

AUG Email 2013-Feb

Last Updated: 03/19/2013

Studies i2b2 query design input solicited [02-27-2013]

From: Bauer, Christian [christian.bauer.med.uni-goettingen.de] **Sent:** Wednesday, February 27, 2013 9:30 AM **To:** i2b2 AUG Members **Subject:** AW: #studies i2b2 query design input solicited

Hi,
for our i2b2 installation in Goettingen we were tasked by researchers to provide a way to query for specific biospecimen and also give a quick overview and export options for the found specimen (patient(1) -> specimen (n) -> facts_about_specimen(n)). Also the normal functionalities of i2b2 for querying facts and getting the patient (patient(1) -> patient_facts (n)) should be available at the same time.

We used modifiers to bundle all the concepts of the ontology which were about the specimen facts. We also gave each specimen an instance_num specific to a patient_num. We then used the patient_mapping table for storing each specimen's original id (patient_num + instance_num -> barcode of a specimen). We further developed a plugin for the i2b2Workbench which uses the qm_id of a standard i2b2-query for getting all the necessary data from the observation_fact to display/export facts about the found specimen.

With our approach we were able to archive our goal to search for and extract data of specific specimen, but we lost the ability to use the instance_num for grouping some information of a specimen. An additional id in the observation_fact table that represents a specimen (or any other entity a patient could have n of) could help here. We have done this for a specific project and want to incorporate it into a more generic import/plugin-workflow. So we are very interested in a core i2b2 solution.

Christian Bauer
Christian Bauer
UNIVERSITÄTSMEDIZIN GÖTTINGEN
GEORG-AUGUST-UNIVERSITÄT
Department of Medical Informatics
Robert-Koch-Straße 40
37075 Göttingen
Germany
Mail: 37099 Göttingen
Phone: +49551 / 39-8227http://www.universitaetsmedizin-goettingen.de/index_en.html

Von: Valtchinov, Vladimir I., Ph.D. [VVALTCHINOV.PARTNERS] **Gesendet:** Dienstag, 26. Februar 2013 22:47 **An:** 'Russ Waitman'; 'i2b2 AUG Members' **Cc:** 'Nathan Graham'; 'Dan Connolly' **Betreff:** #studies i2b2 query design input solicited

Hello Russ-
thanks for your informative response, and sorry for the delayed response on my part.

Thanks also for reviewing the similar initiatives you have at KUMC regarding ordering data – for clinical diagnostic labs, for radiology studies, and orders for medications. To your question regarding existence of a standard ontology related to DSS – no, we were not aware of any such pre-existing ontology. However, certain requirements of the CMS-based Medical Imaging Demonstration project predetermined the workflow and the structure of Decision Support process for radiology ordering CPOE, and we have used that to define a novel i2b2 view of the steps it takes to order/perform a radiology study. Let me know if you'd like to hear more specifics, we can chat about it offline. As a side note – I'd be very interested to chat about RStudio <-> i2b2 Web Client integration as well.

Here I'd like to solicit another set of inputs.

Shawn and I were talking about the patient-count centric design of i2b2 (this is by design – the system is designed for discovering patient cohorts after all), and what is needed to make it suitable for structuring and querying #studies types of aggregate information, i.e. suitable for operations research of the type of the CMS radiology study. It is evident that the i2b2 schema is general enough to allow for structuring the information. The system then needs to allow for a general #studies query to be generated, and the corresponding count and studies list (i.e. encounter_num-s) to be returned.

So Shawn and I agreed it might make sense to actually add to the type of queries i2b2 can generate, defining a generic '# studies' query. Here I wanted to address the wider i2b2 community to gather opinions what people might have seen, so when we at the Brigham make those changes the new functionality serves the community and can be incorporated in future i2b2 back-end releases.

For our specific case (radiology CPOE and DSS), we have one-order-generating-one-exam type of a relationship, and also each exam has a unique encounter_num as recorded in Visit_Dimension. This results in a very simple change of the underlying SQL:

take 'distinct encounter_num' instead of 'distinct patient_num' in the normal i2b2 queries.

Please write back to me if you have similar use cases for i2b2 and might suspect you need a different, more complex

#studies query.

Best,

vlad

CEBI, i2b2

From: Russ Waitman [rwaitman.kumc] **Sent:** Tuesday, February 12, 2013 6:41 PM **To:** Valtchinov, Vladimir I., Ph.D.; 'i2b2 AUG Members' **Cc:** Nathan Graham; Dan Connolly **Subject:** RE: need help: i2b2 queries against radiology CPOE data

Hi Vlad,

Sorry for the delayed reply and I don't know if others responded to you individually. We are doing some related work. I may have struggled to follow you so a call might be warranted.

That said, I've had related thoughts regarding the life cycle of an "order", related decision support and processes across i2b2 observations and is there a standard ontology?

We also have billing CPTs for the professional radiology interpretation, and also now the CPOE orders. I'd like to bring in decision support suggestions but we haven't gotten there yet with Epic.

In the lab domain, we'd have the CPOE order, the lab panel results, and in some cases an interpretation.

In a more concrete area, we're using the modifier feature to tag "medications" as either home meds/historical meds, orders (inpatient or outpatient), dispenses, nursing administration (MAR), and PRN flag.

In those cases like you, yes, we have to load a new row for each modifier. But, in some cases it make sense as it's a separate timed observation (physician order versus nurse activity) but in other cases it's redundant (ex: PRN flag).

Ideally as you indicating, one might want to explicitly tie orders to diagnoses or other observations. When you start down this path, the EAV star model might get a bit inefficient. I could thought we could use a modifier to hold the pointer and then build some kind of advanced tools or plugin to exploit the pointer in the user interface. Then you also start adding performance implications and Nathan Graham has the most experience in this area on our

team. Our current release is pretty fast on 1 billion facts except when you try to use the same financial encounter feature and then it crawls but Nathan says it will be fixed in our next monthly release after we sacrifice 20 pineapples in honor of Larry Ellison's purchase of the Hawaiian island Lanai (it's either that or voodoo stats/hints/index magic).

For the most part, i2b2 just gives you the patient count in the UI and you need to pull out the data to do sub analysis. So, I think it's working as designed. If you want the number of studies, you don't get that as an immediate count (at least we don't using 1.6.07 and the web client). Today, we'd wind up having the user define that cohort and then do a data request to pull the data and then they can count studies or do whatever analysis they desire. In the near future we're working on an i2b2 -> RStudioServer plugin so the could pipe the cohort and study data into a data frame and do the sub analysis online. I can see times though, especially for inpatient processes, you'd want the UI to return patient counts and encounter counts.

Russ Waitman, PhD

Director of Medical Informatics

Assistant Vice Chancellor for Enterprise Analytics

Associate Professor, Department of Biostatistics

913-945-7087 (office)

rwaitman.kumc@informatics.kumc.edu

From: Vladimir I.Ph.D. Valtchinov [VVALTCHINOV.PARTNERS] **Sent:** Thursday, January 17, 2013 4:23 PM **To:** Russ Waitman; 'i2b2 AUG Members' **Subject:** need help: i2b2 queries against radiology CPOE data

Hello All-

we have implemented a dual-purpose i2b2 repository at the BWH Radiology Department for the CMS-based Medical Imaging Demonstration project (<http://www.diagnosticimaging.com/practice-management/content/article/113619/1792788>).

It would support both a standard, patient-cohort discovery with criteria related to patient record and medical imaging findings. Additionally, it would allow for operations research queries, interrogating the system generally about number of studies performed in a given imaging modality domain.

The source for the operations research data is the complete record of the CPOE process that takes place between the ordering physician (i.e. the provider), for a specific patient, and the 'CPOE system' and include things like a) signs and symptoms for the patient as up to 10 Dx, their ICD9s, the b) original image order, the CPT4, then c) as many as 5 iterations between the PCP and the CPOE Decision support software where DSS had different imaging studies suggested and the PCP accepted or rejected those, with a structured menu for the reason, and finally d) recording the final imaging study that got performed (CPT4) as well as the final Dx as a result of the study.

We have modeled each one of the events/DSS suggestions from above as a separate concept_id with a modified_id for the multiplicity the measure (i.e. there were 10 modifier_id-s for the ICD9-s in item a) above. This would result in up to ~35 FACT tables rows for each of the many fields recorded in each row of the image ordering appropriateness data source, MID.

An example of a question one would 'ask' the system is: What is the number of studies done by (say BWH Providers) that have ordered MRI of the head but the DSS system suggested CT of the head w/w/o Contrast (CPT=70470), which was performed as appropriate, and the final study was (insert a CPT)'. The mapping of the CPOE source to the i2b2 schema would we think support queries like that.

They do seem to work in most of the cases, but there are some cases when I get underestimation of numbers.

We think it all boils down to using 'distinct patient_num' as part of the default (#patients) query in the i2b2.

This is ok for identifying patient cohorts (you need the list of unique pt id-s with that specific combination of parameters) but it gives lower count numbers when you need 'number of studies' type of count as in operational research such as CPOE data.

The exact degree of underestimation is the number of distinct pt_id-s vs. the number of visits – if we have each and every visit number to be associated w/ different pt_id (as it happens most of the time), the counts are accurate.

The I have looked at the 'Occur in the Same encounter' grouping option under the Time Constraint box.

This I think corresponds to our case – you have multiple 'measurements' in the same encounter – we end up having up to ~35 Fact rows for the same pt_num, provider_id and visit_num.

Using SQL Profiler I was able to track the sql statements to (apart from some DDL):

insert data

```
insert into i2b2radiology2.i2b2demodata.#global_temp_table ( encounter_num, patient_num , panel_count)
select encounter_num, patient_num , 1 as panel_count from
( select /*+ index(observation_fact observation_fact_pk) */ encounter_num, patient_num
from i2b2radiology2.i2b2demodata.observation_fact where provider_id IN
(select provider_id from i2b2radiology2.i2b2demodata.provider_dimension
where provider_path LIKE 'MID Providers%') group by encounter_num , patient_num ) t
```

update stats

```
UPDATE STATISTICS i2b2radiology2.i2b2demodata.#global_temp_table
```

insert into new temp table

```
insert into i2b2radiology2.i2b2demodata.#dx ( patient_num , encounter_num )
select * from ( select distinct patient_num , encounter_num from i2b2radiology2.i2b2demodata.#global_temp_table
where panel_count = 1 ) q
```

The select distinct patient_num , encounter_num seems to be doing the trick – I get the correct counts when I execute the sql in to this point in SQL studio.

The Web Client returns the old number (taking just distinct pt_num), and I don't know why.

So, in general, I'd like to request for your guys comments here. Even more importantly --have any of you seen this before or maybe are pursuing the same structure – maybe this is a known issue w/ known workaround (others have maybe done it?)

Thank you much in advance, and sorry for the long message and messy formatting.

Best,

vlad

CFP – Bio-Ontologies SIG 2013, July 20,2013, Berlin, Germany [02-21-2013]

From: Nigam Shah [nigam.stanford] **Sent:** Thursday, February 21, 2013 9:38 PM **To:** bioontologies.gmail **Subject:** CFP - Bio-Ontologies SIG 2013, July 20, 2013, Berlin, Germany

Dear Colleagues,
[50 days to submission deadline]

What: Call for participation in the Bio-Ontologies SIG 2013 (www.bio-ontologies.org.uk)

Where: July 20, 2013, Berlin, Germany. (Co-located with ISMB/ECCB 2013) *When: *Submissions Due: April 12th, 2013 (Fri) @ www.bio-ontologies.org.uk/submissions

Notifications: May 10th, 2013 (Fri)

Final Version Due: May 17th, 2013 (Fri)

Please forward to everyone who might be interested.

The Bio-Ontologies SIG provides a forum for discussion of the latest research in the application of ontologies and in the organisation, presentation and dissemination of knowledge in the life sciences. In its 15th year, Bio-Ontologies is one of the longest running SIG at ISMB. Papers are invited in areas, such as the applications of bio-ontologies, newly developed bio-ontologies, and the use of ontologies in data sharing standards. Example topics include (but not limited to):

Applications of ontologies in bioinformatics

Hypothesis Testing Platforms

Use of Ontologies in Data Standards

"Flash updates" on Newly Developed or Existing Bio-Ontologies

Bio-Curation Platforms

Automated Annotation Pipelines

Efforts using ontologies for Bio-NLP or Information Retrieval

Semantic Web Enabled Applications

Advances in development of biomedical ontologies

Collaborative Ontology Authoring and Peer-Review Mechanisms

Automated Ontology Learning

Mapping between Ontologies

Research in Ontology Evaluation

Using games for Ontology review and evaluation

We invite three types of submissions. Short papers, up to 4 pages. Poster abstracts, up to 1 page. Flash updates, up to 1 page. Following review, successful papers will be presented at the Bio-Ontologies SIG. Poster abstracts will be provided poster space and time will be allocated for a flash update on the poster. Flash updates are for short talks (5 min) giving the salient new developments on existing public ontologies. Unsuccessful papers will automatically be considered for poster presentation.

Best regards,

SIG Organizers: Nigam Shah, Susanna Sansone, Michel Dumontier, Larisa Soldatova

Using SQLSERVER: USER, SCHEMA, DATABASE, LOGIN [02-20-2013]

From: Sebastian Mate [Sebastian.Mate.imi.med.uni-erlangen.de] **Sent:** Wednesday, February 20, 2013 11:04 AM **To:** 'Darren W Henderson' **Cc:** 'i2b2 AUG'; Mendis, Michael E. **Subject:** AW: Using SQLSERVER: USER, SCHEMA, DATABASE, LOGIN ...

Hi Darren,

it works – the Ontology is loading! There was indeed an error in the ont-ds.xml file.

Thanks again!

Sebastian

PS: This is my setting, which also works:

```
<!-- Ontology DS for Demo -->
<local-tx-datasource>
  <jndi-name>OntologyDemoDS</jndi-name>
  <connection-url>jdbc:sqlserver://xxxxxxx:xxx;DatabaseName=l2B2METADATA</connection-url>
  <driver-class>com.microsoft.sqlserver.jdbc.SQLServerDriver</driver-class>
  <user-name>l2B2METADATA</user-name>
  <password>xxxxxxx</password>
  <idle-timeout-minutes>1</idle-timeout-minutes>
  <set-tx-query-timeout>true</set-tx-query-timeout>
  <metadata>
    <type-mapping>MS SQLSERVER2000</type-mapping>
  </metadata>
</local-tx-datasource>
```

Von: dwhend0.gmail **Im Auftrag von** Darren W Henderson **Gesendet:** Mittwoch, 20. Februar 2013 16:41 **An:** Mendis, Michael E. **Cc:** Sebastian Mate; i2b2 AUG **Betreff:** Re: Using SQLSERVER: USER, SCHEMA, DATABASE, LOGIN ...

Sebastian and Mike, in our sqlserver setup I can assure it works with: ontology.bootstrapdb.metadataschema=dbo

Your new error looks more like an issue with your connection strings in your CRCAApplicationContext.xml, ont-ds.xml. (I think)

They should look similar to the following connection string:

```
<driver-class>com.microsoft.sqlserver.jdbc.SQLServerDriver</driver-class>
<connection-url>jdbc:sqlserver://<IP address>;user=<uname>;password=<passwd>;</connection-url>
<metadata>
  <type-mapping>MS SQLSERVER2000</type mapping>
</metadata>
```

On Wed, Feb 20, 2013 at 10:25 AM, Mendis, Michael E. wrote:

for ontology.bootstrapdb.metadataschema=dbo I think it should be ontology.bootstrapdb.metadataschema=i2b2hive.dbo

On Feb 20, 2013, at 10:16 AM, Sebastian Mate wrote: Hi Mike, thanks, I forgot about that. It was actually set to SQLSERVER by the demodata loader

script. Thanks to some feedback from the AUG I have reset the schema to be "dbo". However, if I start the server I get the following error message (Ungültiger Objektname = invalid object name): 07:05:37,868 ERROR [MetadataDbDao] PreparedStatementCallback; bad SQL grammar [select * from I2B2HIVE.ont_db_lookup where LOWER(c_domain_id) = ? and LOWER(c_project_path) like ? and (LOWER(c_owner_id) = ? or c_owner_id = '@') order by c_project_path]; nested exception is com.microsoft.sqlserver.jdbc.SQLServerException: Ungültiger Objektname 'I2B2HIVE.ont_db_lookup'.07:05:37,869 ERROR [AxisServlet] org.apache.axis2.AxisFault: Database error; nested exception is: edu.harvard.i2b2.common.exception.I2B2DAOException: Database error; nested exception is: org.apache.axis2.AxisFault: Database error; nested exception is: edu.harvard.i2b2.common.exception.I2B2DAOException: Database error; nested exception is: org.apache.axis2.AxisFault: Database error; nested exception is: edu.harvard.i2b2.common.exception.I2B2DAOException: Database error; nested exception is: org.apache.axis2.AxisFault: Database error; nested exception is: edu.harvard.i2b2.common.exception.I2B2DAOException: Database error If I set ontology.bootstrapdb.metadataschema=dbo (is this the correct setting when using SQLSERVER?) in jboss-4.2.2.GA/server/default/conf/ontologyapp/ontology.properties, I get the following error: 07:11:48,044 WARN [JBossManagedConnectionPool] Throwable while attempting to get a new connection: nullorg.jboss.resource.JBossResourceException: Could not create connection; - nested throwable: (com.microsoft.sqlserver.jdbc.SQLServerException: The TCP/IP connection to the host has failed. java.net.ConnectException: Connection refused) at org.jboss.resource.adapter.jdbc.local.LocalManagedConnectionFactory.createManagedConnection (LocalManagedConnectionFactory.java:179) at org.jboss.resource.connectionmanager.InternalManagedConnectionPool.createConnectionEventListner(InternalManagedConnectionPool.java:577) at org.jboss.resource.connectionmanager.InternalManagedConnectionPool.getConnection(InternalManagedConnectionPool.java:262) at org.jboss.resource.connectionmanager.JBossManagedConnectionPool\$BasePool.getConnection(JBossManagedConnectionPool.java:500) at org.jboss.resource.connectionmanager.BaseConnectionManager2.getManagedConnection (BaseConnectionManager2.java:341) at org.jboss.resource.connectionmanager.TxConnectionManager.getConnection (TxConnectionManager.java:315) at org.jboss.resource.connectionmanager.BaseConnectionManager2.allocateConnection(BaseConnectionManager2.java:396) at org.jboss.resource.connectionmanager.BaseConnectionManager2\$ConnectionManagerProxy.allocateConnection (BaseConnectionManager2.java:842) at org.jboss.resource.adapter.jdbc.WrapperDataSource.getConnection(WrapperDataSource.java:88) at org.springframework.jdbc.datasource.DataSourceUtils.doGetConnection(DataSourceUtils.java:111) at org.springframework.jdbc.datasource.DataSourceUtils.getConnection(DataSourceUtils.java:77) at org.springframework.jdbc.core.JdbcTemplate.execute(JdbcTemplate.java:463) at org.springframework.jdbc.core.JdbcTemplate.query(JdbcTemplate.java:529) at org.springframework.jdbc.core.JdbcTemplate.query(JdbcTemplate.java:562) at org.springframework.jdbc.core.JdbcTemplate.query(JdbcTemplate.java:571) at org.springframework.jdbc.core.JdbcTemplate.query(JdbcTemplate.java:610) at org.springframework.jdbc.core.simple.SimpleJdbcTemplate.query(SimpleJdbcTemplate.java:107) at edu.harvard.i2b2.ontology.dao.ConceptDao.findRootCategories(Unknown Source) at edu.harvard.i2b2.ontology.delegate.GetCategoriesHandler.execute(Unknown Source) at edu.harvard.i2b2.ontology.ws.ExecutorRunnable.run(Unknown Source) at java.lang.Thread.run(Thread.java:679)Caused by: com.microsoft.sqlserver.jdbc.SQLServerException: The TCP/IP connection to the host has failed. java.net.ConnectException: Connection refused at com.microsoft.sqlserver.jdbc.SQLServerException.makeFromDriverError(Unknown Source) at com.microsoft.sqlserver.jdbc.SQLServerConnection.connectHelper(Unknown Source) at com.microsoft.sqlserver.jdbc.SQLServerConnection.loginWithoutFailover(Unknown Source) at com.microsoft.sqlserver.jdbc.SQLServerConnection.connect(Unknown Source) at com.microsoft.sqlserver.jdbc.SQLServerDriver.connect(Unknown Source) at org.jboss.resource.adapter.jdbc.local.LocalManagedConnectionFactory.createManagedConnection(LocalManagedConnectionFactory.java:171) ... 20 more07:11:48,046 ERROR [ConceptDao] Could not get JDBC Connection; nested exception is org.jboss.util.NestedSQLException: Could not create connection; - nested throwable: (com.microsoft.sqlserver.jdbc.SQLServerException: The TCP/IP connection to the host has failed. java.net.ConnectException: Connection refused); - nested throwable: (org.jboss.resource.JBossResourceException: Could not create connection; - nested throwable: (com.microsoft.sqlserver.jdbc.SQLServerException: The TCP/IP connection to the host has failed. java.net.ConnectException: Connection refused)) I'm not sure if this is a different error and the previous one has been fixed. I will have to do some additional debugging ... :-/ Sebastian

From: Mendis, Michael E. Sent: Tuesday, February 19, 2013 8:34 AMTo: Sebastian MateCc: i2b2 AUGSubject: Re: Using SQLSERVER: USER, SCHEMA, DATABASE, LOGIN ...

Sabastian,
In the CDC_DB_LOOKUP, WORK_DB_LOOKUP and ONT_DB_LOOKUP the C_DB_SERVERTYPE would be SQLSERVER
mike

On Feb 19, 2013, at 8:00 AM, Sebastian Mate wrote: Hi all, I'm trying to extend our i2b2-Wizard (see <http://www.imi.med.uni-erlangen.de/~matesn/i2b2wizard/>) to also support SQLSERVER. It will also support other Linux distros besides Ubuntu! However, I'm having some difficulties in getting i2b2 (1.6.08) running with SQLSERVER. I don't have a clear understanding on how I should prepare the database. I've tried to set USER = SCHEMA = DATABASE = LOGIN (see SQL below). E.g. for I2B2HIVE, the schema is no longer "dbo" but "I2B2HIVE". This seems to be the correct setting because the server source code is trying to access a schema "I2B2HIVE" during start-up (when it reads the *_DB_LOOKUP tables). On the other hand, the "create_procedures_release_1-6" job of the demo data complains that it can't access the "dbo" schema. Can someone please give me a hint how I should handle USER, SCHEMA, DATABASE and LOGIN? Other than that I've configured the *-ds.xml (JBoss data sources) and *ApplicationContext.xml files for SQLSERVER, but I haven't really tested this, yet. Is there anything else that I need to change in order to use SQLSERVER? Many tanks & Greetings from ErlangenSebastian

Code to create a new user (I2B2_USR and I2B2_PWD are being replaced with real values during execution):CREATE DATABASE I2B2DB_USR;CREATE LOGIN I2B2DB_USR WITH PASSWORD = 'I2B2DB_PWD', DEFAULT_DATABASE = I2B2DB_USR, CHECK_POLICY=OFF; USE I2B2DB_USR; CREATE SCHEMA I2B2DB_USR; CREATE USER I2B2DB_USR FOR LOGIN I2B2DB_USR; GRANT ALL TO I2B2DB_USR; ALTER USER I2B2DB_USR WITH DEFAULT_SCHEMA = I2B2DB_USR; ALTER AUTHORIZATION ON SCHEMA::I2B2DB_USR TO I2B2DB_USR; Dipl.-Inf. Sebastian Mate

Friedrich-Alexander-University of Erlangen-Nuremberg Department for Medical Informatics Krankenhausstrasse 12, 91054 ERLANGEN, GERMANY
Phone +49 9131/85-26722, Fax +49 9131/85-26754 Sebastian.Mate@imi.med.uni-erlangen.de<http://www.imi.med.uni-erlangen.de>

Save the Dates [02-19-2013]

From: Churchill, Susanne E. **Sent:** Tuesday, February 19, 2013 3:50 PM **To:** i2b2 AUG Members **Subject:** Save the Dates
Please mark your calendars:

1. **i2b2 Workshop at AMIA Translational/Clinical Informatics Meetings** in San Francisco, Wednesday, March 20, 2013, 6-8 pm, Parc 55 Hotel (Room TBA).
2. Our back-to-back **i2b2/SHRINE Annual Meetings** in Boston (New Research Building, 77 Avenue Louis Pasteur)

Tuesday, June 18, 2013

NLP Workshop in morning

AUG Meeting Part 1 (12-5 pm)

Wednesday, June 19

AUG Meeting Part 2 (9-12)

SHRINE Meeting Part 1 (12-5)

Thursday, June 20

SHRINE Meeting Part 2 (9-12)

Agendas/logistics to follow; content requests welcome.

Regards,

Susanne