RecFIN Database MS-SQL Migration

FY 2015 Proposal

Ed Hibsch
Created: 05/13/2015

1. Overview

1.1. Sponsor
Edward Hibsch

1.2. Focus Group
Information Management

1.3. Background
PSMFC's Recreational Fishing Information Network (RecFIN) database was developed many years ago to process, store and report estimates and sample data for the three West coast states; California, Oregon and Washington. The current database stores all of this information in a network of SAS data files. The database has served RecFIN's needs quite well since its inception. Over the past year, RecFIN staff has engaged users to gather feedback about their user experiences accessing data from RecFIN. We have conducted a web-based RecFIN Users survey, elicited feedback from participants in two CRFS (California Recreational Fishing Survey) Training/Workshops, and also met with stock assessors and economists at the Northwest Fisheries Science Center. We compiled a great deal of detailed feedback that has led us to the conclusion that the RecFIN database is in need of an upgrade. The main theme that emerged was that the RecFIN database needs to be in a true relational database format.

1.4. Project Description
This project proposes to design and develop a new RecFIN Microsoft SQL database. When completed, the new database will replace the current SAS-based RecFIN database. RecFIN data from 1980 to the present will be stored in MS-SQL data tables and user access/reporting will be designed using MS-SQL Reporting Services (or compatible) query tools.

1.5. Public Description

1.6. Objectives
1. Simplified Data Access - The primary objective of this effort is to allow users to be able to access RecFIN sample data and estimates in a user-friendly, intuitive manner; eliminating the need to perform additional steps to get the data in a usable format that they can run analyses on. Data will be clearly documented so that users understand what each data element consists of.2. Enhanced Reporting – We strive to incorporate state-of-the-art reporting tools to give fishery managers the information they need concisely, accurately and reliably. Additionally, we hope to provide new modeling and simulation tools to empower users to gain additional insight into their data.3. Consistency – Our goal is to report the same data that each of the states are using for management, and to provide stock assessors with the data they need across all three states in an intuitive, predictable format. This will save them considerable time when conducting stock assessments.

1.7. References

2. Methodology

2.1. Methodology
1. Gather Information: We have already elicited a great deal of feedback from our users by the methods described in the Background section. Information from these efforts have been synthesized into a list of features that RecFIN users would like to see added or enhanced. Every day we continue to learn of things that users would like to see modified in order to make RecFIN data easier to use. As a result, the requirements document is an ever evolving work in progress.2. Chart the Course: Using existing funds, we have contracted with a technology firm to build a RecFIN SQL Migration "Roadmap". We will begin to meet with them regularly to answer their questions and provide detailed technical specifications for what are current state exists of and what we desire for the future state of RecFIN. The deliverable for this phase will be a detailed description of the steps in SQL migration, cost and timing for each step, and an assessment of risk and consequences of not completing steps successfully.3. Build the Infrastructure: The detailed process steps will be specified in the Roadmap document that our technology vendor delivers. The general stages of the build will be as follows:• Import the legacy RecFIN data from delimited files converted from RecFIN SAS data files. This stage will also include making the relational connections between data elements in MS-SQL• Redesign the RecFIN website to provide direct data links to the MS-SQL supporting database. This will most likely be done using Microsoft .NET technology as it appears to offer the most seamless transfer of query information to the SQL and return od tabular data for user view.• Design and develop a self-service reporting environment based on the recommendations from the Roadmap. This will likely be built in SQL Server Reporting Services, unless the analysis done to create the Roadmap yields a solution that will better meet RecFIN's needs. The reporting environment will be two-fold, one that is public-facing and one that is available to authenticated users who are granted access to specific data views. • Design and develop query tools that will enable users to easily download the sample data records that they desire with all of the data elements intact. Users will no longer be required to write additional data management code to put the data elements together into a format they can use.• Create new tools to enable each state to easily upload their data to RecFIN. The new approach will streamline this process by adopting tools

that will directly import data from other Microsoft products (Excel, Access) into MS-SQL.4. Analytic Extension: Once all of the data processes are in place, it is our hope that we can assist regular users of the data by incorporating some of the existing models they use for catch projection, trend analysis and forecasting. This will enable users to run these models on the most current data at any time, alleviating the need to download data and run them independently.

2.2. Region
Pacific

2.3. Geographic Coverage
Data from California, Oregon and Washington will be included.

2.4. Temporal Coverage
Data from MRFSS will be included from 1980-2003. From 2004 to the current date (and beyond), data wi

2.5. Frequency
NA

2.6. Unit of Analysis
NA

2.7. Collection Mode
NA

3. Communication

3.1. Internal Communication
Monthly updates will be provided in the MDMS. Detailed progress updates will be presented to the RecFIN Technical committee in their semiannual meetings and by email when appropriate.

3.2. External Communication
A Steering committee may be formed to engage frequent, experienced users to provide consultation and beta testing.

4. Assumptions/Constraints

4.1. New Data Collection
N

4.2. Is funding needed for this project?
Y

4.3. Funding Vehicle
RecFIN Grant

4.4. Data Resources
Existing data resources are all that are required. This includes:Marine Recreational Fishery Statistical Survey (MRFSS) will be included from 1980-2003. California Recreational Fishery Survey (CRFS): 2004 - currentOregon Recreational Boat Survey (ORBS): 2004 - currentWashington Ocean Sampling Project (OSP): 2004 – currentPotential New Data Sources:Ocean Salmon Project (Salmon data for all 3 states)Commercial Passenger Fishing Vessel (CPFV) Logbook

4.5. Other Resources

4.6. Regulations

4.7. Other

5. Final Deliverables

5.1. Additional Reports

5.2. New Data Set(s)
The new RecFIN database will be a Microsoft SQL Server database that resides on a virtual server at

5.3. New System(s)
The new RecFIN website will reside on a PSMFC web server.

6. Project Leadership

6.1. Project Leader and Members

First Name	Last Name	Title	Role	Organization	Email	Phone 1	Phone 2
Edward	Hibsch	RecFIN Programmer /Analyst	Team Leader	PSMFC	ehibsch@psmfc.org	503-595-3100	
Craig	Miller		Team Member	PSMFC	cmiller@psmfc.org	503595-3100	

7. Project Estimates

7.1. Project Schedule

Task #	Schedule Description	Prerequisite	Schedule Start Date	Schedule Finish Date	Milestone
1	Development Roadmap		11/10/2014	02/27/2015	
2	Create RFP for Migration programming	1	03/02/2015	03/13/2015	
10	Obtain and perform ETL on MRFSS era legacy data	9	04/23/2015	07/31/2015	
13	Redesign website to integrate access to SSRS reporting	8	08/03/2015	09/25/2015	
18	Develop an interface that will feed data to PacFIN to fill scorecard reports	14	10/05/2015	12/31/2015	
14	Develop sample data query tools	8	08/03/2015	09/25/2015	
16	Develop a Web Services interface for remote state data upload	14	10/05/2015	11/27/2015	

Task #	Schedule Description	Prerequisite	Schedule Start Date	Schedule Finish Date	Milestone
17	Develop enhanced analytics and integrate for use through the RecFIN web site	14	10/05/2015	03/31/2016	
3	RFP response period	2	03/16/2015	04/10/2015	
4	RFP Review and vendor selection	3	04/13/2015	05/15/2015	
5	Project kickoff/Initiation/ Project plan development	4	06/01/2015	06/12/2015	
6	Finalize data model	5	06/15/2015	06/26/2015	
9	Final decision on what historical data to maintain (State data vs. MRFSS data, combination)		04/21/2015	04/22/2015	
19	Reengage Beta user group to test and refine second wave of features	17	04/04/2016	05/27/2016	
7	Develop Extract/Transform/Load (ETL) programs and process for monthly state data		06/29/2015	07/31/2015	
8	Import post MRFSS legacy data (2004-2014)	6	06/29/2015	07/31/2015	
11	Develop Infrastructure: Security, User Management, Error Logs, Data Backups	6	06/29/2015	08/28/2015	
15	Engage Beta testers to explore the new database and reporting functionality	13	10/05/2015	11/27/2015	
12	Develop reporting framework in SSRS	8	08/03/2015	09/25/2015	

7.2. Cost Estimates

Cost Name	Cost Description	Cost Amount	Date Needed
PSMFC Overhead Costs		$2415.00	03/02/2015
Application programming	Programming contractor	$150000.00	03/02/2015
TOTAL COST		$152415.00	

8. Risk

8.1. Project Risk

Risk Description	Risk Impact	Risk Probability	Risk Mitigation Approach
Secured funding is inadequate to complete all of the features outlined in project plan	Some of the desired features will have to pushed into a new development project next year	Medium	Scrutinize the costs of each development phase and associated tasks to be sure we are not spending funds on features of secondary importance or that can be built with our own staff
SQL Server Reporting Services (SSRS) does not fully meet the needs of RecFIN users	Development would be held up until a third party reporting tool can be identified, evaluated and implemented. There may also be a cost implication associated with this change.	Medium	Build a reporting prototype solution early on in the process to be sure that SSRS will meet RecFIN's needs.
Development does not mesh with ongoing changes in each state's data management procedures.	Possible duplication of effort or development of system interfaces that are incompatible.	Medium	Frequent communication with state stakeholders. Engage the RecFIN Data and Technology subcommittee more often to keep each other apprised of state development efforts.

9. Supporting Documents