OpenMRS-DHIS2-SDMX-HD integration Project Plan

Project Details

This page is my proposal on what this project is and how it can be taken forward and what are it's challenges.

What is this project's objective?

Creating a renderer for the report module for mapping report output elements to the exchange format (SDMX-HD/DHIS2).

How i propose to develop it? project plan?

(Preferably the project would be focusing on cohort Indicator Data set reports to convert to SDMX-HD/DHIS2 in the beginning.)

<table>
<thead>
<tr>
<th>Week</th>
<th>Activities</th>
<th>Deliverables</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-3</td>
<td>1. Investigate(formats of sdmx-hd/dh2)</td>
<td>UI Mockups, Finalized Project schedule, Exact workflow of how the integration</td>
</tr>
<tr>
<td></td>
<td>2. Familiarize(Reporting module and messaging module)</td>
<td>between the reporting module and messaging module can be achieved.</td>
</tr>
<tr>
<td>4-6</td>
<td>User Interface</td>
<td>The user interface on the reporting module-selection of the format and the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>final display of the message.</td>
</tr>
<tr>
<td>7-9</td>
<td>SDMX-HD/DHIS2 integration</td>
<td>The code of the renderer.(java classes providing the functionality of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>transforming the data)</td>
</tr>
<tr>
<td>10-11</td>
<td>Buffer time for Cleaning and Testing</td>
<td>The code must be reviewed and test cases must be provided.</td>
</tr>
<tr>
<td>12</td>
<td>Documentation and tutorials</td>
<td>Written documentation on project with project details and technical details</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and how to use it, Video tutorial on how to use it.</td>
</tr>
</tbody>
</table>

Objectives of the project

1. Locate the DHIS2 instance
2. Get a template of all datasets and keep it in the file system
3. For a particular dataset, create a return template
4. Select a report
5. Display all the columns in the report and all the fields in the template
6. Allow the user to map from one to the other
7. Save the mapping in the return template
8. After the report is run, there will be a new return template with substituted values. The new return template is specific to the parameter values chosen when the report was run. This report can be sent to the DHIS2 instance corresponding to the dataset using the parameters that we keep on each DHIS2 installation

What to expect at the end of the project?

The following functionality can be expected at the end of the project:

**Step 1:** The user selects the output type to be SDMX/DHIS2+destination(where the message is to be sent) in the reporting module's interface. (the administrator would already have set up any transmission parameters on an administration page).

**Step 2:** The report's data is converted to the SDMX/DHIS2 and is forwarded to the messaging module.

**Step 3:** The message is sent to the destination address specified by the user through the messaging module. The destination address might be an email or a destination where the user can receive such aggregate data from multiple websites.

**Step 4:** The log of what happened is displayed i.e. as the message is sent through the messaging module it would be displayed to the user that the message has been successfully sent.

Doubts and Resolutions

The above content of the page will get updated based on the clarifications, so the doubts might seem irrelevant.
<table>
<thead>
<tr>
<th>No.</th>
<th>Date</th>
<th>Doubt</th>
<th>Resolution</th>
<th>Date</th>
<th>Resolved by</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>May 9, 2013</td>
<td>Is my flow of understanding correct? On taking data from reports and converting it into sdmx xml and displaying it to the user?</td>
<td>I was thinking that we would have a report renderer: from the user's point of view, it would be just like running a report and choosing the output type as SDMX/DHIS2+destination; the administrator would already have set up any transmission parameters on an administration page (see the messaging module). The period indicator report might be a good type of report, but maybe not the only type; it's probably a good place to start, but keep type-specific code isolated maybe in a helper class which implements an interface/extends a base class with a generic representing the type.</td>
<td>May 10, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>2</td>
<td>May 9, 2013</td>
<td>Is this only for taking aggregate data as those are finally viewed as indicators?</td>
<td>Probably we will deliver numerators and denominators separately, but some indicators will be counts only (not a ratio).</td>
<td>May 10, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>3</td>
<td>May 9, 2013</td>
<td>In step 1 I do we provide the option to select the aggregate data for multiple questions or single question per conversion.</td>
<td>Per 1, I would prefer to keep 1 report definition = 1 message, but if it is not easy to duplicate parts of report definitions, then it might be necessary to let users select those outputs to be packaged into the message.</td>
<td>May 10, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>4</td>
<td>May 9, 2013</td>
<td>After the conversion the xml must be displayed to the user? (ynl)</td>
<td>No, but it might be nice to have an xslt to allow them to see it in a simple format.</td>
<td>May 10, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>5</td>
<td>May 9, 2013</td>
<td>The xslt structure would remain constant or this must change on user selection? (this doubt is related to doubt 4).</td>
<td>I think the xslt structure would remain constant unless the message definitions changed.</td>
<td>May 10, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>6</td>
<td>May 9, 2013</td>
<td>Is the xslt the only integration between DHIS2 and this xml or do we provide some other kind of integration?</td>
<td>I think we should provide a message transport service like the messaging module and/or something special for MIRTH.</td>
<td>May 10, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>7</td>
<td>May 9, 2013</td>
<td>I suppose the user interface consists of providing an option for selection of question and display of xml in the reporting module itself? should it be different?</td>
<td>I think what the admin interface needs is a way to associate reports with messaging parameters and dataset contents with indicator definitions. I think in most cases there will be a message template which includes the indicator IDs, the association with dataset contents will be managed through the interface; you may need the ability to start with an empty template and build it as you go.</td>
<td>May 10, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>8</td>
<td>May 30, 2013</td>
<td>What can we take down as the projects objectives or the workflow could be?</td>
<td>- Extending reporting module's tables to persist the requested report's data for this project.</td>
<td>May 30, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>9</td>
<td>June 6, 2013</td>
<td>This is also a continuation to the previous doubt but this was discussed with the project champions.</td>
<td>1.export of dhis2 report                                                                 2.metadata export of dhis2 –questions form dhis2 3.create an OpenMRS report having all the indicators user wants mapped to the template report/template indicator/dataset/report show the disaggregations dhis2 wants and map them to the values in OpenMRS.</td>
<td>June 6, 2013</td>
<td>Roger Friedman</td>
</tr>
<tr>
<td>10</td>
<td>June 6, 2013</td>
<td>Todo list for me to get started on the project.</td>
<td>1. Use the DHIS2 webAPI to download the dataset definition (template) 2. Create a new annotation for different OpenMRS reporting framework indicators 3. Fill the annotations with created indicator ids (uuids) 4. Load the value into this and send them as data/ValueSet into DHIS2 (as the dhisreport module does)</td>
<td>June 6, 2013</td>
<td>Saptarshi Pukayastha and Roger Friedman</td>
</tr>
</tbody>
</table>
11 June 2013  
Project Beginners introducing on hisp india module and integration with dhis2

1. The hisp module assumes that the DHIS2 instance is accessible to the OpenMRS instance. I think that we still have to support a distributed (i.e. intermittent synchronization) model. Perhaps the easiest way to do this would be to have a DHIS2 instance co-located on the OpenMRS server; it could accept DHIS2 metadata updates from the central server. That might not work if there were multiple DHIS2-based systems in the country, e.g. a routine health data system run by the Ministry of Health and a TB indicator system run by an NGO. Alternatively, we could process the metadata export ourselves, perhaps using code from DHIS2, although I don’t like that tight coupling.

2. The hisp module assumes that the OpenMRS data maps to exactly one organizational unit in DHIS2. I think we must allow for an OpenMRS containing data on multiple organizational units; the central-and satellite form of deployment using the synch module is a common (but not well supported) use case for OpenMRS.

3. I would like to keep the message structure the same and use the <annotation> tag to store our mapping data. I would suggest we add an "annotation type" attribute to the XSD which would default to "sql" using the body of the tag; our mappings should be expressed in other attributes on the annotation. We are also going to have to define a way of mapping the category option combos to OpenMRS concept values or reporting cohorts. More on this later.

4. The existing module has a single location and a single data element and a single category combo value per annotation (SQL query). Our hope is to collect multiple locations, data elements and category combo values per OpenMRS report definition, then generate multiple data rows (perhaps in multiple data files) from them. We would be sorting report outputs into DHIS2 bins rather than iterating over DHIS2 bins and getting their contents. Note that it is a very common DHIS2 use case for multiple data elements to be used in lieu of a single data element with categories, so we might be using filter values to distinguish between data elements as well as between categories. So we are mapping location, data element and category combo value on the DHIS2 side to dataset, cohort, variable, filter values on the OpenMRS side (all within a single DHIS2 dataset or OpenMRS report specification).

5. The existing sdmx module (Ryan Crichton’s) is already integrated into the reporting framework. I chatted with Mike Saason this morning and he suggested looking at Reporting Module release notes subsequent to 0.7.3 on which the sdmx module is based to see what non-backwards-compatible changes have been introduced. Much of the guts of the hisp module would be used (the transfer file format and operations based on it); I have to look at both.

12 June 2013  
First Testing Use case for trying out the functionality

1. OpenMRS large dataset is basis

2. Data will be obs, datatime in [2006/02,2006/06)

3. Data will be obs of concepts 5356 CURRENT WHO STAGE and 5497 CD4 COUNT. The answers to concept 5356 are concepts 1204-1207, WHO STAGE 1-4 ADULT.

4. Monthly indicators:
   a) NUMBER OF PATIENTS OBSERVED WITH WHO STAGE 1-4 (4 indicators), potentially disaggregated by age (0-5,5-14,14-21,21-49,50+) and sex.
   b) NUMBER OF PATIENTS OBSERVED WITH ANY WHO STAGE, sum of indicators in A, potentially disaggregated by age (0-5,5-14,14-21,21-49,50+) and sex.
   c) NUMBER OF PATIENTS WITH A CD4 COUNT, potentially disaggregated by CD4 Count<=250 and CD4Count>250; and
   d) NUMBER OF NEW PATIENTS BY CD4 COUNT, where a new patient is one whose first CD4 count is in the reporting month, potentially disaggregated by CD4 Count<=250 and CD4Count>250; and
   e) NUMBER OF NEW PATIENTS BY CD4 COUNT, where a new patient is one whose first CD4 count is in the reporting month, potentially disaggregated by age (0-5,5-14,14-21,21-49,50+) and sex.

5. The test data has only one location. It will be modified on the following basis: new locations will be created 2 REGION A, 3 HOSPITAL X and FACILITY Y. (See ticket REPORT-288)

13 June 2013  
Tips for starting the module

1. Exploring Reporting module with test data and period indicator reports.

14 June 2013  
What exactly are disaggregations

Disaggregations are segregation of the report data obtained into a more detailed data parts having information about them. Meaning if a data such as current who stage is generated, separating the data obtained based on age groups can be called disaggregation.

15 June 2013  
Suggestion by Roger to start coding

Create a basic module using the maven archetype which can take in data definition.

16 June 2013  
I am facing difficulties in creating the disaggregations in the report

That's an area we need to investigate and discover how we need to achieve that.

17 June 2013  
Is the following flow

1a> I'm trying to run the definitions that are existing in the database (the ones created as cohort queries)  
2a> I'll work on giving controls to take in definitions from the user which can be added from a function  
3a> Then I plan to take in the DHIS2 xml and map it to our data elements  
4a> Run the Reports and Generate the SDMX-HD message.

Rather than giving an option or writing a function it's better to use the preexisting api.

In the first step also instead of using the services try to use the reporting api directly.

18 June 2013  
In the first step mentioned above is it fine if I use the services such as the ones used in the query parameter class? to obtain the

It's better if in the beginning we take the test data and analyze that and provide the functionality from the testing procedure and later on we can add more complexities based on the requirement.