



To: Distribution

From: Stephen Smith, Technical Writer

Thru: Ernestine Bryant, STOrage and RETrieval (STORET), Project Manager (PjM),
Solutions Delivery Center (SDC)

Subject: Minutes of STORET Change Control Board (CCB) Meeting

1.0 Purpose

A CCB Meeting was held on April 7, 2005 at the SDC. The purpose of the meeting was to review and address the status of current project activities, resolve project issues, and ensure that activities are within the scope of the Job Assignment (JA).

2.0 Attendees

Bryant, Ernestine - Project Manager
Christian, Kevin - Job Assignment Manager
Manning, Lee
Munthali, Chimwemwe
Smith, Stephen
Tarquinio, Ellen
Thadkamalla, Ganesh

3.0 Discussion Topics

The discussion covered the following topics:

- C Action Items.
- C Taxonomy Update Utility.
- C Central Warehouse.

3.1 Introduction

Attendees opened the meeting with a summary of project activities and a status of open action items. The Software Incident Report (SIR) Log (Attachment A) and User Support Log (Attachment B) are included as deliverables with these CCB Minutes.

E. Tarquinio was introduced as the biologist and taxonomist that will be handling taxonomic issues in STORET.

3.2 Taxonomy Update Utility

The taxonomy update process was reviewed and key decision points are as follows.

When duplicate Taxon Serial Numbers (TSNs) are found in the Integrated Taxonomic Information System (ITIS) data, apply the following rules:

- C If one of the duplicate TSN taxons is Approved, incorporate the Approved record into STORET and ignore other synonyms.
- C If all the duplicate TSN taxons are Synonyms, incorporate the first fetched Synonym record into STORET and ignore other synonyms.
- C If there are multiple Approved duplicate TSN taxons, stop processing and ensure duplicate records are identified as such in a log file.

When duplicate Display Names are found in the ITIS data, perform the following:

- C Identify the records before incorporation into STORET.
- C Load the records into a table in the STORET schema providing columns for the TSN, display name as extracted from ITIS, and modified display name to be loaded into STORET as determined by reference table maintenance personnel. An unallocated SIR was created to provide a maintenance area in the Reference Table Application for editing the modified display names.

When TSNs or parent TSN values are missing in the ITIS data as compared to the data present in STORET at execution time, perform the following:

- C Leave the record in STORET.
- C Remove associated taxonomy information (e.g., Taxon Sort Code, Rank, Parent TSN).

- C Identify the record in a log and continue with processing.

The extraction and appending of ITIS comments/notes to display name values was discussed. It was realized that the content of the comments/notes has significant variation, and could change over time. The current rules extrapolated from the Practical Extraction and Report Language (PERL) and Statistical Analysis Systems (SAS) programs were deemed insufficient for determining the comment information for concatenation to the taxon display name. L. Manning will provide the rules and conditions for ITIS comment information extraction.

Non-American Standard Code for Information Interchange (ASCII) characters were present in the ITIS data. L. Manning will provide translation rules for these characters.

3.3 Central Warehouse

K. Christian brought up two items for consideration:

- C Modify the TSRUOM table for support of unit conversion and deliver the modified table with next software release. In conjunction, modify Central Warehouse software as necessary to accommodate the change.
- C Create an archival version of the Central Warehouse. This would help provide long-term control of the Central Warehouse size. Legacy data could also be considered for inclusion in this archive.

4.0 Action Item Summary

Action items and their status are as follows:

Number	Action Item	Assignment	Date Issued	Status	Date Completed
0001	Research VB.NET capabilities and provide response.	B. Norris	06/02/2004	Closed	09/08/2004
0002	Provide STORET team with copy of Dasler software.	B. King	06/02/2004	Closed	08/25/2004
0003	Provide BEACHES program with STORET architecture information (e.g., Entity Relationship Diagram, data definitions).	J. Wilson	06/28/2004	Closed	07/28/2004
0004	Run Extract, Transform, and Load (ETL) software week of July 12 th .	J. Wilson	06/28/2004	Closed	07/28/2004

Number	Action Item	Assignment	Date Issued	Status	Date Completed
0005	Provide copy of email discussing required changes to implement Uniform Resource Locator (URL) encoding for the Station Home Page.	B. Norris	06/14/2004	Closed	08/06/2004
0006	Send Report Module v2.0.3 deliverable.	S. Smith	07/28/2004	Closed	07/28/2004
0007	Include scripts to remove orphaned rows and nullify type_name column in TSRCHALS for next reference table upgrade.	L. Manning	07/28/2004	Closed	08/23/2004
0008	Investigate Integrated Taxonomic Information System (ITIS) data acquisition and supplementary data processing.	L. Manning	07/28/2004	Closed	01/12/2005
0009	Examine Native American Land (NAL) data for compliance to the Environmental Data Registry (EDR).	B. Norris	08/23/2004	Closed	09/08/2004
0010	Review structure of TSMNAICS table to see if it was expanded from its original source.	S. Smith	08/23/2004	Closed	09/08/2004
0011	Send Data Entry Module v2.0.1 deliverable.	S. Smith	09/08/2004	Closed	09/08/2004
0012	Prepare Entity Relationship Diagrams of the Central Warehouse and STORET Database for the National Conference.	B. Norris	10/8/2004	Closed	12/01/2004
0013	Prepare graphs showing results converted to target units for possible use at the National Conference.	S. Smith	10/08/2004	Closed	10/21/2004
0014	Email URL to access STORET Station Home Page and Extensible Markup Language (XML) stream to B. King.	B. Norris	12/01/2004	Closed	12/02/2004
0015	Update BLOBs on preformatted reports matrix for review at next CCB.	S. Smith	12/01/2004	Closed	12/16/2004
0016	Aggregate pertinent Reference Table SIRs for review and disposition.	S. Smith	01/12/2005	Closed	03/03/2005
0017	Schedule ITIS data load meeting.	E. Bryant	01/12/2005	Closed	02/16/2005
0018	Provide Practical Extraction and Report Language (PERL) and Statistical Analysis System (SAS) programs supporting ITIS data load to SDC.	L. Manning	01/12/2005	Closed	02/16/2005

Number	Action Item	Assignment	Date Issued	Status	Date Completed
0019	Add a comment to the ALL_TAB_COMMENTS view identifying the full name of each STORET and STORET1 table.	K. Christian	02/16/2005	Open	
0020	Provide L. Manning a report of duplicate Display Names and Taxon Serial Numbers found in ITIS data.	G. Thadkamalla	03/23/2005	Closed	04/07/2005
0021	Send duplicate Taxon Serial Numbers found in data during development to L. Manning when duplicates are identified as Approved.	G. Thadkamalla	04/07/2005	Open	
0022	Send duplicate display names to L. Manning when duplicates are within the same hierarchical branch.	G. Thadkamalla	04/07/2005	Open	
0023	Provide rules and conditions for ITIS comment information extraction.	L. Manning	04/07/2005	Open	

5.0 Next Meeting

The next meeting was scheduled for April 21, 2005 at the SDC.

6.0 Distribution

Name	Email
Bryant, Ernestine	
Christian, Kevin	Christian.Kevin@epamail.epa.gov
Manning, Lee	Manning.Lee@epamail.epa.gov
McElhinney, Cary	Mcelhinney.Cary@epa.gov
Szajgin, Tracey	
Wilson, Joseph	Wilson.Joe@epamail.epa.gov

7.0 Approval of Minutes

Kevin Christian
Job Assignment Manager

Date

ATTACHMENT A

SIRs Since Last CCB

Software Incident Reports Since Last CCB - Period: 03/23/2005 - 04/06/2005			
System Version	Date Initiated	SIR #	Description
C2.0.5	03 31 2005	1690	Update unit conversion scripts and associated spreadsheet to incorporate reference table changes through dbfix019 (released March 4, 2005). 04/05/2005 - Delivered to EPA as interim release of C2.0.5 of script unit_conversion.sql only.
R	04 06 2005	1691	Allow selected Projects to be interrogated against Project Activity assignments.

Legend for System Version:

System Letter is concatenated with Version Number. System Letter Key:

C = Central Warehouse, R = Report Module

ATTACHMENT B

User Support Log

User Support Log - Period: 03/23/2005 - 04/06/2005			
Date Received	System Version	Issue	Response
3/23/05	P2.0	One of the training machines (XP) is not opening up the Document/Graphic window.	Provided diagnostic assistance as follows: - Ensure "Oracle Provider for OLEDB" is loaded. - Ensure Adobe Acrobat is loaded. - Try logging into machine as Administrator. Change privileges to give Write and Modify. - Try logging into machine as User. Change privileges to give Write and Modify. Assign user to Group. The problem is related to Windows XP security. Not able to reproduce at SDC.

User Support Log - Period: 03/23/2005 - 04/06/2005			
Date Received	System Version	Issue	Response
3/24/05	P2.0	<p>User wanted information regarding the correct representation of Replicate activities and Sample from Sample activities. Note: user question was not entirely clear.</p> <p>Email: Using parent/child relationship model for data entry--</p> <p>Sample A Parent Sample Activity Category = Routine Sample Child Activity Category = Created from Sample</p> <p>Sample B (Duplicate) Parent Sample Activity Category = Field Replicate/Duplicate (and Replicate Number =1) Child Activity Category = Created from Sample</p> <p>This appears to be a solution to my earlier question. Will I have any problems with this setup?</p>	<p>Not exactly sure what the "earlier question" was, but here is some feedback on Question 1.</p> <p>The scenarios depicted in Sample A and B can be entered into the Data Entry Module (DEM); however...</p> <p>A single Activity replicate would seem incomplete. If someone wants to perform Replicate activities (Field Duplicate or Depletion), I believe the proper way to represent this is to create both samples as Replicates, assigned replicate numbers of 1 and 2 (or more) to each respectively, and use the same Activity ID for each replicate. This ties the samples together via the Activity ID. One of the Activities could be flagged as quality Control (QC) (via the check box), thereby allowing it to be filtered out of Result outputs if the QC Activity sample was being considered a control (e.g., distilled water).</p> <p>Sample from Sample should only be used if a portion of the parent sample is being used for further analysis.</p>

User Support Log - Period: 03/23/2005 - 04/06/2005			
Date Received	System Version	Issue	Response
3/28/05	P2.0	<p>Migrating the STORET database using the manual process for converting data from 1.2 to 2.0 according to the instructions in migrate.txt.</p> <p>Following issues were raised.</p> <ol style="list-style-type: none"> 1. While checking the logs after completing step # 6: Grant permissions, I found that the file: storet-x\sto2\upgrade\stor20b1\sir712.sql did not run correctly. There is no connection string included in the SQL script. What user should I add as a connection script to this file? Will it hurt to run this file out of sequence? 2. I received a synonym error while running the file: storet-x\sto2\upgrade\stor20b1\sir1025.sql. The error stated: create synonym tsrca1t for storet.tsrca1t * ERROR at line 1: ORA-00955: name is already used by an existing object. Will this affect the migration at all? 3. There is a text file named: storet-x\sto2\upgrade\stor20b1\sir729.txt which has instructions for modifying the tnsnames file. Is this something that belongs here? 	<p>The following responses were made to the issues:</p> <ol style="list-style-type: none"> 1. Add "connect storuser/sto2user" before the line "spool log\sir712.log" and re-run the script. Running out of sequence should not cause any problem in this case. 2. This should not affect the migration. But the existing TSRCALT synonym should point to the STORET.TSRCALT table. This can be confirmed by the following query. In a SQL*Plus session, login as storuser (connect storuser/sto2user) and run the following: select * from user_synonyms where synonym_name='TSRCALT'; This should yield: SYNONYM_NAME = TSRCALT TABLE_OWNER =STORET TABLE_NAME =TSRCALT 3. This file gives instructions to help update the tnsnames.ora file manually. This can be achieved using Oracle provided Network Configuration Assistant tools. The STORET v2.0 Data Entry Module "Client-Server Installation" and "Standalone Installation" options will add such entries automatically. You can ignore this file.

User Support Log - Period: 03/23/2005 - 04/06/2005			
Date Received	System Version	Issue	Response
3/31/05	C2.0.4	Regarding the Central Warehouse Extract, Transform, and Load (ETL) process, the script related to the biological result fact table is running extremely slow. This script is using f_char_name function to fetch characteristic name. The concern was if rewriting the function f_char_name to use the di_characteristic table rather than TSRCHAR table would help resolve the issue.	Upon investigation, required indexes including the primary key index was missing on TSRCHAR. Creating this index helped the performance problem. Impact of changing the f_char_name function also forwarded.

User Support Log - Period: 03/23/2005 - 04/06/2005			
Date Received	System Version	Issue	Response
04/06/2005	C2.0.5	<p>At this point, I believe that users in the field of STORET version 2 do not possess the amended TSRUOM table, and need not run any script affecting it. Am I mistaken? Have we (through the DBFIX process) replaced TSRUOM with the version amended to show the conversion factors?</p> <p>Where do I install these fixes so that the warehouse inherits them? On K. Christian's ETL Instance (which I call STORFCH) ? Are any other instances that need to be fixed? Is there any reason to distribute these changes to the field?</p> <p>Row number (ISN) 289, Unit of Measure (UOM) "Grains" is a reference to a unit of mass/weight = 1/7000 pounds. I assume your conversion factor uses this value.</p> <p>We have recently added row 290, "Granules" in response to a request to store counts of mineral particles recovered by "panning" (requested by the National Park Service in support of reported outcomes in Alaska). I assume it should also be "converted" using 1.0, target unit "Granules". These changes will be released in DBFIX020.</p>	<p>The following answers were provided:</p> <ul style="list-style-type: none"> - Users do not need to run the script provided. - Users do not possess the amended TSRUOM table with Unit Conversion (at least not through dbfix019). - To incorporate the updated unit conversions, replace the file "unit_conversion.sql" provided with the one you already have; it was last delivered in directory "Release C2.0.4/software\storetw". Run the script as part of the ETL that builds the Central Warehouse database. I will assume this is K. Christian's ETL instance (STORFCH) which builds the Central Warehouse for export to RTP. Do not run the script separately as it will invalidate dependent objects, and will not generate converted results values stored in the fact tables. No other implementation is required. - The script is for the Central Warehouse. The script essentially creates a copy of TSRUOM called lu_unit_conversion with unit conversion table and data changes. This is not the implementation that will be distributed to the users upon release of the Reference Table Application; TSRUOM will be modified directly. Users who write programs dependent upon the "TSRUOM copy" may get confused when the "user-targeted" implementation is delivered. - In a second email listing UOM values slated for dbfix020, the only new value not in dbfix019 is "Granules". After researching this, it seems that the intended use provided in the UOM description would make this value unique. As a default, the unit_conversion.sql script will convert Granules to itself automatically; no unit conversion update is required.

Legend for System Version:

System Letter is concatenated with Version Number. System Letter Key: C = Central Warehouse, P = Data Entry Module