloguing, finding, and retrieving the information that is essential for making sound management decisions.

NPSpecies Application Executive Summary

Introduction

Knowledge of the species that occur within parks is fundamental to good park management, and is of high interest to cooperators, researchers, and the general public. The NPSpecies Application documents the occurrence and status of species in more than 270 national park units that have significant natural resources.

NPSpecies, working in conjunction with other services such as units and taxonomy, provides fundamental information on species and their park status such as (e.g., present, probably present, historic), abundance, residency, and nativity.

Data available through NPSpecies are posted on the Natural Resource Information Portal and are available through web services. Data are accessible to all National Park Service staff; no login is required.

The original NPSpecies application is still the primary means for users to enter or edit park species data. Migrations from the original application to the portal occur monthly.

Current Features

Certified Park Species List

NPSpecies allows users to view and download park-species lists for vertebrates and vascular plants that have been reviewed by subject-matter experts. These species lists contain species scientific and common names with park status.

Full Park Species List

The full park species list contains scientific names for any taxonomic category (i.e. invertebrates) that have been documented to occur in a park. This list includes names at taxonomic levels other than species (e.g., subspecies, genus, family). Data are preliminary and may not have been reviewed by subject-matter experts.

Park-Species Profile

These profile pages associated with park species provide more detailed taxonomic and park status information.

Web Services

Web services are available to allow users to fetch resources (i.e. species list) directly by using a URL composed with specific patterns.

Match List Service Integration

Full Park Species Lists are now matched to up species with designations maintained by other agencies or organizations, e.g., U.S. Fish and Wildlife Service T&E data, ozone sensitive plants, NatureServe global conservation status ranks, and state species of concern.

What's Planned

Upcoming releases of NPSpecies will include observation, voucher, and reference data affiliated with park species records, which will allow users to view evidence associated with park species records.

Tools will also be developed for controlling access to sensitive records as well as reviewing records for quality.

Conclusion

The goal of NPSpecies is to provide easy access to reliable and up-to-date park species information. This and related applications are expected to be fully developed and deployed in late 2010, at which time the original NPSpecies application will be retired.

Unit Service
Executive Summary

Introduction

The National Park Service comprises a diverse collection of parks, monuments, offices, programs, networks and other functional entities. The term "unit" refers to any one of these entities.

Some units represent operational or programmatic groupings, such as vital signs networks, regions or clusters. There are also unit collections that must be maintained. An example collection is those parks with significant natural resources that are included in the Inventory & Monitoring Program.

Almost all natural resource information gathered and managed by NPS is affiliated with one or more units. The Unit Service is being developed to maintain current and legacy information for all NPS units that are pertinent to natural resources. While all NRInfo Portal applications and services will eventually adopt this service to ensure data consistency, the Unit Service is not the official repository for NPS park or organizational codes.

The Unit Service groups all units of similar coarse-grained function by assigning a unit "type." Park, directorate, office, operation, and program are examples of unit types.

A unit's purpose is further refined by assignment of a unit "subtype. " Legislated park, affiliated area, and pending legislation are examples of park subtypes. Operation subtypes include region, I&M network, and cluster.

Relationships among various units are managed using " link types." A link's assigned type defines its purpose, and therefore the types of units with which it may be associated. For instance, a park may be linked to a region via the 'assigned to operation' link type.

Links may be used in conjunction with unit types and subtypes to ask either generalized or specific questions. A generalized question might be, "show me all parks in my network," whereas a specific question would be, "show me how parks in my network have changed."

Current Features

Unit Service currently serves as the source of unit information for NPSpecies and the Reference Service on the NRInfo Portal (http://nrinfo.nps.gov). Unit records are actively maintained and managed so the service returns up-to-date, accurate information.

Users of the NRInfo Portal often need to create or download a master list of units or certain subsets that they can use locally. A utility is available ("unit search" on the Tools tab) to meet this need, including searching, reporting, and download features.

What's Planned

Geographic representations of units will be added soon. These geographic locators will be used to refine searches based on Units. The Unit Service will also allow the creation of "sub-units." These are distinct geographic areas within a unit; for example, one island among several, or individual non-contiguous areas that comprise a unit. Users will be also able to create their own sub-units, for the purpose of better-managing or grouping affiliated data.

Conclusion

The Unit Service streamlines natural resource information management by providing a single, flexible, centrally-managed source of unit codes and associated attributes, thereby ensuring data consistency and reliability.

June, 2010

Taxonomy Service
Executive Summary

Introduction

Taxonomy is the combined practice of science and classification. There are numerous biological classification sources and systems used by federal and state agencies, organizations, and individuals, and one organism can often be defined by many distinct names, both scientific and common. These multiple systems and names can make it difficult to compare or evaluate information from different geographic regions, agencies, or persons.

A goal of the Taxonomy Service is to provide a flexible structure with fluency in multiple taxonomies. This allows for name translations and crosswalks among key taxonomic systems, and allows users to assess information combined from different systems regardless of the names used.

Taxonomic names may be provided or recognized by one or more classification sources. Well-known sources include official taxonomic systems such as the Integrated Taxonomic Information System (ITIS), the American Ornithological Union, or the U.S. Department of Agriculture PLANTS system. Other commonly- accepted classification sources include, for example, publications of state or regional flora such as Welsh (Utah) or Jepson (California).

Current Features

The Taxonomy Service currently provides basic read-only capabilities so that users can find, view, and download taxonomy records. Currently-available data sources for approximately 750,000 taxonomic records are ITIS and NPSpecies. Records include the names (scientific and common), codes, synonyms, crosswalks, and associated hierarchical information (e.g., order, family). The service presents a consistent interface for viewing and searching for all taxonomy records, regardless of the record source.

The Taxonomy Service is also the taxonomic information source for all records in the NPSpecies application on the NRInfo Portal.

What's Planned

Features planned in future releases include: links to additional external classification sources (e.g., USDA PLANTS), scientific authorities, advanced searching, REST services that can be called by URL, and write and edit capabilities.

The Taxonomy Service is designed to accommodate taxonomies other than biological, such as vegetation classes, geology, or paleontology, should this information be integrated into the NRInfo Portal in the future.

Conclusion

Taxonomic information is complex, changing, and has myriad variations based on scientific discipline, geographic region, and personal preferences. The Taxonomy Service is designed to be flexible, accommodate individual preferences, and leverage the multiple information systems that are already in existence.