

2017-06-19 - Reporting SIG Notes

Date

19 Jun 2017

Attendees

- [Doreen Herold](#)
- [Anne L. Highsmith](#)
- [Joanne Leary](#)
- [Peter Murray](#)
- [Nassib Nassar](#)
- Erin Nettifee
- [Lydia Pettis](#) (convener)
- [Unknown User](#) (simona)
- [Charlotte Whitt](#) (minute taker)
- [Michael Winkler](#)

Goals

- Demo reports / Cornell
- Discuss "Reporting Database"

Discussion items

Time	Item	Who	Notes
	Welcome	Lydia	
20	Demo Cornell Reports	<p>Lydia shows Annex Moves, moving books into long term storage (50% of the current reporting activities; 3 mill. books in the past 15 years, and expect to move 1 mill. books the next 5-8 years).</p> <ul style="list-style-type: none"> • 40% of Cornell's stack is in foreign languages (250+ different languages), tremendous need for diacritics • Holding Statements can be quite long (more than 1 page), very large data sets • Lydia shows databases in Voyager Access. • Using MARC 852 866 867 868 780 785, ... for holdings data. Tag 880 for Chinese character set. These tags are concatenated together to create holdings. • Collection Development: Voyager Fund Selectors, managing about 100 Funds. Open orders with commitments. Question: which items have been received and payed for ? Calculate amount and subtotals in each section. Need to link persons as the owners of the fund. Each of the selectors have a copy of this database. • Circulation Summary • Circulation Statistics, 26 different queries are combined into one screen of circulation statistics. 	Capture of screen dumps
10	Presentation - Reporting Database	<p>Nassib Nassar (Index Data) presents IndexData's approach to the Reporting Service.</p> <p>FOLIO will have a reporting service separate from modules that actually provide data. Data for reporting will be accessed by a report generating tool. This tool will have direct access to a database schema.</p> <p>FOLIO generally reduces the dependency on a database schema, but in reporting there is no way of getting around this.</p> <p>Two approaches:</p> <p>1st approach: replicate the operational database; operational database = "aggregates", a collection of databases</p> <p>2nd approach: make two distinct databaes. 1. operational database 2. distinct database for reporting, "data warehouse"</p> <p>IndexData favors the data warehouse approach, they want to have two different schemas.</p> <p>For reporting: store data in a relational database in a de-normalized form which does not acquire views. Needs to be read-only. Communicates with the modules only via their documented service interface.</p> <p>1st approach has points of failure, for example if operational database schema is changed.</p> <p>Reporting database is going to look quite differenrent from the operational database.</p> <p>Reporting database can nevertheless be synchronized within not much less than real time.</p> <p>Anything which needs real data should use the operational database.</p>	

20	Discussion re: Rpt DB	<p>Ann has two remarks: "1." read-only will not be convenient / will not fit us. We will want to build apps on top of the reporting database and store data. "2." don't see a reason why not all the data should be in the reporting database.</p> <p>Nassib replies:</p> <p>A solution for "1." would be to replicate the reporting database.</p> <p>"2." : the operational database does not only store data which are interesting for reporting (e.g. administrative metadata to the data sets).</p> <p>Lydia: it is not unusual, in annex reports, to select bibs and items. One of the specifications might be: Has this item circulated within the last 7 years ? I.e., it is not unusual to cross modules, i.e. cataloging data and circulation data. It is not clear to Lydia what "aggregating" the data means (in Nassib's talk).</p> <p>Nassib: Aggregation = Pulling the data from the different modules into a single (relational) database. But the tables in a relational database are virtually independent. Nassib will not use the term "aggregation" for this anymore.</p> <p>Lydia: All the holdings should be pulled over to the reporting database.</p> <p>Nassib: The issue "diacritics" falls under "internationalization". Charlotte: if all bibliographical data is in UTF-8 in the Voyager database, it needs to be in UTF-8 in FOLIO as well. Michael: we want to go beyond that, e.g. support right-to-left written languages.</p>	
5	Planning for next session	<p>20 minutes Reporting Demo by Ann Highsmith.</p> <p>Lydia retires within the next weeks, we are looking for a convener until Cornell hires someone to replace Lydia (also as a convener).</p>	

Action items

