

# AUG Email 2010\_Jul-Dec

Last Updated: 12/09/2010

---

## AMIA CRI Summit – Request for Information [12-09-2010]

*On 12/09/10 9:33 AM, Shawn Murphy wrote:*

Dear Colleagues,

The Scientific Program Committee (SPC) of the 2011 AMIA Clinical Research Informatics Summit (<http://jointsummits2011.amia.org/>) is exploring options to provide members of the CRI with content targeting practical, applications-oriented solutions to common data, information, and knowledge management needs commonly found in the clinical research environment. As part of these efforts, we would like to solicit your input and feedback, particularly with regard to the following four questions:

1. What types of practical, application-oriented solutions would you be most interested in learning about if you if you were to attend the 2011 AMIA Clinical Research Informatics Summit? Examples could include i2b2, REDCap, Profiles, Vivo, and caGrid/TRIAD.
2. Would the availability of this type of content make the 2011 AMIA Clinical Research Informatics Summit more desirable for you to attend?
3. What type of format would be best for delivering content targeting practical, application-oriented solutions to common data, information, and knowledge management needs commonly found in the clinical research environment? Examples could include professional networking opportunities, workshops, demonstrations, or tutorials.
4. Do you have any other comments or suggestions regarding the format and content of the 2011 AMIA Clinical Research Informatics Summit?

Thank you in advance for your input and comments – we appreciate your help in making the 2011 AMIA Clinical Research Informatics Summit an exciting and highly valuable meeting for members of the CRI community.

Please contact Shawn Murphy, Paul Harris, or Philip Payne with your ideas.

---

## Using i2b2 in a Cancer Registry [12-03-2010]

*On 12/03/10 9:12 AM, Zak Kohane wrote:*

Two efforts that I am aware of:

1. IBD registry led out of Cincinnati
  2. JRA registry led out of CARRANet
- members of both on this mailing list and I think each network has 30-60 sites.
- Zak

*On 12/03/10 8:05 AM, Jeff Cowall wrote:*

Is anybody out there using or considering i2b2 in relation to a traditional Cancer Registry, or with NAACCR or Impac/MetriQ data?

I'd like to get in contact if you are.

Thanks,

Jeff Cowall

---

## i2b2 Web Client Additional Plug-ins [12-03-2010]

*On 12/03/10 8:04 AM, Shawn Murphy wrote:*

Hi Paul,

There are some plug-ins included with the main web client distribution, but none currently distributed separately. They will be coming in the release of 1.6 however, and more yet as the community web client project develops.

Thanks,

Shawn.

*On 12/02/10 3:03 PM, Paul Avillach wrote:*

Hi,

I'm looking for additional plugins already developed to analyze patient sets in the web client 1.5.

Is there a central repository?

Best Regards

Paul Avillach

---

#### SNOMED-CT with i2b2 [11-19-2010]

**On 11/19/10 9:32 AM, Keith Marsolo wrote:**

There was a group at UTHealth that presented a poster at this week's AMIA conference that talked about doing a mapping between SNOMED and i2b2. I'm not sure if anyone involved in that research is on the AUG mailing list, but it seemed like it would be possible to leverage their work. I talked to the main author, Krishna Kookal and he said they weren't quite ready to release their code, but would do so sometime soon.

Here's a link to the poster via the infuriating new AMIA proceedings website: <http://proceedings.amia.org/120tpv>

Since it's also indexed via Google, I'm assuming the site is open to the public, but I don't know if that's unintentional or by design.

Keith

**On 11/19/10 9:17 AM, Bill Adams wrote:**

At BU we are moving incrementally starting with EHR-based vitals and growth concepts (only 3000+ terms to go after that...) to support our cardiovascular outcomes project.

Happy to share our vitals to SNOMED-CT map file when ready.

Bill Adams

**On 11/19/10 5:28 AM, Nick Holden wrote:**

Thanks, Lori.

I suspect that our questionnaire covers more than one category of SNOMED codes, although it has a cardiovascular focus - we ask about the participant's medical history, their family history, their risk factors and their current and previous treatment and interventions.

I've seen several projects around the NLP challenge which talk about taking free text and seeking to code it in SNOMED for use in i2b2 which was why I hoped someone had already built a SNOMED terms ontology.

Nick

**On 11/17/10 8:47 AM, Lori Phillips wrote:**

Hi Nick,

You would need to create a metadata table that is comprised of SNOMED terms. i.e. the ontology in the Navigate Terms view would be SNOMED terms. In addition, the concept\_dimension table would need to be populated with these terms as well.

As luck would have it, I am currently working on an effort to provide SNOMED terms in a format that i2b2 understands. The primary challenge I am facing right now is handling the volume of data that is generated. Is there is a sub category of SNOMED terms that you are primarily interested in? It would be easier to try to convert a subset of SNOMED instead of the entire ontology.

Lori

**On 11/17/10 8:02 AM, Nick Holden wrote:**

Hi all,

This is probably a really basic question but if we have a source of data which is currently coded only with a locally-devised questionnaire coding system, and we want to map the data to a recognised standard, are there preferences for doing so in order to make the process of loading the data into i2b2 as simple as possible? For instance, if we mapped the questionnaire to SNOMED-CT terms (which is widely used in the NHS), would that be understood by i2b2 without further work?

Thanks,

Nick

---

#### Expanding Patient Recordset Tree View Hangs [12-02-2010]

**On 12/02/10 5:36 PM, Dan Connolly wrote:**

Indeed, replacing nulls with '@' in language\_cd, marital\_status\_cd, and religion\_cd of patient\_dimension relieved the symptoms. Thanks again for the clue, James.

**On 12/02/10 2:51 PM, Peter Beninato wrote:**

Hi,

I think there is a mechanism that uses different pools in i2b2 for longer running queries.

So if the query runs for 1800 seconds it goes to a different pool. Maybe there is a handoff problem in the webclient...

What is the parameter for setting the query time limit (referenced below)?

Can anyone elaborate on the query pool mechanisms in i2b2?

Thx.

Peter Beninato

**On 12/02/10 2:46 PM, Dan Connolly wrote:**

OK, thanks; that's an interesting clue...

It's quite likely that I didn't replace nulls with @ in the relevant data set here.

I have never understood what those @ signs are doing there.

Is it documented anywhere? Can anybody explain it?

**On 12/02/10 2:37 PM, Jacob Weiser wrote:**

Hi,

This has happened to me as well(The animated arrow image spins forever) we are on ver 1.5 when trying to debug the error files I have found that this can be caused either due to having characters that the webclient/xml does not like or when running passed a installation defined query time limit(in my case one minute) will cause the arrow to spin indefinitely with the only way out is to log out/in again.

Another caveat to this is in some instances it seems to be hanging without any results but when logging out/in the query results do show up?

The issue here seems to be with I2B2 spinning forever whenever some error is encountered?

Jacob Weiser

**On 12/02/10 2:15 PM, James Law wrote:**

Hi,

Not sure if you're seeing this in both the webclient and the eclipse client, but we've seen this problem when our data had characters that the webclient did not like, or when we haven't replaced nulls with @ signs.

There were Jira items related to these issues, but they may be closed/resolved with 1.5

James

**On 12/02/10 1:36 PM, Dan Connolly wrote:**

Hi,

We've got 2 or 3 i2b2 projects that seem to work fine in this respect, but in one of them, we've got the following problem:

1. Start i2b2 and select, say, ages 10-17 from the concept hierarchy and drag it to the query.

2. Hit Run Query

3. Check "Patient Set" in addition to "Patient Count"; click OK

4. Wait for the query to complete

5. In the Previous Queries box, expand the query, then expand the "Results of..."

6. Click the patient set in an attempt to expand it

expected result:

7e. A list of patient descriptions shows up actual result:

7a. The animated arrow image spins forever

We've spent quite some time with firebug and such trying to find the bug, but we're getting lost in the sdx routers and such. We don't really understand how it's supposed to work.

Suggestions, anyone? Has anyone else seen symptoms like this?

Perhaps we should use the i2b2 jira bug tracker to report this problem?

Oh... as to versions and such... all our i2b2 installations are running 1.4, tweaked to work with CAS.

Dan Connolly

---

#### **i2b2 Patient Demographics [11-16-2010]**

**On 11/16/10 8:13 AM, Shawn Murphy wrote:**

Hi Jacob,

Yes, the default setup for the web query satisfies the "obfuscated user" who is only able to see obfuscated counts. Try loading the timeline plug-in for the web client if it is not already loaded.

Thanks,

Shawn.

**On 11/15/10 5:45 PM, Jacob Weiser wrote:**

Hi All,

We installed I2B2 web client ver. 1.5 I am able to run queries and get back the count of patients but when trying to run the plugins - PDO request for patient detail data I only get back some xml script? Where would I see the actual patient detail? does the plugin cell need to be installed to have this functionality work?

Thanks

Jacob Weiser

---

#### **i2b2 v1.4 Tuning JBOSS etc [11-16-2010]**

**On 11/16/20 7:39 AM, Jeremy Nix wrote:**

Peter,

We're currently using the following java options in our production environment: "-server -Xms128m -Xmx3g -XX:PermSize=256m -XX:MaxPermSize=256m". These settings more or less help better support concurrent users running queries that result in large cohorts. We still need to do some load testing to find out what our limits are.

As far as performance. You will want to make sure that your tables have been properly indexed. A good way of finding out is to take the query string (generated\_sql column) from your i2b2demodata.qt\_query\_master table and review the execution plan for that query.

Something else you will want to do is make sure that your logging is not set to INFO or DEBUG in your jboss-log4j.xml file (in /opt/jboss/server/default/conf dir). You only want to enable those log levels if there is an issue you are trying to resolve.

You mentioned the crc.properties file. If you make PDO queries (queries that return the actual patient data, fact data, or visit data), then you do want to make sure that your edu.harvard.i2b2.crc.pdo.paging.observation.size is not set too high. This is a server side limit that prevents too many observation\_fact records from being returned (which can cause out of memory errors). These limits can also be controlled on the client side in the i2b2 request (by setting the min/max of each input\_option\_listType), though I'm not familiar enough with the i2b2 web client and Eclipse based client to know where these settings are configured.

Jeremy Nix

**On 11/15/2010 12:36 PM, Peter Beninato wrote:**

Hi,

Looking for some tips and hints on tuning an i2b2 installation.

Noticed in the PM documentation regarding changing run.conf for JVM setting ( -Xms512m Xmx1024m) which is how system is currently set. Are there any recommendations for setting this to higher amounts like 2gig?

What about tuning the edu.harvard.i2b2.crc.pdo.paging.observation.size in the etc/spring/crc.properties?

I run queries that will return counts in 80 seconds (40,000+ patients), but the **\*\*same\*\*** queries will at other times run for 15+ minutes, they appear to complete, but then the clock in the lower right hand window on the webclient, keeps running for several hundred seconds more.

I was thinking this might be related to writing out XML as part of the PDO?

Thanks for any tuning tips.

Peter Beninato

---

### C\_METADATAXML Column [11-10-2010]

**On 11/10/10 3:20 PM, Mike Mendis wrote:**

It is a new feature, which is part of 1.5, of course you can create it yourself. Take a look at the 1.5 Ontology Message Doc 4.1.3 Optional <metadatxml> content  
mike

**On 11/10/10 3:08 PM, Arvinder Choudhary wrote:**

Mike:

We are currently using 1.4, is there any way we can get this working?

Arvinder

**On 11/10/10 2:57 PM, Mike Mendis wrote:**

Arvinder,

If you are using 1.5, you can create this using the ontology edit terms, which is in the ontology view. This is in the eclipse client.

mike

**On 11/10/10 2:54 PM, Arvinder Choudhary wrote:**

Team:

We are currently loading the vital observations values and Lab data into i2b2 and C\_METADATAXML column that is in i2b2 table is to be populated. I see examples in the demo database. My question is how is this column populated. Is there any program that creates this xml script for this column  
Here is the Sample script I pulled from the demo database.

```
<?xml version="1.0"?>
<ValueMetadata>
<Version>3.02</Version>
<CreationDateTime>11/18/2005 14:53:45</CreationDateTime>
<TestID>APCO2</TestID>
<TestName>Arterial PCO2</TestName>
<DataType>PosInteger</DataType>
<CodeType>GRP</CodeType>
<Loinc>2019-8</Loinc>
<Flagstouse>HL</Flagstouse>
<Oktousevalues>The units in this set of tests are not equivalent, therefore numeric comparisons between the tests are not possible. </Oktousevalues>
<MaxStringLength></MaxStringLength>
<LowofLowValue></LowofLowValue>
<HighofLowValue></HighofLowValue>
<LowofHighValue></LowofHighValue>
<HighofHighValue></HighofHighValue>
<LowofToxicValue></LowofToxicValue>
<HighofToxicValue></HighofToxicValue>
<EnumValues></EnumValues>
<CommentsDeterminingExclusion>
<Com></Com>
</CommentsDeterminingExclusion>
<UnitValues>
<NormalUnits>mm</NormalUnits>
<NormalUnits>hg</NormalUnits>
<NormalUnits>mmhg</NormalUnits>
<NormalUnits>mmhg.</NormalUnits>
<NormalUnits>mm/hg</NormalUnits>
<EqualUnits>mm</EqualUnits>
<EqualUnits>hg</EqualUnits>
<EqualUnits>mmhg</EqualUnits>
<EqualUnits>mmhg.</EqualUnits>
<EqualUnits>mm/hg</EqualUnits>
<ExcludingUnits></ExcludingUnits>
<ConvertingUnits>
<Units></Units>
<MultiplyingFactor></MultiplyingFactor>
</ConvertingUnits>
</UnitValues>
<Analysis>
<Enums />
<Counts />
<New />
</Analysis>
</ValueMetadata>
```

Regards

Arvinder Choudhary

---

**New Support and Install Mailing List [11-09-2010]**

***On 11/09/10 4:14 PM, Mike Mendis wrote:***

With the growing install and supports questions on the AUG mailing list (a good thing), we have created a new and exciting mailing list located at

<http://groups.google.com/group/i2b2-install-help>

This will be the future home for any support and install questions, the current AUG mailing list, will be for AUG material.

The mailing list location and can also be found on the community wiki site

<http://community.i2b2.org>

Thanks

Mike Mendis

---

## i2b2 Lockout Logic [11-04-2010]

**On 11/04/10 10:49 AM, Shawn Murphy wrote:**

Hi Jeremy,

Many of the details of the algorithm (and the logic behind it) are nestled in the paper:

Murphy, S.N., Chueh, H. A Security Architecture for Query Tools Used to Access Large Biomedical Databases. Journal of the American Medical Informatics Association, Symposium Supplement 2002; 552-556.

We do look at the resulting number of patients and not the logic of the query for repeats. Clearly it is much simpler to look at the number, but I would also assert that looking at comparable logic in queries is an intractable problem when one considers all the ways that the same logical query could be formulated (and further ways that are coming down the pipeline, such as query in queries and patient/encounter set in queries). I had your same concern about arbitrary lockouts, but in the 8 years of implementing the algorithm at Partners Healthcare with 2500 users on the system, we have had less than 10 lockouts, and the majority were people blindly repeating the same query. This makes sense when one considers that most queries on a large database return between 1000-300,000 patients, so there is lots of opportunity for different numbers.

Thanks,  
Shawn.

**On 11/03/10 3:12 PM, Jeremy Nix wrote:**

I had some questions with respect to the obfuscated role user lockout logic. I'm looking at the query that gets generated and trying to determine exactly what the lockout criteria is. Let's start with the query (QueryResultInstanceSpringDao:251):

```
SELECT COUNT(r1.result_instance_id) result_count
FROM qt_query_result_instance r1
INNER JOIN qt_query_result_instance r2 ON r1.real_set_size = r2.real_set_size,
qt_query_instance qi
WHERE r1.start_date BETWEEN (sysdate - (30)) AND sysdate
AND r2.start_date BETWEEN (sysdate - (30)) AND sysdate
AND r1.result_type_id = (4) - PATIENT_COUNT_XML
AND r2.result_type_id = (4) - PATIENT_COUNT_XML
AND qi.user_id = ('demo') - Current user
AND qi.query_instance_id = r1.query_instance_id
AND qi.query_instance_id = r2.query_instance_id
AND r1.real_set_size = (0) - Patient count of current query
GROUP BY r1.real_set_size
HAVING COUNT(r1.result_instance_id) > (7)
```

This is formatted example of the oracle version using the defaults: a user will be locked out if they execute the same query 7 times within the last 30 days... (at least that is my interpretation as to what the requirement is). First off, I believe that the inner join is an unnecessary step. The qt\_query\_result\_instance record ends up joining in on itself. The following query produced the same results:

```
SELECT COUNT(r1.result_instance_id) result_count
FROM qt_query_result_instance r1,
qt_query_instance qi
WHERE r1.start_date BETWEEN (sysdate - (30)) AND sysdate
AND r1.result_type_id = (4) - PATIENT_COUNT_XML
AND qi.user_id = ('demo') - Current user
AND qi.query_instance_id = r1.query_instance_id
AND r1.real_set_size = (0) - Patient count of current query
GROUP BY r1.real_set_size
HAVING COUNT(r1.result_instance_id) > (7)
```

That's minor though. What I'm really trying to understand is why i2b2 is assuming that any query that has the same result set size and result type is the same. After all, if a given user is designated the task of coming up with patient cohorts for specific studies, there's a relatively good chance that they may get the same result count within 30 days (I do realize this is configurable), assuming they're running queries multiple times a day. We were worried about the case where running 7 queries that return no patients would lock out the user, but it appears that the application does handle the 0 patient count as a special case.

Do we think there would be benefit in comparing the queries more closely to see if they are actually the same query? One thought (although it has flaws of its own) is to compare the GENERATED\_SQL (or the REQUEST\_XML) columns to see if they used the same query criteria when running the query. As stated, it has its flaws. If you take the simple approach (as I probably would) and compare the GENERATED\_SQL directly, you miss the case where the users have constructed the same query using different panels to formulate the criteria.

A much more involved option would be to create a sql function that compared the 2 xml request documents for equality. For queries to be equal, each panel in 1 query master request xml would have to be equal (minus panel name and panel number) to a panel in a 2nd query master request xml. Again, this isn't impossible to do, just more involved.

Is there any interest in adopting any one of these 2 lockout lookup alternatives? We would be happy to develop and contribute back either solution. If not, are there plans to enhance this functionality in some other way in the future?

Thanks,

Jeremy Nix

---

## IRB-Friendly Architecture and Processes [11-03-2010]

**On 11/03/10 9:29 AM, Russ Waitman wrote:**

Brian,

I built off Keith's protocol and am attaching our IRB protocol which may help. We have slightly different challenges here because the hospital, clinics, and university (medical center) are separate organizations. In addition to the IRB protocol we had to negotiate a new legal agreement and develop an oversight process with the different organizations. You may not require that. We've established i2b2 for the participant registry since we already had that data. We're in the process of loading the hospital data from Epic and are beginning the discussion on how we'll do ETL from the clinics' IDX system.

Russ Waitman

**On 11/02/10 1:09 PM, Michael Kamerick wrote:**

I started a governance archive on the CTSA wiki a couple of years ago. It contains IRB documents from UCSF, OHSU, Cincinnati, and Kaiser, as well as supporting documents around data usage and access. If you are a CTSA institution, you should have access to it. Here is the link:

<https://www.ctsawiki.org/wiki/pages/viewpageattachments.action?pagelId=20152842>

For those who are not in the CTSA, I have attached the UCSF IRB document along with a few other docs.

At UCSF, like many sites, we provide a concierge service for identified data. Users can browse the de-identified data in i2b2, develop a cohort, and then request the identified version of the cohort from our staff. We check their IRB protocol and provide the data accordingly.

However, at UCSF we do take the process a significant step further, and that step has won us a lot of favor with our IRB and Medical Center IT organization.

We have developed a secure research portal we call MyResearch. It provides a remoted environment, somewhat like a Citrix installation (though we use a different product). Users access the i2b2 workbench, their data, and numerous applications through this portal. Data remains on our servers and applications run on the servers but display on the local workstation/laptop.

Any data set containing PHI that is extracted from our backend data bases (which feed i2b2), or straight from the major clinical systems in the Medical Center, can only be delivered to the user by depositing it in their MyResearch account. This workflow is mandated by a policy developed by our Privacy Office, our IRB (which we call CHR), Medical Center IT, and my group, which is part of the campus IT organization and the CTSA.

I have attached our IRB protocol document, a simplified schematic of our systems environment, and a short document explaining our policies on access to clinical data for research.

Michael Kamerick

**On 11/02/10 7:27 AM, Brian Bush wrote:**

Greetings – we at Virginia Commonwealth University are in the process of implementing an i2b2 instance, and would be interested to know what architectural and procedural approaches you are taking to satisfy patient privacy concerns, particularly from your IRB.

Keith at Cincinnati Children's posted an IRB Protocol document <<http://sites.google.com/a/i2b2aug.org/discussions/Home/release-discussions-postings/i2b2irb>> over a year ago, which was very helpful in explaining their approach (at least at the time), and I'd be very interested in hearing what others have done, particularly your process for handling "re-identification."

Many thanks in advance,

Brian J. Bush

---

## i2b2 Query by Value (ENUM) Example Needed [11-01-2010]

**On 11/1/10 11:32 AM, Peter Beninato wrote:**

Hmmm...

So perhaps this is an issue in v1.4

**On 11/1/10 11:21 AM, Jack London wrote:**

We use the web client Query & Analysis Tool, v1.5.

Jack

**On 11/1/10 11:18 AM, Peter Beninato wrote:**

Hi,

Do you use the thick client or the web client?

And what version of i2b2 are you running?

Thanks.

Peter

**On 11/1/10 10:21 AM, Jack London wrote:**

Peter,

I was able to get this to work for our biospecimen ontology where we wanted to use query-by-value to select either "malignant" or "normal" tissue (or both). (BTW, we will probably use the cd\_modifier functionality in v1.6 in the future to accomplish this.) Below is the C\_METADATAXML field in the metadata table for the biospecimens:

```
=====
<?xml version="1.0"?>
<ValueMetadata><Version>3.02</Version>
<CreationDateTime>05/19/2010 04:01:06</CreationDateTime>
<TestID>Specimen Path Status</TestID>
<TestName>Specimen Path Status</TestName>
<DataType>Enum</DataType>
<CodeType>GRP</CodeType>
<Loinc>No LOINC</Loinc>
<Flagstouse></Flagstouse>
<Oktousevalues>Y</Oktousevalues>
```



```

<MaxStringLength></MaxStringLength>
<LowofLowValue></LowofLowValue>
<HighofLowValue></HighofLowValue>
<LowofHighValue></LowofHighValue>
<HighofHighValue></HighofHighValue>
<LowofToxicValue></LowofToxicValue>
<HighofToxicValue></HighofToxicValue>
<EnumValues>
<Val description="">Normal</Val>
<Val description="">Malignant</Val>
</EnumValues>
<CommentsDeterminingExclusion>
<Com></Com>
</CommentsDeterminingExclusion>
<UnitValues>
<NormalUnits></NormalUnits>
</UnitValues>
<Analysis>
<Enums /><Counts /><New />
</Analysis>
</ValueMetadata>

```

=====

In the OBSERVATION\_FACT table, the VALTYPE\_CD = "T" and TVAL\_CHAR = "Normal" or "Malignant" (I also set MODIFIER\_CD = "Normal" or "Malignant" but I do not think this is in any way necessary.)

Jack

**On 10/29/10 6:18 PM, Peter Beninato wrote:**

Hi,  
 Working on trying to allow a Query-by-Value to work for Text Values.  
 I'm using the demo and I can't say I have ever seen it work. (I have query-by NUMERIC Value working)  
 I have noticed the use of Enum Values in the ontology c\_metadatxml field.

By way of example I'm using  
 i2b2.c\_basecode = 'LOINC:9317-9'

Or  
 select \*  
 from i2b2  
 where c\_basecode LIKE 'LOINC:9317-9%'  
 In the c\_metadatxml there is a datatype of ENUM as in <DataType>Enum</DataType>

I have used Query-by-Value before with "NUMBER" data. In those cases the DataType is set to "PosFloat" and in the observation the valtype\_cd is set to 'N'.

In the example for = 'LOINC:9317-9', The valtype\_cd is set to 'T' and the XML Datatype is "ENUM"  
 However, when I run for different values that are shown in the ENUM like 'NL', 'NORMAL', these values **\*\*ARE NOT\*\*** included in the SQL found in qt\_query\_master, so it is like no filter is being applied.

What does the XML need to be to get a text/enum query-by-value to work? How do the fields in observation\_fact need to be set?

Is there a "LAB" in the demo that does work for ENUM values?

Thanks.

#### LDAP Integration i2b2 v1.4-v1.6 [11-01-2010]

**On 11/1/10 9:08 AM, Dan Connolly wrote:**

Regarding checking for training, we're also planning to do that.  
 We plan to do it largely external to i2b2, and only create i2b2 accounts once the check has been done.  
 Since there are at least two of us interested in this feature, I added a request for it in JIRA:  
 check for HIPAA training etc. before granting access to i2b2 projects  
<https://community.i2b2.org/jira/browse/LEOPARD-48>

**On 10/30/10 2:40 PM, Robert Schuff wrote:**

thanks Phillip.

**On 10/29/10 5:04 PM, Phillip Reeder wrote:**

I looked at the authentication piece at one point. If memory serves me correct, the PM cell uses the jcifs library for AD authentication. According to the documentation for jcifs, it can also be used for LDAP so I think it should just be a matter of configuring it to use your ldap server but I didn't actually attempt it.

I think Mike Mendis had sent me some information about the AD configuration settings.

Phillip

**On 10/29/10 5:42 PM, Robert Schuff wrote:**

Hi Folks,

We are beginning to think about using our institutional LDAP system for authentication to i2b2 thin client. I saw that in 1.4 Active directory support was supposed to be enabled and documentation was forthcoming but I'm unable to find anything on the community wiki about it. My two questions are:

1. How does one enable LDAP in i2b2? Are there differences between 1.4 and 1.6 for doing this?

1. In addition to authenticating against LDAP, we need to also check a web service for investigators' having taken various HIPAA and RCR training. What mechanisms are available for us to hook into the authentication process to accomplish this?

thanks  
Rob

---

### Concepts for Nursing Observations and Fundamental Vital Signs [10-28-2010]

**On 10/28/10 10:08 AM, Brian Ostasiewski wrote:**

We plan to address measurement location/type via concept modifier\_cd values. Concepts like high/low blood pressure or under/overweight are typically first-line filtering our researchers want to do on patient sets.

Brian

**On 10/27/10 1:31 PM, Russ Waitman wrote:**

Brian,

Thanks for the reply. So you're providing definitions for these observations. I was thinking less at interpretation of the numeric result and more at the level of organizing the type of observation.

We're looking to see if we could load everything out of the EPIC IP\_FLWSHT\_MEAS table (the core inpatient vitals and nursing observations) and map them to some combination of LOINC or SNOMED. We'd then be interested if there are existing hierarchical ontologies for this kind of data similar to lab results.

I'd imagine they might be sorted like

Vital Signs

Blood Pressure

NIBP non-invasive blood pressure

ABP arterial blood pressure

CVP central venous Pressure

Respiratory Rate

Nursing Assessments

Pain

Skin

Braden Score nutrition intake patter

Braden Score physical activity

Glasgow Coma Scale.

eye opening

etc...

Judy Warren, nursing informaticist, is also going to check with the LOINC nursing people to see if this structure already exists.

Russ

**On 10/26/10 2:53 PM, Brian Ostasiewski wrote:**

We incorporated some basic vitals with the structure:

Blood Pressure

Normal

Diastolic < 80

Systolic < 120

Prehypertension

Diastolic 80-89

Systolic 120-139

Mild Hypertension

Diastolic 90-99

Systolic 140-159

Severe Hypertension

Diastolic >99

Systolic < 159

Body Mass Index

Underweight (BMI < 18.5)

Severe Thinness <16

Moderate Thinness 16-17

Mild Thinness 17-18.5

Normal Weight (BMI 18.5-24.99)

Low 18.5-23

High 23-25

Overweight (BMI >= 25.0)

Pre Obese

25-27.5

27.5-30

Obese

Class I

Low 30-32.5

High 32.5-35

Class II

Low 35-37.5

High 37.5-40

Class III > 40

Temperature

<= 95

96

97

98

99

100

101

102

103

104

>= 105

Brian

***On 10/26/10 3:33 PM, Russ Waitman wrote:***

Hi all,

Arvinder, Dan and I are now wading through our Epic data. We are working through demographics, labs, meds but also want to get started on basic nursing observations and vital signs (say blood pressure, respiratory rate, pressure ulcer data).

What are others approaches? Is there a best practice vital sign ontology that's being used? I guess ideally there would be a subset derived from the UMLS or bioontology.org but we're looking for anything as a starting point.

Russ Waitman

---

#### **Terminology Mapping Tool [10-22-2010]**

***On 10/22/10 7:59 AM, Nigam Shah wrote:***

Hi Phillip,

Have you considered uploading your researcher's terminology into BioPortal? (<http://bioportal.bioontology.org>). It will get mapped to existing reference terminologies (such as ICD9, SNOMEDCT etc) via lexical methods described in this paper (<http://www.ncbi.nlm.nih.gov/pubmed/20351849>).

The cross-maps are then available for access via REST services (and very soon the researcher's custom terminology will also be importable into i2b2 directly from BioPortal).

Regards,

Nigam.

***On 9/16/10 9:42 AM, Phillip Reeder wrote:***

What is the preferred i2b2 terminology mapping tool? Basically, we will be loading a representation of several researchers' non-standardized data collection forms into i2b2. Then, we need to create a map from the researcher's terminology to a common terminology. What tools have you used and what would you recommend for others to adopt?

Thanks,

Phillip Reeder

---

#### **Social Security Number [10-22-2010]**

***On 10/22/10 7:01 AM, Shawn Murphy wrote:***

Hi Arvinder,

The social security number can be stored in the patient mapping table as another type of encrypted patient identifier.

Thanks,

Shawn.

***On 10/21/10 4:42 PM, Arvinder Choudhary wrote:***

Mike:

How and where you all storing the social security number of the patients in i2b2 application in order to join this with the social security death index and also to collaborate with other databases outside your organization.

Regards

Arvinder

---

#### **Sharing Queries, etc., via the Workplace Cell [10-21-2010]**

***On 10/21/10 9:25 AM, Mike Mendis wrote:***

Take a look at the workplace\_access table in the workplace schema.

You should have a row with a c\_name of 'SHARED' and c\_group\_id with your project id.

If not, then add one by modifying the following insert statement:

```
INSERT INTO WORKPLACE_ACCESS(C_TABLE_CD, C_TABLE_NAME, C_PROTECTED_ACCESS, C_HLEVEL, C_NAME, C_USER_ID,
C_GROUP_ID, C_SHARE_ID, C_INDEX, C_PARENT_INDEX, C_VISUALATTRIBUTES, C_TOOLTIP)
VALUES('demo', 'WORKPLACE','N', 0, 'SHARED', 'shared', 'demo', 'Y', 100, NULL, 'CA', 'SHARED');
```

mike

***On 10/20/10 2:07 PM, Dan Connolly wrote:***

I don't see a 'shared' folder. I'm using the 1.4 web client... should it work there?

**On 10/20/10 7:50 AM, Mike Mendis wrote:**

Yes there is a view/cell called 'workplace'. If you are a manager or admin for that project than you will be able to see all the queries in the workplace for all the users in that project.

In the workplace there is a folder called 'shared', if your colleague moves their query into the shared folder then any user in the project can see that query.  
mike

**On 10/19/10 4:55 PM, Dan Connolly wrote:**

Can I share workspaces with other i2b2 users?

I see something about managers being able to see workspaces of other project users in the documentation, but I don't quite understand how to use it. What I'm actually trying to do is debug a problem with the timeline view.

A colleague made a query and a couple patient sets. Re-creating the query is a little tedious, so I'd like to just re-use it.

---

## Removing a Project [10-21-2010]

**On 10/21/10 9/20 AM, Mike Mendis wrote:**

Michael,

I am wondering if that project somehow got corrupt and that id is not being sent correctly to the backend. But the easy hack way is to log onto the database where the pm resides. Look at the pm-ds.xml in the deploy directory of jboss for login info. Then locate the Pm\_project\_table table, and find the project you want to delete. Might be safer to just remove it, you can change the status\_cd to 'D' also. This states it is deleted.

When you locate this project, can you send me what was filled out, maybe I can figure out why it would not delete.

Thanks

mike

**On 10/20/10 1:38 PM, Michael Ford wrote:**

I am just starting to work with the product, please forgive if this is not the correct forum for this and a pointer to where I ask such questions would be helpful.

I am trying to remove a project from my I2B2 configuration (Demo2)

It was 1/2 created by someone before I began to work on the project.

When I try and delete the site I receive the following errors in the log and the site still shows. (Release 1.4)

Frontend error

```
<ns2:response xmlns:tns="http://ws.pm.i2b2.harvard.edu"
xmlns:ns3="http://www.i2b2.org/xsd/hive/msg/version/"
xmlns:ns4="http://www.i2b2.org/xsd/cell/pm/1.1/"
xmlns:ns2="http://www.i2b2.org/xsd/hive/msg/1.1/">
<message_header> <i2b2_version_compatible>1.1</i2b2_version_compatible>
<hl7_version_compatible>2.4</hl7_version_compatible>
<sending_application> <application_name>PM Cell</application_name>
<application_version>1.4</application_version>
</sending_application> <sending_facility>
<facility_name>i2b2 Hive</facility_name> </sending_facility>
<receiving_application> <application_name>PM Cell</application_name>
<application_version>1.4</application_version>
</receiving_application> <receiving_facility>
<facility_name>i2b2 Hive</facility_name>
</receiving_facility> <datetime_of_message>2010-10-20T08:43:24.881-07:00</datetime_of_message>
<message_control_id> <message_num>8uVB8DF3v8W94DBOKn0Ao</message_num>
<instance_num>1</instance_num>
</message_control_id> <processing_id> <processing_id>P</processing_id> <processing_mode>I</processing_mode>
</processing_id> <accept_acknowledgement_type>AL</accept_acknowledgement_type>
<application_acknowledgement_type>AL</application_acknowledgement_type> <country_code>US</country_code>
<project_id>Demo</project_id> </message_header> <response_header> <result_status>
<status type="ERROR">Project not updated, does it exist?</status> </result_status> </response_header> </ns2:response>
[Break on this error] $('pmAdmin-projStatus').value = 'A';\r\n
```

Back end log error

```
2010-10-20 09:32:50,671 ERROR [edu.harvard.i2b2.pm.delegate.ServicesHandler] Project not updated, does it exist?
2010-10-20 09:32:50,672 ERROR [STDERR] edu.harvard.i2b2.common.exception.I2B2DAOException: Project not updated, does it exist?
2010-10-20 09:32:50,672 ERROR [STDERR] at edu.harvard.i2b2.pm.dao.PMDbDao.deleteProject(Unknown Source)
2010-10-20 09:32:50,672 ERROR [STDERR] at edu.harvard.i2b2.pm.delegate.ServicesHandler.runDeleteProject(Unknown Source)
2010-10-20 09:32:50,673 ERROR [STDERR] at edu.harvard.i2b2.pm.delegate.ServicesHandler.execute(Unknown Source)
2010-10-20 09:32:50,673 ERROR [STDERR] at edu.harvard.i2b2.pm.ws.ExecutorRunnable.run(Unknown Source)
2010-10-20 09:32:50,673 ERROR [STDERR] at java.lang.Thread.run(Thread.java:595)
2010-10-20 09:32:50,977 DEBUG [edu.harvard.i2b2.pm.ws.PMService] my pm repsonse is: <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:response xmlns:ns4="http://www.i2b2.org/xsd/cell/pm/1.1/" xmlns:ns3="http://www.i2b2.org/xsd/hive/msg/version/" xmlns:ns2="http://www.i2b2.org
/xsd/hive/msg/1.1/">
<message_header>
<i2b2_version_compatible>1.1</i2b2_version_compatible>
<hl7_version_compatible>2.4</hl7_version_compatible>
<sending_application>
<application_name>PM Cell</application_name>
<application_version>1.4</application_version>
</sending_application>
<sending_facility>
<facility_name>i2b2 Hive</facility_name>
</sending_facility>
<receiving_application>
```

```

<application_name>PM Cell</application_name>
<application_version>1.4</application_version>
</receiving_application>
<receiving_facility>
<facility_name>i2b2 Hive</facility_name>
</receiving_facility>
<datetime_of_message>2010-10-20T09:32:50.674-07:00</datetime_of_message>
<message_control_id>
<message_num>8aNGWiKieTT1WamY6XU5Z</message_num>
<instance_num>1</instance_num>
</message_control_id>
<processing_id>
<processing_id>P</processing_id>
<processing_mode>I</processing_mode>
</processing_id>
<accept_acknowledgement_type>AL</accept_acknowledgement_type>
<application_acknowledgement_type>AL</application_acknowledgement_type>
<country_code>US</country_code>
<project_id>Demo</project_id>
</message_header>
<response_header>
<result_status>
<status type="ERROR">Project not updated, does it exist?</status>
</result_status>
</response_header>
</ns2:response>
2010-10-20 09:32:50,978 DEBUG [edu.harvard.i2b2.pm.ws.PMService] my return is: <ns2:response xmlns:ns3="http://www.i2b2.org/xsd/hive/msg/version
/" xmlns:ns4="http://www.i2b2.org/xsd/cell/pm/1.1/" xmlns:ns2="http://www.i2b2.org/xsd/hive/msg/1.1/">
<message_header>
<i2b2_version_compatible>1.1</i2b2_version_compatible>
<hl7_version_compatible>2.4</hl7_version_compatible>
<sending_application>
<application_name>PM Cell</application_name>
<application_version>1.4</application_version>
</sending_application>
<sending_facility>
<facility_name>i2b2 Hive</facility_name>
</sending_facility>
<receiving_application>
<application_name>PM Cell</application_name>
<application_version>1.4</application_version>
</receiving_application>
<receiving_facility>
<facility_name>i2b2 Hive</facility_name>
</receiving_facility>
<datetime_of_message>2010-10-20T09:32:50.674-07:00</datetime_of_message>
<message_control_id>
<message_num>8aNGWiKieTT1WamY6XU5Z</message_num>
<instance_num>1</instance_num>
</message_control_id>
<processing_id>
<processing_id>P</processing_id>
<processing_mode>I</processing_mode>
</processing_id>
<accept_acknowledgement_type>AL</accept_acknowledgement_type>
<application_acknowledgement_type>AL</application_acknowledgement_type>
<country_code>US</country_code>
<project_id>Demo</project_id>
</message_header>
<response_header>
<result_status>
<status type="ERROR">Project not updated, does it exist?</status>
</result_status>
</response_header>
</ns2:response>
2010-10-20 09:33:01,200 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] run: [id=1285950320162,target=[target=jboss.j2ee:service=EJB3,ear=QP1.ear,jar=QP-
An-EJB.jar,name=CronEjb],remaining=-2380298,periode=60000,active]
2010-10-20 09:33:01,200 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] run: [id=1285950320163,target=[target=jboss.j2ee:service=EJB3,ear=QP1.ear,jar=QP-
An-EJB.jar,name=LargeCronEjb],remaining=-2380229,periode=60000,active]
2010-10-20 09:33:01,200 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] setTimerState: in_timeout
2010-10-20 09:33:01,200 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] setTimerState: in_timeout
2010-10-20 09:33:01,201 DEBUG [org.springframework.jdbc.core.JdbcTemplate] Executing SQL query [ select * from i2b2hive.CRC_ANALYSIS_JOB
where queue_name = ? and status_type_id = ? order by create_date]
2010-10-20 09:33:01,201 DEBUG [org.springframework.jdbc.datasource.DataSourceUtils] Fetching JDBC Connection from DataSource
2010-10-20 09:33:01,201 DEBUG [org.springframework.jdbc.core.JdbcTemplate] Executing SQL query [ select * from i2b2hive.CRC_ANALYSIS_JOB
where queue_name = ? and status_type_id = ? order by create_date]
2010-10-20 09:33:01,201 DEBUG [org.springframework.jdbc.datasource.DataSourceUtils] Fetching JDBC Connection from DataSource
2010-10-20 09:33:01,201 DEBUG [org.springframework.jdbc.core.StatementCreatorUtils] Setting SQL statement parameter value: column index 1,
parameter value [LARGE_QUEUE], value class [java.lang.String], SQL type unknown

```

2010-10-20 09:33:01,201 DEBUG [org.springframework.jdbc.core.StatementCreatorUtils] Setting SQL statement parameter value: column index 1, parameter value [MEDIUM\_QUEUE], value class [java.lang.String], SQL type unknown  
2010-10-20 09:33:01,202 DEBUG [org.springframework.jdbc.core.StatementCreatorUtils] Setting SQL statement parameter value: column index 2, parameter value [1], value class [java.lang.Integer], SQL type unknown  
2010-10-20 09:33:01,202 DEBUG [org.springframework.jdbc.core.StatementCreatorUtils] Setting SQL statement parameter value: column index 2, parameter value [1], value class [java.lang.Integer], SQL type unknown  
2010-10-20 09:33:01,204 DEBUG [org.springframework.jdbc.datasource.DataSourceUtils] Returning JDBC Connection to DataSource  
2010-10-20 09:33:01,204 DEBUG [org.springframework.jdbc.datasource.DataSourceUtils] Returning JDBC Connection to DataSource  
2010-10-20 09:33:01,204 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] Timer was not registered with Tx, resetting state: [id=1285950320162,target=[target=jboss.j2ee:service=EJB3,ear=QP1.ear,jar=QP-An-EJB.jar,name=CronEjb],remaining=-2320302,periode=60000,in\_timeout]  
2010-10-20 09:33:01,204 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] setTimerState: active  
2010-10-20 09:33:01,204 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] Timer was not registered with Tx, resetting state: [id=1285950320163,target=[target=jboss.j2ee:service=EJB3,ear=QP1.ear,jar=QP-An-EJB.jar,name=LargeCronEjb],remaining=-2320233,periode=60000,in\_timeout]  
2010-10-20 09:33:01,204 DEBUG [org.jboss.ejb.tx.timer.TimerImpl] setTimerState: active  
2010-10-20 09:33:02,240 DEBUG [org.jboss.resource.connectionmanager.IdleRemover] run: IdleRemover notifying pools, interval: 30000

Michael C Ford

---

### Problems unzipping 1.6-RC1 VMWare [10-14-2010]

**On 10/14/10 2:54 PM, Jack London wrote:**

Thank you Mike!

Installed p7zip – all is well. I'm looking forward to trying out the cd modifiers with our biospecimen data.

Jack

**On 10/14/10 2:29 PM, Mike Mendis wrote:**

It is because this file is greater than 4gb. But I only experienced this on Windows extractor. Installing a third party one such as 7-zip resolved it on windows. I am going to d/l and try on a mac to see. But for linux you can install p7zip, which should resolve it

Centos/redhat

sudo yum install p7zip-full

7za e filename.zip

Ubuntu

sudo aptitude install p7zip-full

7za e filename.zip

**On 10/14/10 2:04 PM, Jack London wrote:**

Have problem unzipping VMWare zip file on both Mac and Linux. Get message:

skipping: i2b2 Server/i2b2 Server-flat.vmdk need PK compat. v4.5 (can do v2.1)

Suggestions?

Jack

---

### NLP Cell Installation Advice [09-30-2010]

**On 9/30/10 12:45 PM, Michael Bobak wrote:**

Yes, we were very careful to get exactly the suggested versions of Tomcat, Axis2, Gate/Weka etc. One thing that I wasn't sure about for the common/lib jars, was if I should replace the default jasper-compiler-jdt.jar (given the jasper errors we got), so I looked & followed a suggestion to use it from a 3rd party Gate lib (text2onto) but not much change\*. Right now I've been asked to run most of the new 1.5.2 server from the VMware image; and we have been trying to get the nlp-cell on either os-x or another Linux image, as I don't think we will have use of a Windows box as a server.

Thanks you,

Michael

**On 9/30/10 12:29 PM, Michael Bobak wrote:**

Thank you! It was suggested that I build axis2 from source, but am having trouble with that at the moment. Another stumbling block, is during the load of our test cases: The text doesn't fit into the tval\_char column of the Observation\_fact table, and so far I have been unable to change the column type via the oracle web-interface.

Thanks again, Mike

**On 9/30/10 10:54 AM, Marius Petruc wrote:**

Michael,

in our setup we use IIS7 (not Tomcat). so to get the NLP cell installed without increasing complexity we decided to install NLP on the already working jboss + axis2, together with the other i2b2 cells. i followed the installation instructions provided.

%CATALINA\_HOME%\common\lib = [jboss\_root\_folder]\server\default\lib. i tired using a newer version of gate, but it didn't work. you need to use the exact version of gate (GATE-3.1.exe) and weka (Weka 3.4.4). in windows i had to install a few other libraries that were missing (pdfbox is an example that I remember), on linux you might or might not need to. use the log file to figure out if/what's missing and where it needs to go.

i thought i have made notes, but can't find them at the moment (i'm running a grep search right now). if i find anything i'll let you know.

yours,

Marius

**On 9/30/10 8:39 AM, Shawn Murphy wrote:**

Hi Mike,

This cell was built by Qing Zeng and Sergey Goryachev (both cd'd).

Thanks,

Shawn.

**On 9/29/10 12:51 AM, Michael Bobak wrote:**

Recently we have been going through NLP-installation-10.pdf to set up the cell either on os-x, the VMware image that I am running off of, or another Linux image. We have had some (jasper related axis) problems even with the exact versions of axis2&tomcat listed in the doc. If you have dealt with something similar, or have any other advice that might speed up the process of being able to try out the nlp-cell, we would be most grateful.

Thanks!

mike

---

#### **Plug-in Packaging Guidance [09-29-2010]**

**On 9/29/10 12:34 PM, Shawn Murphy wrote:**

Hi Bill,

There is a description of how to package and deploy a Workbench plug-in in the document "i2b2Workbench\_Developers Guide. It is on the web site under the documentation hexagon for the i2b2 Workbench. It's a big file, so I will send it to you separately so that it does not clog everyone's mailboxes.

Thanks,

Shawn.

**On 9/23/10 12:12 PM, Bill Adams wrote:**

Hello,

We have developed a new i2b2 workbench plug-in that we will be demonstrating at the upcoming CTSA informatics meeting. We are looking for help/guidance related to packaging and deploying our plug-in. Has anyone done this who could offer some help/guidance?

Bill Adams

---

#### **Temporal Query [09-28-2010]**

**On 9/28/10 9:39 AM, Bill Adams wrote:**

At BU we are working on a plug-in as well. We will be presenting it on October 14 at the CTSA informatics meeting. The plug-in uses temporal- and value- and age-constraints to monitor and assess health outcomes in i2b2. Once it is out of prototype stage we will be happy to share and improve within the i2b2 AUG.

Bill Adams

**On 9/28/10 8:39 AM, Daniele Segagni wrote:**

Hello David,

Here in Pavia, Italy, we're working on temporal query to automatically discover interesting regularities or relationships in our i2b2 data.

Currently we're planning the development work in order to create a dedicated temporal analysis i2b2 cell and a plug-in for the i2b2 web client.

Since we are in the earliest part of our project we can collaborate on this problem to find a good solution that could be used by both.

Regards

Daniele Segagni

**On 9/27/10 3:05 PM, Daniel Nigrin wrote:**

I'm sure Griffin will chime in, but yes, he and I and others here in Boston are working on this. At the core, we are implementing concepts that me and Zak Kohane described in this JAMIA article from many moons ago:

<http://jamia.bmj.com/content/7/2/152.abstract>

Best,

Dan

**On 9/27/10 2:58 PM, Keith Marsolo wrote:**

David,

Griffin Weber has done some work related to this. He presented it at the AUG meeting held in conjunction with the AMIA Summit on Translational Bioinformatics back in March. He may have some updates.

Keith

**On 9/27/10 2:06 PM, David Norris wrote:**

Temporal query was cited as weakness of i2b2 in the review by Deshmukh, Meystre and Mitchell in BMC Medical Research Methodology 2009, 9:70. I find that moderately complex temporal queries in my own recent clinical research would have been impossible to implement in the current Workbench.

It would appear that a Temporal Query plug-in for the Workbench could make a substantial contribution to i2b2. Has anyone explored this possibility?

Would anyone be interested to collaborate on this problem?

David Norris

---

#### **ICD9 Ontology Missing Terms Export of NIH/CDC ICD9 Hierarchy [09-21-2010]**

**On 9/21/10 7:10 AM, Shawn Murphy wrote:**

Thanks Philip, that would be great, there are no legal entanglements ☺.

Shawn.

**On 9/20/10 5:40 PM, Phillip Reeder wrote:**

Is there any legal reason why i2b2 is not distributing the full hierarchy? I have created the hierarchy from the UMLS and would be willing to share it with any and all. Perhaps it could be added to the demodata?

Phillip Reeder

**On 9/20/10 9:41 AM, Mike Mendis wrote:**

Peter,

That is correct, the i2b2 demodata ICD9 Diagnoses is not a complete ICD9.

Here is a link to the a access database that contain a list of the ICD9 codes [http://www.cms.hhs.gov/MinimumDataSets20/07\\_RAVENSoftware.asp](http://www.cms.hhs.gov/MinimumDataSets20/07_RAVENSoftware.asp) <[http://www.cms.hhs.gov/MinimumDataSets20/07\\_RAVENSoftware.asp](http://www.cms.hhs.gov/MinimumDataSets20/07_RAVENSoftware.asp)> from the United States Department of Health and Human Services.

I have not used or tried it, but you should be able to create a compatible i2b2 metadata from it.

**On 9/17/10 2:31 PM, Peter Beninato wrote:**

Hi,

Is there a current ICD9 "Diagnoses" ontology available?

It seem there are gaps in demo "Diagnoses."

I have loaded observation\_fact and have approximately 820 codes that don't map to concept\_dimension/i2b2 table(s) (see after signature for listing).

There could be more as this was run against only a three-month dataset.

Does anyone have a CSV, tab-delimited, oracle dmp file of the NIH/CDC ICD9 hierarchy? Or even better an "i2b2" compliant version?

Thanks.

Peter Beninato

---

#### Best Practices for EPIC Clarity – i2b2 ETL Scripts [09-19-2010]

**On 9/19/10 12:03 PM, Keith Marsolo wrote:**

All,

We've posted some documentation that describes our work with the Epic Clarity database. It can be found at the following location:

<https://bmi.cchmc.org/svn/i2b2/i2b2/public/data/Documents/>

The first document: "Moving Data from Epic to I2B2.docx" describes, well, our approach to moving data from Epic to i2b2. It's not complete, but it provides an overview of how we load demographics, diagnoses, and medications. There's some information on how we create metadata XML for each content type. This is similar to the upcoming "modifier" functionality in version 1.6. Until we release our code, it's more or less CCHMC-specific, but the content would be applicable to either approach.

The other document "epic\_dw\_master\_tables.pdf" details some of the steps used to create our "Master" Epic datamart that we provide to users for reporting and other purposes. Not all of the tables are included, but we've provided a few of the more frequently used ones. The scripts provide an overview of what columns/tables to use to look for certain content.

Words of "wisdom":

1. The tables/columns in Epic may have misleading names. Always check the Epic UserWeb/Clarity data dictionaries to determine the true purpose of the field.

2. Never assume that the Epic documentation is accurate and/or complete. Each institutions Epic implementation is different, and data may be in unique locations, particularly if that data is fed by an interface. To verify the data you are working with is correct (i.e. what you think it means), I would encourage you to work with your Clarity Reporting Team or end users to ensure that the data in the database matches what appears on the screen.

Thanks.

Keith

**On 9/19/10 1:03 PM, Keith Marsolo wrote:**

Russ,

We have scripts that we can share. We will send by the end of the week.

Keith

**On 9/15/10 12:56 PM, Russ Waitman wrote:**

Hi,

Thank you all again for the vast help on adopting i2b2. We've completed our business agreements with the affiliated hospital and clinic and data sharing and are looking to establish the ETL processes.

Most of our clinics use IDX for billing and registration. Arvinder developed an ETL process for an existing subset of that data that we receive (demographics, diagnoses, procedures). He also worked with several members to load CPT from the UMLS. That's been the beta environment we've been working with to bring up the system and work on security integration. We'll now start building out more clinic encounter information.

We can now start on EPIC. It is almost deployed in our inpatient and ED environments and they are starting to rollout throughout the clinics. Specifically we're going to follow many of you that seem to start by building off a backup extract of their Clarity database. We've received a copy of the "Build/Test" environment for development and are in the process of moving over the production data.

Are people sharing best practices or scripts for how you're getting from things like the PATIENT, PAT\_ENC... tables in Clarity into the i2b2 schema for patient, visit, and observer dimension? We'd like to not reinvent the wheel or at least avoid common mistakes.

Sincerely,

Russ Waitman

---

#### Web Client Issue [09-17-2010]

**On 9/17/10 8:39 AM, Mike Mendis wrote:**

What is the list you have in the <http://localhost:9090/i2b2/services/listServices>? Also this is a new install? If so than you completed the crc install doc? mike

**On 9/17/10 8:15 AM, Jeremy Nix wrote:**

Are you trying to pull up the webclient or test out the CRC Cell webservice? The url for the webclient is <http://localhost:80/index.php>.

Jeremy Nix

**On 09/16/2010 04:45 PM, Maggie Qiu wrote:**

Hi, All:

We had 1.5.1 installed and it works fine with fat client workbench. We can see all services are up at <http://localhost:9090/i2b2/services/listServices>. When



we try to launch web client <http://localhost:9090/i2b2/rest/QueryToolService>, we got page unavailable error. Can someone help? Thanks a lot.  
Maggie Qiu

---

#### **i2b2 Hive Account Locked [09-15-2010]**

**On 9/15/10 5:18 PM, Dan Connolly wrote:**

Whew... finally figured it out... it wasn't that my new instance was broken, but I didn't kill my old instance, and the old instance was trying to use old passwords.

Thanks for the various bits of help, everybody.

Dan

**On 9/14/10 12:52 PM, Dan Connolly wrote:**

Really? Regardless of name? Interesting. But no, I don't see any extra files:

```
jboss-4.2.2.GA/server/default/deploy> grep i2b2hive *  
crc-jms-ds.xml: <user-name>i2b2hive</user-name>  
ont-ds.xml: <user-name>i2b2hive</user-name>  
work-ds.xml: <user-name>i2b2hive</user-name>
```

And the other one? How often is that i2b2hive.CRC\_ANALYSIS\_JOB query supposed to run? It seems to run once per minute here.

**On 9/14/10 12:22 PM, Mike Mendis wrote:**

Do you have any backup files in the deploy directory. Jboss will deploy and use any file in the deploy regardless of what it is name. So if you have a file called crc-ds.xml and one called crc-ds.xml-orig, it will try to deploy both, and the orig might have a old user/pass which after deploying a couple of times will cause a lockout on the oracle side.

The message about the timing already exists is fine.

**On 9/14/10 12:03 PM, Dan Connolly wrote:**

Looks like that wasn't the heart of the problem. The symptoms persist.

I feel like I'm losing my mind trying to figure this out; it all worked until about a week ago, and I'm not sure what I did to break it...

**On 9/13/10 1:27 PM, Dan Connolly wrote:**

Looks like our client and servers didn't agree on the i2b2 hive password, and once that i2b2hive.CRC\_ANALYSIS\_JOB ran enough times, blammo, the account got locked.

I was misled by a goofy timezone configuration on our database server.

**On 9/13/10 12:08 PM, Dan Connolly wrote:**

Our i2b2 hive keeps getting stuck because the i2b2hive oracle account keeps getting locked out.

I can fix the problem, for a while, by doing:

```
alter user i2b2hive account unlock;
```

But the problem seems to come up at some random time later.

In the most recent case, Oracle gives a LOCK\_DATE of 10-SEP-2010 01:49:09

Looking in the logs, I see that i2b2hive does this once an hour (though it generates 2 DEBUG messages):

```
2010-09-10 01:49:58,545 DEBUG [org.springframework.jdbc.core.JdbcTemplate] Executing SQL query [ select * from i2b2hive.CRC_ANALYSIS_JOB  
where queue_name = ? and status_type_id = ? order by create_date]
```

I can't see why that would cause the account to get locked.

I see this error from time to time, but I don't understand how it would cause the i2b2hive account to get locked either:

```
2010-09-09 08:43:52,758 DEBUG [org.jboss.mq.pm.jdbc2.PersistenceManager] Could not create table with SQL: CREATE TABLE JMS_MESSAGES (  
MESSAGEID INTEGER NOT NULL  
, DESTINATION VARCHAR(255) NOT NULL, TXID INTEGER, TXOP CHAR(1), MESSAGEBLOB BLOB, PRIMARY KEY (MESSAGEID, DESTINATION) )  
java.sql.SQLException: ORA-00955: name is already used by an existing object
```

Help? Any ideas? We have a demo early tomorrow morning, so it would be really great to get to the bottom of this today.

Dan Connolly

---

#### **i2b2 Source, Ant and Eclipse [09-15-2010]**

**On 9/15/10 11:11 AM, Jeremy Nix wrote:**

I've picked up the knowledge by experimentation and previous knowledge of the technologies used. I believe there is a master script that can be run which walks through each of the projects and executes the dist target, which has the effect of executing the jaxb\_gen and generate-ejb targets. I can't remember where that ant build file (and target) is currently.

As for the use of ant. It is heavily used in the i2b2 project. I believe some have even converted the ant scripts into maven POM files, though we have not done the same quite yet. Using JAXB you can define a schema (in this case, the schemas represent the xml transmitted to/from the REST based webservices), and have that schema transformed into Java source code (via XJC) so that the underlying XML can be manipulated against within Java quite easily. Castor is another such project that provides similar code generation and XML marshalling/unmarshalling capabilities.

Once you get your Eclipse environment setup, you will not interact with these ant files except for when you are trying to build a deployment, or when you are changing either an EJB definition or a schema that is generated into java code via JAXB. So, pretty much, its just an initial hurdle that you have to jump over.

Jeremy Nix

**On 09/15/2010 10:08 AM, Dan Connolly wrote:**

OK, thanks for the details.

2 questions:

1. How did you pick up that bit of knowledge? Is it just a normal J2EE thing, or is it specific to i2b2? (I'm learning this whole J2EE world as I go, and I haven't read up on JAXB yet).
2. Do i2b2 developers normally use ant directly like this? Or is this a way of making up for some step that I missed?

**On 9/15/10 7:00 AM, Jeremy Nix wrote:**

Dan, it looks like you need to create the JAXB generated objects that are used as apart of the i2b2 Axis2 web services (see schemas from edu.harvard.i2b2.xml project). From Eclipse, right click on the build.xml in the CRC cell project (edu.harvard.i2b2.crc) and choose "Ant Build...". Choose the targets "clean, jaxb\_gen, generate-ejb" in that order, and click the Run button. Watch the console and make sure that no errors occurred during the JAXB code generation process. If all goes well, right click on that CRC cell project and click Refresh. You should notice that a gensrc folder appears and now has all of the source code that you were missing. You may have to repeat this for other projects if similar issues occur. Some projects do not have EJB's so there will not be a generate-ejb target in their build.xml. Always make sure to refresh the project after you run this as Eclipse does not pick up the fact that new files were generated within the project.

Jeremy Nix

**On 09/14/2010 05:51 PM, Dan Connolly wrote:**

Well, that document helped, but it only seems to address building the Workbench, in detail. I'm trying to use eclipse to navigate the hive sources. When I follow these instructions and then try to import the i2b2-core-source.zip into the same workspace, I get a bunch of errors again:

Description Resource Path Location Type  
PasswordType cannot be resolved to a type CRCAxisAbstract.java /edu.harvard.i2b2.fr/src/test/edu/harvard/i2b2/fr/ws line 57 Java Problem  
PasswordType cannot be resolved to a type CRCAxisAbstract.java /edu.harvard.i2b2.fr/src/test/edu/harvard/i2b2/fr/ws line 57 Java Problem  
Project 'edu.harvard.i2b2.crc.loader' is missing required source folder: 'gensrc' edu.harvard.i2b2.crc.loader Build path Build Path Problem  
Project 'edu.harvard.i2b2.crc' is missing required source folder: 'gensrc' edu.harvard.i2b2.crc Build path Build Path Problem  
Project 'edu.harvard.i2b2.ontology' is missing required source folder: 'gensrc' edu.harvard.i2b2.ontology Build path Build Path Problem  
Project 'edu.harvard.i2b2.pm' is missing required source folder: 'gensrc' edu.harvard.i2b2.pm Build path Build Path Problem  
Project 'edu.harvard.i2b2.workplace' is missing required source folder: 'gensrc' edu.harvard.i2b2.workplace Build path Build Path Problem  
The import edu.harvard.i2b2.fr.dataavo.i2b2message.PasswordType cannot be resolved CRCAxisAbstract.java /edu.harvard.i2b2.fr/src/test/edu/harvard/i2b2/fr/ws line 30 Java Problem  
The method getAlgorithm() is undefined for the type File SendfileRequestHandler.java /edu.harvard.i2b2.fr/src/server/edu/harvard/i2b2/fr/delegate line 101 Java Problem  
The method getOverwrite() is undefined for the type File SendfileRequestHandler.java /edu.harvard.i2b2.fr/src/server/edu/harvard/i2b2/fr/delegate line 90 Java Problem  
The project cannot be built until build path errors are resolved edu.harvard.i2b2.crc Unknown Java Problem  
The project cannot be built until build path errors are resolved edu.harvard.i2b2.crc.loader Unknown Java Problem  
The project cannot be built until build path errors are resolved edu.harvard.i2b2.ontology Unknown Java Problem  
The project cannot be built until build path errors are resolved edu.harvard.i2b2.pm Unknown Java Problem  
The project cannot be built until build path errors are resolved edu.harvard.i2b2.workplace Unknown Java Problem

More clues, please?

**On 9/14/10 4:12 PM, Dan Connolly wrote:**

Yes, this seems to cover the information I need.

For reference... [https://www.i2b2.org/software/projects/workbench/i2b2Workbench\\_Developers\\_Guide\\_14.pdf](https://www.i2b2.org/software/projects/workbench/i2b2Workbench_Developers_Guide_14.pdf) <https://www.i2b2.org/software/repository.html?p=14>

**On 9/14/10 3:58 PM, Lori Phillips wrote:**

Dan,

I don't know if you are using the i2b2Workbench\_Developer's\_Guide\_1-5; it should answer your questions.

Lori

**On 9/14/10 3:52 PM, Dan Connolly wrote:**

The use of ant to build i2b2 from source is documented clearly enough, but there are also .project files in the source that suggest using eclipse.

Can someone who uses eclipse explain how they do it?

What I tried is:

1. start eclipse; pick a workspace
  2. File/Import/General/Existing Projects Into Workspace
  3. pick the i2b2 src directory containing edu.harvard.i2b2.common etc., and import
- I end up with a bunch of errors:

Description Resource Path Location Type

Project 'edu.harvard.i2b2.crc' is missing required source folder: 'gensrc' edu.harvard.i2b2.crc Build path Build Path Problem  
Project 'edu.harvard.i2b2.fr' is missing required library: '/edu.harvard.i2b2.eclipse.plugins.webservicesAxis2/annogen-0.1.0.jar' edu.harvard.i2b2.fr Build path Build Path Problem  
Project 'edu.harvard.i2b2.ontology' is missing required source folder: 'gensrc' edu.harvard.i2b2.ontology Build path Build Path Problem  
Project 'edu.harvard.i2b2.pm' is missing required source folder: 'gensrc' edu.harvard.i2b2.pm Build path Build Path Problem  
Project 'edu.harvard.i2b2.workplace' is missing required source folder: 'gensrc' edu.harvard.i2b2.workplace Build path Build Path Problem  
src/core/, gensrc/ is not included in any "source." build entry build.properties /edu.harvard.i2b2.common line 1 Plug-in Problem  
The project cannot be built until build path errors are resolved edu.harvard.i2b2.crc Unknown Java Problem  
The project cannot be built until build path errors are resolved edu.harvard.i2b2.fr Unknown Java Problem  
etc.

I'm pretty new to eclipse and I'm stumped by these errors.

I found an Ant tab, and using that I can build the the dist target of the common project just fine.

Is that how people use eclipse with the i2b2 source code, by just ignoring the errors above and using ant rather than eclipse's built-in build stuff?

Can eclipse still navigate the code (find all usages, navigate to source) even with the problems above? It seems to work OK, I guess.

Dan

---

#### What F12 Is Supposed to Do When Ajax Error [09-14-2010]

**On 9/14/10 10:26 AM, Mike Mendis wrote:**

Dan,

Usually this occurs when the webclient is unable to connect to a i2b2 cell. The Lite version of firebug was included which the F12 button was suppose to activate and allow the user more detailed info on the issue. I will look into this and get it working.

Thanks

mike

**On 9/10/10 2:08 PM, Dan Connolly wrote:**

We see this message in the web client occasionally:

"An error has occurred in the Cell's AJAX library.

Press F12 for more information"

Nothing happens when we press F12.

What is supposed to happen?

So far, we've been able to diagnose the problems using jboss logs, firebug, etc., but we're starting to beta test with users outside out development team, and that sort of thing won't be an option.

Oh... searching the source suggests F12 is supposed to bring up firebug lite:

webclient/js-ext/firebug/firebug-lite.js: var iconTitle = "Click here or press F12, (CTRL|CMD)+SHIFT+L or SHIFT+ENTER to show Firebug Lite. CTRL|CMD click this icon to hide it.";

Does that work for anyone else?

---

#### Default Credentials [09-14-2010]

**On 9/14/10 10:13 AM, Mike Mendis wrote:**

By default all passwords are demouser

The 'demo' user has full data access so it can run all queries and get back all results

The 'i2b2' user has full admin access so it can create and delete users

If you login to the admin tool ([http://]{url}/admin and use the 'i2b2' user, you can create new users and change the password for the i2b2 user.

mike

**On 9/13/10 7:57 PM, Dustin Key wrote:**

Hello,

Does anyone know whether there a configuration that controls the default login credentials for the I2B2 web client and whether that can be disabled? Our install preloads the username and password.

Thanks!

Dustin Key

---

#### Hive Cell/Oracle Password Security Approach [DATE]

**On 9/13/10 12:39 PM, Dan Connolly wrote:**

Thanks for the detailed reply.

I looked into that login-module:

org.jboss.resource.security.SecureIdentityLoginModule

and I discovered

"... It uses a hard-coded password to encrypt/decrypt the datasource password."

<http://community.jboss.org/wiki/EncryptingDataSourcePasswords>

So I'm not sure that particular technique helps very much in our case.

But this does give me a bit of a feel for how login-config.xml and such work, and perhaps I'll use one of the other techniques on that wiki page.

Dan

**On 9/10/10 6:21 PM, Raj Kuttan wrote:**

Attached ([encryptingPasswords.pdf](#)) are the steps to encrypt the datasource password.

Raj

**On 9/10/10 2:16 PM, Dan Connolly wrote:**

When we follow the i2b2 hive installation instructions, we end up with these datasource files that have passwords in them:

/usr/local/jboss-4.2.2.GA/server/default/deploy/crc-ds.xml

/usr/local/jboss-4.2.2.GA/server/default/deploy/crc-jms-ds.xml

/usr/local/jboss-4.2.2.GA/server/default/deploy/ont-ds.xml

/usr/local/jboss-4.2.2.GA/server/default/deploy/pm-ds.xml

/usr/local/jboss-4.2.2.GA/server/default/deploy/work-ds.xml

Does anyone have advice on how to secure those passwords?

The approach I'm using now is to create a dedicated 'jboss' user account, have jboss run as that user, and have those \*-ds.xml files readable only by the jboss user.

I've been trying to get my head around the whole world of JBoss and J2EE...

I gather these \*-ds.xml files come from the Java Connection Architecture (JCA).

Now I see something about JAAS and login-config.xml in JBoss documentation about JDBC connections <<http://docs.jboss.org/jbossas/docs>

/Server\_Configuration\_Guide/4/html/Connectors\_on\_JBoss-Configuring\_JDBC\_DataSources.html> . Does anybody have experience using that stuff?

---

#### SQL Server and #TEMP\_PDO\_INPUTLIST error [09-10-2010]

**On 9/10/10 12:11 PM, Brian Eliason wrote:**

Hi Raj,

Attached are the xml request/response files when selecting multiple concepts and rendering table within the table viewer plugin. Again, we're using SQL Server 2008 R2.

This might be helpful: If we only select one concept, the problem doesn't occur.

Thanks,

Brian Eliason

**On 9/9/10 2:41 PM, Raj Kuttan wrote:**

Hi Brian,

Can you post the PDO request xml body, just to verify the xml not missing something.

Thanks

Raj

**On 9/9/10 4:25 PM, Prakash Lakshminarayanan wrote:**

Brian

Could you try creating a global temp table outside of the source and modify the source-code to just insert or delete records on that temp table? Inherently data inserted onto a temp table is available only for the session inserting them hence multiple sessions could easily operate on the same temp table as long as the data is cleaned up properly by the session. This approach worked for us on Sybase IQ for a different scenario in i2b2.

Prakash

**On 9/9/10 12:38 PM, Brian Eliason wrote:**

Hi,

We are preparing to migrate from Oracle to SQL Server EE 2008 R2 on i2b2 1.5.1. We've noticed/documented (confirmed by other groups) strange behavior in the table viewer plugin of the 1.5.1 fat client release when running on SQL Server (this behavior doesn't occur on Oracle). Basically, when selecting multiple concepts within the table viewer, the patient demographic information returns NO rows. The visit information as well as the observation information for both concepts come back fine. Again, we've confirmed this behavior with other groups running 1.5.1 on SQL Server. In reviewing the logs, we consistently find the following errors in different points of the workflow:

com.microsoft.sqlserver.jdbc.SQLServerException: There is already an object named '#TEMP\_PDO\_INPUTLIST' in the database.

com.microsoft.sqlserver.jdbc.SQLServerException: Invalid object name '#TEMP\_PDO\_INPUTLIST'

To me, this suggests that for some reason SQL Server is attempting to create a temp table which already exists, and that in the process of attempting to create this temp table (#TEMP\_PDO\_INPUTLIST), it is somehow being truncated or deleted, resulting in a blank patient demographic result set within the table viewer.

Has anyone else encountered this error? Is there a simple work around in the source?

Thanks,

Brian Eliason

---

#### Where is "Export Data" in workbench 1.5 (Windows) [09-01-2010]

**On 9/1/10 1:37 PM, Brian Eliason wrote:**

That's if you installed it; doesn't sound like that's happened yet. So, first, perhaps you should go to HELP->UPDATE and grab the plug in of interest.

Brian Eliason

**On 9/1/10 1:29 PM, Marius Petruc wrote:**

if you installed the Export Data plug in (i2b2exportdata-15.zip), you should see the option under Window->Show view->Other, between the Cluster Category and Help.

Marius Petruc

**On 9/1/10 1:23 PM, Qiu Haijun wrote:**

Hi,

I installed latest i2b2 workbench 1.5 for windows. I can not find the "Export Data" tab. It's not available in "Windows-Show View" options.

How can enable "Export Data" in workbench 1.5?

Thanks,

Haijun Qiu

---

#### Can't get the VMServer to work [08-31-2010]

**On 8/31/10 1:37 PM, Mike Mendis wrote:**

Try adding the name of your computer to your windows hostfile

c:\windows\system32\drivers\etc\hosts

Add to the one listed as localhost, so my computername is mikepc, the line would look like:

127.0.0.1 localhost mikepc

**On 8/31/10 1:25 PM, Terry Shen wrote:**

It's so strange! I keep getting this error (see attached screenshots) when I restart my computer. But, if I uninstall VMServer and re-install it, the first time it will work (but only on Internet Explorer, not Firefox). But, if after I've got it to work the first time, and I restart my computer, then it will no longer work. Also, if you have any intro slides for i2b2, and how it compares to caBIG or EPIC, can you please send it my way?  
Thank you,  
~Terry

---

#### Encrypting Sensitive Data [08-30-2010]

**On 8/30/10 1:34 PM, Russ Waitman wrote:**

Sebastian, Keith and Shawn,

Thank you all for your timely help. We are working with our partner organizations that contribute data to understand security options in i2b2 and source systems.

I think they may have read i2b2 uses encryption at rest somewhere but it wasn't obvious which columns that is regarding or how one configures such a setting.

We are also trying to make them understand that if the source file you send is unencrypted (say an EPIC Clarity backup), then when we re-hydrate it into Oracle to start our ETL process, it's also unencrypted. (Unless we go ahead and encrypt the entire OS file system). Since it's sitting on a secured server in the locked data center, physical theft is not a high probability. Arvinder was referring to the fact that column level explicit encryption of a source EPIC Clarity build (using something like Oracle Transparent Data Encryption, TDE) is a non trivial task since out of the box it contains thousands of tables; each of which with many columns.

We're trying to get to consensus that for "in motion" data, we can use secure communication methods.

We then have two servers storing data "at rest": one for staging and linking identified data and the second to store the de-identified data that will be accessed by i2b2.

For these physically large servers screwed into a rack in the secured data center, the possibility of theft is close to zero. That's why it may be acceptable for the data sitting on the identified staging server and the de-identified server to be unencrypted. If someone penetrated security enough to get root or superuser access to where they can see the oracle data files, your pretty much compromised anyway.

If we were to apply various encryption methods, it would seem that encrypting the entire filesystem would be the easiest most all encompassing option and wonder if anyone has experience with what kind of performance hit you suffer. We naively guesstimate 10%.

We don't see how we could encrypt at rest data inbound from EPIC Clarity without code to transform every column in every table in Clarity using something like TDE or AES.

We also don't see how we could encrypt every column in the star schema underneath i2b2 unless we also modified i2b2 code.

Thanks again to the community for any shared advice regarding best practices,

Russ Waitman

**On 8/30/10 12:29 PM, Arvinder Choudhary wrote:**

Team:

You are right Keith. My Bad. I forgot to mention that it is with regards to Source tables.

Arvinder

**On 8/30/10 12:25 PM, Keith Marsolo wrote:**

Arvinder,

Are you talking about encrypting the i2b2 tables, or the tables of your source data (I'm assuming that's the reference to the 1000s of tables).

**On 8/30/10 12:18 PM, Shawn Murphy wrote:**

Hi Arvinder.

Built into i2b2 is an option to encrypt the "blob" column of the fact table using encryption at rest. The formulation of the algorithm is AES, and the decryption is managed with a key distributed to the client. We use it commonly for text notes of various kinds.

Thanks,  
Shawn.

**On 8/30/10 12:16 PM, Sebastian Mate wrote:**

Hello Arvinder,

I'm not sure what you mean with "1000 tables".

But it might be that we have the same problem over here. We need to implement access-restrictions on a record level, that means for every entry in OBSERVATION\_FACT.

Some time ago we have succeeded in a proof-of-concept. This requires one simple modification somewhere in the CRC's source code to work. Again, this is proof-of-concept and I can't say anything about the performance with millions of observation facts. The attached PPT shows the idea.

If anyone is interested in the modified code, please drop me mail ...

Sebastian

**On 8/30/10 11:21 AM, Arvinder Choudhary wrote:**

Team:

We are in the process of receiving sensitive hospital data and we want to make it secured. What is the best to go.

1. (At a Database level) My opinion is; Securing data at a column level will be enormous effort and little unrealistic as we are looking at 1000 of tables.

Does anyone has better idea

2. Has anyone done encryption at file level or device level and what has been your experience.

Regards

Arvinder Choudhary

---

#### Vmware Server Console Question [08-27-2010]

**On 8/27/10 12:17 PM, Mike Mendis wrote:**

I would select VMware Server Home Page.

**On 8/27/10 11:07 AM, Terry Shen wrote:**

I was wondering if someone could help me with this question. I'm trying to download the VM Server from the i2b2 website. I followed the directions, but under Program -> VMware -> VMServer, I only have two options, Manage Virtual Networks and VMware Server Home Page. I can't find the VMware Server Console application?

Terry Hsin-Yi Shen

---

#### Using CAS with i2b2 Web Client and PM Cell [08-27-2010]

*Note: This discussion is a continuation of the original discussion titled: Using CAS with i2b2 web client and PM cell. The date of the discussion was 07-23-2010*

**On 8/27/10 10:54 AM, Dan Connolly wrote:**

OK, I got it pretty much working, for my purposes... it's not as flexible as it should be:

CAS support in the PM cell should be a configuration option or something, but what I've done just assumes you want to use CAS and leaves no other options.

The client assumes there will be a ticket in the URL, as a result of a link from our portal, via the CAS server.

And after the user logs in on cas.kumc.edu, they are redirected to the i2b2 webclient a la:

the CAS server address is hardcoded (as is the i2b2 domain). This should obviously be a configuration option.

And it's not polished; I haven't done anything about logout yet, for example.

The code I wrote is attached in the form of patches.

[.pm-diffs.PATCH](#)

[.webclient-cas.PATCH](#)

I also put the version history, starting with the i2b2 1.4 code, on bitbucket. The Bitbucket provides .zip downloads.

Before I forget... I found the webclient source confusing in a few places...

1. This diagnostic...

```
console.info("AJAX Login Successful! Updated: i2b2.PM.model");
```

... fires even when login is actually **unsuccessful**, i.e. when the username/password is incorrect.

2. In `./webclient/js-i2b2/hive/master_view_ctrlr.js`:

```
alert('
```

```
Analysis tools are disabled - This access attempt has been logged to the  
server
```

```
');
```

As far as I can tell, that attempt has **not** been logged to the server.

Dan Connolly

---

#### Trouble Building CRC Cell:

**"package edu.harvard.i2b2.common.exception does not exist" [08-26-2010]**

**On 8/26/10 3:37 PM, Raj Kuttan wrote:**

Looks like the common project is not built first.

Try this sequence

```
cd edu.harvard.i2b2.common
```

```
ant dist
```

```
cd ../edu.harvard.i2b2.crc.loader
```

```
ant dist
```

```
cd ../edu.harvard.i2b2.crc
```

```
ant dist deploy
```

**On 8/26/10 2:45 PM, Dan Connolly wrote:**

I think I figured out how to customize the PM cell to integrate with CAS for single sign-on, but I'm struggling to deploy the other cells to test it.

In particular, I can't for the life of me get the crc.loader part of the CRC cell to build.

I'm getting piles of errors like this:

```
[javac] /usr/local/i2b2-backend/i2b2/src/edu.harvard.i2b2.crc.loader/src/server/edu/harvard/i2b2/crc/loader/dao/ProviderDAO.java:16: package edu.harvard.
```

```
i2b2.common.exception does not exist
```

```
[javac] import edu.harvard.i2b2.common.exception.I2B2Exception;
```

```
[javac] ^
```

```
[javac] /usr/local/i2b2-backend/i2b2/src/edu.harvard.i2b2.crc.loader/src/server/edu/harvard/i2b2/crc/loader/dao/VisitDAO.java:18: package edu.harvard.
```

```
i2b2.common.exception does not exist
```

```
[javac] import edu.harvard.i2b2.common.exception.I2B2Exception;
```

```
[javac] ^
```

```
[javac] /usr/local/i2b2-backend/i2b2/src/edu.harvard.i2b2.crc.loader/src/server/edu/harvard/i2b2/crc/loader/dao/PatientDAO.java:58: cannot find symbol
```

```
[javac] symbol : class I2B2Exception
```

```
[javac] location: class edu.harvard.i2b2.crc.loader.dao.PatientDAO
```

```
[javac] String tempPatientMappingTableName) throws I2B2Exception {
```

It seems that the ant build stuff for crc.loader somehow isn't connecting with the i2b2.common stuff, but I'm pretty new to ant and finding out where the problem is looks like finding a needle in a haystack.

Has anybody else seen this error?

A complete log of 'ant -f build.xml clean dist' is attached, in case it helps anybody diagnose this problem.

Dan Connolly

---

### Mixed Type Labs [08-18-2010]

**On 8/18/10 5:12 PM, Shawn Murphy wrote:**

No, we don't have that kind of data object available. However, is it really a set of enums masquerading as a number? If there are only a few number-like results available, is usually that the machine they are reading off of is only supporting 4 or 5 measurements, but they make them more "interpretable" by giving the specific ranges that roughly match their calibrations. In that case, we just set the values up as a set of enumerations. If it really is a number, perhaps it would be possible to change "Negative" to zero.

Thanks,  
Shawn.

**On 8/18/10 1:00 PM, Keith Marsolo wrote:**

When populating the c\_metadatatxml field in the ontology to enable the "Set by Value" functionality, it appears that a lab either needs to be classified as "text" or "numeric," but not both. Is that a correct assumption?

For instance, we have a test (Urine Glucose), where the normal results are text ("Negative"), but the abnormal results are numeric ( $\geq 150$ ,  $\geq 1000$ , etc.). We'd like to be able to encapsulate both types in the metadata XML, so that we can filter on either the text or the values. Is there a way to support this, or plans to enable it? We're happy to make the change on our end, but we didn't want to duplicate anyone else's effort.

Thanks.  
Keith

---

### NLP Cell Distribution [08-18-2010]

**On 8/28/10 2:00 PM, Shawn Murphy wrote:**

Hi Marius,  
Qing Zeng and Sergey Goryachev are the creators of that code base.  
Thanks,  
Shawn.

**On 8/17/10 11:04 AM, Marius Petruc wrote:**

Hi all,  
i am trying to install the NLP cell. The package i downloaded form the i2b2 site includes two archives (NLP-client-src-10.zip and NLP-server-src-10.zip), and not the four folders described in the NLP-installation-10.pdf (included in the package). my impression at this point is that the server-src needs to be built on the middle tier (on jboss, like all the other cells), and that the client-src needs to be configured inside the workbench. is this correct? does anybody know what's the ant command to build the server? and how do you configure the client?  
thanks a lot  
marius

---

### i2b2 1.5.1 PM Cell Installation Error [08-13-2010]

**On 08/13/2010 10:40 AM, Qiu, Haijun wrote:**

Thanks a lot for all the responses. I initially installed the latest jdk 1.6 update 21. After I switched to jdk 1.5 update 22, PM cell is working.  
Haijun Qiu

**On 8/12/10 3:41 PM, Jeremy Nix wrote:**

Haijun,  
You will get that error from the PM admin screens if **any** error is passed back from the PM Cell. I'd start looking at the jboss logs to see if anything useful is being logged. The log file is located here: /opt/jboss/server/default/log/server.log. You'd want to scan the file looking for the word ERROR. You may want to add the following lines in your /opt/jboss/server/default/conf/jboss-log4j.xml file in order to make the logging more verbose. Be sure to switch the log level back to WARN or ERROR once you have figured out what the problem is.

```
<!--=====-->
<!--i2b2 specific logging-->
<!--=====-->
<category name="edu.harvard.pm" additivity="false">
<priority value="DEBUG" />
<appender-ref ref="FILE"/>
</category>
<category name="edu.harvard" additivity="false">
<priority value="WARN" />
<appender-ref ref="FILE"/>
</category>
Good luck.
Jeremy Nix
```

**On 08/12/2010 03:23 PM, Qiu, Haijun wrote:**

Yes, it is listed as active. I tried tcpdump to track the tcp packets. when I connect from the i2b2workbench, there is some traffic to port 9090, although I got "The PM Cell is either down or the PM address in the properties file is in correct". When I connect from the PM cell admin login page, there is no traffic to port 9090. Any clue?

<http://ourhost:9090/i2b2/services/listServices>:

Available services  
PMService



Service EPR : <http://ourhost:9090/i2b2/services/PMService>  
Service Description : PMSERVICE  
Service Status : Active  
Available Operations  
getVersion  
getServices

Thanks,  
Haijun

**On 8/12/10 2:23 PM, Dan Connolly wrote:**

Were you able to verify that the PM web service is running, as discussed on page 6 of the PM install guide?

7. Verify webservice is running

a) Check url 'http://yourHost:9090/i2b2/services/listServices' in a browser.  
Verify that PMSERVICE is listed as active.

I happen to be doing some development around authentication, and I discovered that the "Your account does not have access to any i2b2 projects." message is given in several cases:

username you gave isn't in the oracle database

password is wrong

no access to any i2b2 projects (i.e. what the message says).

The webclient might be giving that message in the case that the php proxy cannot reach the PM service at all, but I'm not sure.

Be sure to check the error logs for the web server that hosts the web client.

**On 8/12/10 1:29 PM, Haijun Qui wrote:**

Hi Marius,

Thanks,

Apache 2.2.3, PHP5.1.6, php is indeed working and jboss was started with -b 0.0.0.0

What's other possible reason?

Thanks,

Haijun

**On 8/12/10 1:19 PM, Marius Petruc wrote:**

1) what web server do you use? please verify that php is working

2) if php is working, please verify that you're starting jboss with the appropriate arguments: jboss -b 0.0.0.0

hope this helps

Marius

**On 8/12/10 12:38 PM, Haijun Qui wrote:**

Hi,

I am installing i2b2 version 1.5.1. After loading data in to oracle, I installed PM Cell. But

PM Cell did not work as expected.

After launching the PM webclient at <http://ourhost/admin> and use i2b2/demouser to login, I got error:

"Your account does not have access to any i2b2 projects."

From i2b2workbench, I got error: "The PM Cell is either down or the PM address in the properties file is in correct".

i2b2workbench.properties:

writeTimelineFile=yes

applicationName=i2b2

messageversion=1.1

demoUser=no

i2b2.1=HarvardDemo.REST,<http://services.i2b2.org/i2b2/rest/PMService/>

i2b2.2=i2b2demo.REST,<http://ourhost:9090/i2b2/rest/PMService/>

<http://ourhost:9090/i2b2/services/listServices>:

Available services

PMSERVICE

Service EPR : <http://ourhost:9090/i2b2/services/PMService>

Service Description : PMSERVICE

Service Status : Active

Available Operations

getVersion

getServices

I must be missing something, please advise.

Thanks in advance,

Haijun Qui

---

#### Data Installation – New Install (v1.5.1) Build Failed [08-10-2010]

**On 8/10/10 4:04 PM, Marius Petruc wrote:**

ok, the problem was caused by trailing spaces in the db.properties file. cleaned it, it worked fine.

thanks a lot everybody!



**On 8/10/10 3:51 PM, Mike Mendis wrote:**

It could be, try clearing out your classpath and running it again  
mike

**On 8/10/10 3:49 PM, Marius Petruc wrote:**

that was it. i was copying and pasting. i typed the command by hand and it worked (sort of). that's odd, since i didn't have this problem with the other builds...  
anyway, it's now complains that the jdbc driver can't be loaded:  
C:\i2b2source\edu.harvard.i2b2.data\Release\_1-5\NewInstall\Metadata\data\_build.xml:33:  
Class Not Found: JDBC driver  
com.microsoft.sqlserver.jdbc.SQLServerDriver could not be loaded  
Can this be due to some conflict with my MS jdbc 3.0 driver already configured in CLASPATH?  
anyway, thanks a lot for your help!

**On 8/10/10 3:23 PM, Raj Kuttan wrote:**

Could be a special character in the command, if you cut & paste from the install doc.  
Please try again by typing the ant command.  
Raj

**On 8/10/10 3:22 PM, Mike Mendis wrote:**

Which version of ant are you running? This is kinda interesting, because if you run ant alone, than your message is expected. But providing the -f command tells ant to a specific xml file, this case data\_build.xml. So I would expect to see at least:  
Buildfile: data\_build.xml does not exist!  
Build failed  
I assume there is not a return hit after the ant was entered?  
mike

**On 8/10/10 3:14 PM, Marius Petruc wrote:**

Hi!  
i've created the required databases in the SQL server and i'm trying to create the tables. when i try to run the build i'm getting an error message related to build.xml not being available:  
C:\i2b2source\edu.harvard.i2b2.data\Release\_1-5\NewInstall\Metadata>ant -f data\_build.xml create\_metadata\_tables\_release\_1-5 Buildfile: build.xml does not exist! Build failed  
did anybody have this problem? any suggestions?  
thanks a lot!  
Marius Petruc

---

#### **i2b2 Upgrade from VM [08-09-2010]**

**On 8/9/10 1:03 PM, Mike Mendis wrote:**

I assume you are currently running 1.4 and want to move to 1.5?  
The easiest path is to following:

1. Update the database from the edu.harvard.i2b2.data project
2. Follow the upgrade scripts for the ontology, CRC, workplace and Project Management cell. You will need to update the properties files in each to point to your external oracle DB. Than do the standard ant build deploy for each.

mike

**On 8/9/10 10:13 AM, John Reber wrote:**

Hi All,  
We currently have the i2b2 Full VM installed, which was originally configured by a third party and points to an external Oracle database server . We want to upgrade to the latest version. Is there a list of which configuration files will have to be migrated over?

John Reber

---

#### **Set Value XML Metadata [08-03-2010]**

**On 08/03/2010 3:45 PM, Lori Phillips wrote:**

The end of the Ontology messaging document describes the Optional metadata\_xml content associated with Value Metadata.  
Release 1.5 of the workbench client has a new Ontology wizard (Edit Terms) that assists in the creation of Ontology terms. This functionality is enabled for users with role = EDITOR. This wizard also has a series of help pages that describe each field in the ValueMetadata.  
Lori Phillips

**On 08/03/2010 3:29 PM, Jack London wrote:**

Is there documentation describing the c\_metadatatxml field in the i2b2metadata table? By trial and error I have determined that "normal units" (e.g., "years") must be specified if you desire setting value constraints with the i2b2 Query tool. Without entering a value for this tag the "set value" pop-up

window hung.

So, is there any document that describes which tags MUST be set to achieve various capabilities (e.g., selecting "high" values).

Thanks for any help.

Jack

---

### **i2b2 Wizard Updated to Support i2b2 1.5.1 [08-03-2010]**

**On 08/03/2010 6:12 AM, Sebastian Mate wrote:**

Hi all,

I've updated the i2b2 Wizard to support i2b2 1.5.1. This is a maintenance update; there are no new features. It only supports Ubuntu 10.04 Linux (Desktop or Server, 32 bit) and Oracle.

The tool can be found here: <http://www.wcip.informatik.uni-erlangen.de/~sisemate/>

There are also two small videos, which show the i2b2 Wizard in action.

Greetings from Erlangen

Sebastian

---

### **JBoss Errors with Remote SQL Server Backend [07-28-2010]**

**On 07/28/10 3:49 PM, Lori Phillips wrote:**

JD,

Going to localhost:9090/i2b2/services/ and attempting to call the above operations on their services gives the following error (changed as appropriate for the operation):

Our cell urls are typically localhost:9090/i2b2/rest/ (not /services/)

Is that your problem perhaps?

Lori Phillips

**On 07/28/10 2:52 PM, JD Wolk wrote:**

Thanks Raj

It's not just the analysis plug-ins and the ontology patient count that aren't working, it's everything. I fire up the webclient, and while I'm able to log in, I get 4 consecutive errors saying an error has occurred in the Cell's AJAX library. There is nothing displayed in the "Navigate Terms" section of the ontology, nor in "Workplace", "Previous Queries", etc.

In the Message Log of the webclient, the first getUserAuth call from PM is ok, but the rest (GetCategories, GetSchemes, getFoldersByUserId, getQueryMasterList\_fromUserId, and getQueryMasterList\_fromUserId) are yellow, and the 'rcvd' section is blank.

Going to localhost:9090/i2b2/services/ and attempting to call the above operations on their services gives the following error (changed as appropriate for the operation):

Umarshaller error: Error during unmarshall <getCategories />; nested exception is:

edu.harvard.i2b2.common.exception.I2B2Exception: Umarshaller error: Error during unmarshall <getCategories />; nested exception is:

org.apache.axis2.AxisFault: Umarshaller error: Error during unmarshall <getCategories />; nested exception is:

edu.harvard.i2b2.common.exception.I2B2Exception: Umarshaller error: Error during unmarshall <getCategories />

**On 07/28/10 2:28 PM, Raj Kuttan wrote:**

You can remove crc-oracle-jdbc2-service.xml, but do you see the file crc-mssql-jdbc2-service.xml.

If there is error in the TimerImpl, then the analysis plug-ins will not Work. The default plugin(patient count) in the ontology view wont work.

Hope it helps.

Raj

**On 07/28/10 2:16 PM, JD Wolk wrote:**

Hi Mike,

Thanks for responding so quickly. The answer is yes and no. I tried with two instances of i2b2, one with the vm image where I JUST changed the crc-ds.xml, crc-jms-ds.xml, ont-ds.xml, pm-ds.xml, work-ds.xml, CRCApplicationContext.xml and CRCLoaderApplicationContext.xml, and a separate one where I built from source (via a slightly modified version of Sebastian Mate's i2b2wizard) and made all the changes I detailed in my previous message. In both instances I ran into the second set of errors below, i.e. "ERROR [TimerImpl] Error invoking ejbTimeout..."

I tried your suggestion, and the first error (DeploymentException) went away. I wasn't too concerned about this error in the first place, but it's nice to have a cleaner JBoss deployment now. Thanks.

Any idea about the second set of errors?

**On 07/28/2010 1:42 PM, Mike Mendis wrote:**

Are you using the vm image as a starting point?

If so than you will need to remove that file:

/opt/jboss-4.2.2.GA/server/default/deploy/jms/crc-oracle-jdbc2-service.xml

Because it appears to have already put the correct sqlserver one in.

mike

**On 07/28/10 1:40 PM, JD Wolk wrote:**

Hi all,

I'm JD Wolk, a developer working under Lisa Dahm and Charles Boicey from the UC Irvine Medical Center. I'm currently trying to deploy a (test) instance of i2b2 v1.5 on an Ubuntu 10.04 machine against a remote SQL Server 2008 backend running on Windows 7.

Whenever I start up JBoss on the Ubuntu machine, I get the two following (and probably unrelated) sets of error messages:

17:39:57,114 ERROR [MainDeployer] Could not create deployment:

file:/opt/jboss-4.2.2.GA/server/default/deploy/jms/crc-oracle-jdbc2-service.xml>l

org.jboss.deployment.DeploymentException: Trying to install an already registered mbean: jboss.mq:service=DestinationManager

Incompletely deployed packages

org.jboss.deployment.DeploymentInfo@1c2fcd1c {url=[file:/opt/jboss-4.2.2.GA/server/default/deploy/jms/crc-oracle-jdbc2-service.xml]}

deployer: [org.jboss.deployment.SARDeployer@e64686](mailto:org.jboss.deployment.SARDeployer@e64686)

status: Deployment FAILED reason: Trying to install an already registered mbean: jboss.mq:service=DestinationManager

state: FAILED

watch:

file:/opt/jboss-4.2.2.GA/server/default/deploy/jms/crc-oracle-jdbc2-service.xml

altDD: null

lastDeployed: 1280191197099

lastModified: 1280191197000

mbeans:

...  
18:25:43,593 ERROR [TimerImpl] Error invoking ejbTimeout: javax.ejb.EJBException: java.lang.IllegalArgumentException: dataSource is required  
...

17:40:32,690 INFO [Server] JBoss (MX MicroKernel) [4.2.2.GA (build: SVNTag=JBoss\_4\_2\_2\_GA date=200710221139)] Started in 1m:5s:210ms 17:40:

33,005 INFO [XmlBeanDefinitionReader] Loading XML bean definitions from class path resource [org.springframework.jdbc/support/sql-error-codes.xml]

17:40:33,098 INFO [SQLExceptionCodesFactory] SQLExceptionCodes loaded: [DB2, HSQL, MS-SQL, MySQL, Oracle, Informix, PostgreSQL, Sybase] 17:40:

33,117 ERROR [TimerImpl] Error invoking ejbTimeout: javax.ejb.EJBException: org.springframework.jdbc.BadSqlGrammarException:

PreparedStatementCallback; bad SQL grammar [ select \* from i2b2hive.CRC\_ANALYSIS\_JOB where queue\_name = ? and status\_type\_id = ? order by create\_date]; nested exception is com.microsoft.sqlserver.jdbc.SQLServerException: Invalid object name 'i2b2hive.CRC\_ANALYSIS\_JOB'.

This last [ERROR] keeps repeating about once every minute or so.

I think the problem may have something to do with the way I've configured things. I know that configuring with SQL Server is tricky, and with a remote SQL Server box doubly so. I've double and triple checked all of the files I modified just to be sure that I didn't make any obvious errors.

Here's a list of those files:

crc-ds.xml, crc-jms-ds.xml, ont-ds.xml, pm-ds.xml, and work-ds.xml from the appropriate cells' /etc/jboss directories

CRCApplicationContext.xml, CRCLoaderApplicationContext.xml, OntologyApplicationContext.xml from the /etc/spring directory of the crc, crc.loader, and ontology directories

.properties files from the /etc/spring directories of the crc and crc.loader directories

db.properties files from /etc/scripts in the ontology and workplace directories

build.properties in the crc directory

I also ran the master\_build.xml from each cell's directory after I finished modifying the appropriate files (especially the .xml files from the first two points)

There may have been an additional file here or there. Anywhere I saw "ORACLE" I changed it to "SQLSERVER", "Oracle 9i" to "MS SQLSERVER2000", <connection-url>, <driver-class>, etc. Also included the appropriate DB names/passwords, and gave the ip address to my DB.

All services at localhost:9090/i2b2/services/listServices on my Ubuntu box say they are Active.

What can I do to fix this? I can provide other information if needed.

Thanks in advance

JD

---

#### Using CAS with i2b2 web client and PM cell [07-23-2010]

**On 07/23/10 9:02 PM, Andy McMurphy wrote:**

CAS has not yet been implemented with shrine, though it is a great area to engage collaboration on if you are interested.

andy mcmurphy

(shrine developer)

**On 07/23/10 3:02 PM, Dan Connolly wrote:**

Hi,

We're gearing up to deploy i2b2 in our enterprise where we use JA-SIG Central Authentication Service (CAS <<http://www.jasig.org/cas>> ) for single sign-on, so I'm trying to get i2b2 to use CAS for authentication.

Has anybody else tried this? Any advice on how to go about it?

I see "How to i2b2 <[http://www.i2b2.org/events/slides/i2b2\\_AMIA\\_Tutorial\\_20100310.pdf](http://www.i2b2.org/events/slides/i2b2_AMIA_Tutorial_20100310.pdf)> " has a "How to implement SHRINE" slide (#130) with a list that includes:

LINK your local user authentication system

But I haven't found any detailed clues about how to go about it.

The best idea I have is to use CAS tokens somewhat similar to the session keys in <http://sourceforge.net/projects/i2b2pmcellsecur/develop> as noted in this comment from RequestHandler.java:

//If password begins with "SessionKey:" its a session key and decrypt it and validate it

I'm also poring over the i2b2webclient code, trying to find where to hook in... I'm making progress, but it's slow.

I see some SHRINE code in there, but I don't really see how it works nor how I would use it to hook in CAS.

Dan Connolly

---

#### i2b2 – v1.4 SP1 – DEBUG error [07-23-2010]

**On 07/23/10 1:47 PM, Raj Kuttan wrote:**

Hi Peter,

This is not related to "c\_totalnum", you can ignore this message.

This message is fixed in 1.5.

Thanks

Raj

**On 07/23/10 1:40 PM, Peter Beninato wrote:**

Hi,

In the log, I see this error when a query is executed via the thin client:  
ERROR [CRDAO] Unable to get Estimated Size: ORA-00942: table or view does not exist  
There is not much more information in DEBUG mode (or I'm missing it).  
Is this related to the "c\_totalnum" column in the table "i2b2" in the metadata schema?  
If so, is it just a matter of granting select on that table to the data schema owner?  
Peter Beninato

---

#### **New User Setup [07-13-2010]**

**On 07/13/10 2:08 PM, Brian Eliason wrote:**

Hi Dustin,

If you're using Oracle, make sure to run the updates. There is an issue with the triggers where they aren't being created correctly; the update addresses that, but you'll still want to verify the triggers are successfully created.

If you're using MSSQL, this isn't an issue (in my experience).

Thanks,

Brian Eliason

**On 07/13/10 11:34 AM, Dustin Key wrote:**

Hello,

When I try to add a role to a user\_name in PM Navigation (Projects/[project name]/Users), something strange happens. For instance, I will try to +add the new user and select roles USER and DATA\_OBFSC. I select save updates. The user\_name I just added will then disappear. However, if I just select DATA\_AGG and update, then I will see the new user\_name with roles DATA\_AGG and DATA\_OBFSC assigned. (The latter won't seem to be able to select counts. I assume that's because I haven't successfully assigned either a User/Manager/Admin role.)

My current way around this is to just add the correct roles directly to the pm\_project\_user\_roles table. That seems to work fine. However, I was wondering if anyone recognizes what I've described with the PM Navigator and whether there is a way to fix it?

Thanks!

Dustin Key

---

#### **Reason for visit / chief complaint ontology [07-13-2010]**

**On 07/13/10 1:31 PM, Andrew McMurry wrote:**

Yes, we have worked on precisely this problem in the past.

<http://chip.org/ihl/publications.htm#biosurv>

We also have ETL engines for doing these kinds of operations by loading data on the fly from Emergency Departments

Apologies in advance that I cannot spend much time helping as I am already way committed, but I can point you towards papers and code that may be of help.

andy mcmurry

**On 07/13/10 12:20 PM, Brian Eliason wrote:**

Hi,

We are doing a public health program (syndromic surveillance) in which we need to look at reason for visit / chief complaint data coming from the ED and Urgent Care. As many of you know, this is usually not tied to any diagnostic data, sometimes it is even entered in free-text.

Does anyone have practical experience in mapping chief complaint data into a standardized chief complaint ontology within i2b2? I'm hoping for some open source web services that might accommodate chief complaints in the same way that the ontomapper has navigated RXNorm. Any suggestions?

Thank you,

Brian Eliason

---

#### **Project Request Plug-in [07-09-2010]**

**On 07/09/10 12:54 PM, Shawn Murphy wrote:**

Hi Jeremy,

The physical creation of the data mart is a manual process that is only facilitated by the scripts provided in that package mentioned below. The project request wizard is essentially a way to gather all the prerequisites and specifics needed to build the project and run the package. One of these prerequisites is that an approval already exists that is registered to the user creating the request. In a specific scenario, the approval is from the IRB for a specific investigator. That IRB approval needs to be registered into the PM cell (perhaps via the client, but could be automated) by an administrator. That approval can then be picked by the investigator when they are making their project request based upon a patient set that fits the criteria set out in the IRB approval.

Thanks,

Shawn.

**On 07/09/10 12:34 PM, Jeremy Nix wrote:**

As a follow up to this discussion, what kind of workflow process is involved with this Project Request plug-in? Is the creation of the new datamart a manual process or does the plug-in do this for you? What is the relationship of project requests to approvals? The reason I ask this is that from reading the documentation it seems as though an approval must be created before a user can request a project to be created. I may have misinterpreted, if so I apologize.

Jeremy Nix

**On 07/08/10 6:25 PM, Shawn Murphy wrote:**

Hi Keith,

Indeed, the current design is for each project to have its own data schema. This is so we can cleanly manage user permissions and data ownership when derived data is created for a project. This can be facilitated with a package of sql scripts that can create the new schema and move the data from an existing schema for specific patient sets and data types. They are under the source\_code section of the software page, in a package named i2b2DerivedDataMarts-src-14.zip  
Thanks,  
Shawn.

**On 07/08/10 3:39 PM, Keith Marsolo wrote:**

According to the 1.5 Release Notes, with the Project Request plug-in, new datasets are created by copying existing data (I'm assuming into a new schema).

Is there any intention to include functionality that will allow users to define multiple projects within the same schema?

In other words, if I want to give data\_obfs access to everyone, but then allow certain users to have data\_deid access for specific patient sets (based on previous queries), can I do that within the same project? Or do I need to create a separate schema?

Keith

---

#### Dimension Query [07-08-2010]

**On 7/8/10 1:07PM, Raj Kuttan wrote:**

Brian,

You able to get the patient list because in 1.5 it is defaulting to PROVIDER\_PATH, will fix this in next version.

Thanks

Raj

**On 7/8/10 12:52 PM, Brian Ostasiewski wrote:**

Are you sure? Because we are matching on the ID and the SQL query the i2b2 server generates works correctly and returns an accurate patient list when executed directly against Oracle. If I have problems still with 1.5 I will try your suggestion.

Brian

**On 7/8/10 10:49 AM, Lori Phillips wrote:**

Brian,

I noticed one thing that may be a problem.

In your metadata the

<C\_COLUMNNAME>PROVIDER\_ID</C\_COLUMNNAME>

should be

<C\_COLUMNNAME>PROVIDER\_PATH</C\_COLUMNNAME>

Lori Phillips

**On 7/8/10 9:39 AM, Brian Ostasiewski wrote:**

Was testing it with 1.3 and 1.4 installs. That is good news about 1.5, I will give that a try when it comes out. Thanks!

**On 7/8/10 9:29 AM, Shawn Murphy wrote:**

Hi Brian,

Are you using 1.4 or 1.5RC?

1.5 (final release coming out next week) supports this for providers, but not patient or visits. 1.6 will support those dimensions as well.

Thanks,

Shawn.

**On 7/6/10 3:42 PM, Brian Ostasiewski wrote:**

It appears i2b2 is able to accept queries against dimensions. I tried this with the Provider dimension and am having problems. In my metadata I have an entry along the lines of:

<C\_FULLNAME>WFUBMC\Provider\00000-09999\00236\</C\_FULLNAME>

<C\_BASECODE>00236</C\_BASECODE>

<C\_DIMCODE>WFUBMC\Provider\00000-09999\00236\</C\_DIMCODE>

<C\_FACTTABLECOLUMN>PROVIDER\_ID</C\_FACTTABLECOLUMN>

<C\_TABLENAME>PROVIDER\_DIMENSION</C\_TABLENAME>

<C\_COLUMNNAME>PROVIDER\_ID</C\_COLUMNNAME>

If I run a query through the client, the SQL the i2b2 server generates and executes is along the lines of:

SELECT patient\_num, 1 FROM (

SELECT patient\_num FROM i2b2data.observation\_fact WHERE (provider\_id IN (

SELECT provider\_id FROM i2b2data.provider\_dimension c where provider\_path LIKE 'WFUBMC\Provider\00000-09999\00236%'

)) group by patient\_num having count★ >=1

) t

If I execute this SQL directly against Oracle, it indeed returns a patient list, and fairly quickly. However in the web client or workbench, it hangs. The workbench times out after 3 minutes and says to refresh later to check progress - but it never changes from PROCESSING.

The web client fares worse - after the 3 minutes the server seems to go into an infinite loop of loading the user's permissions and the cell locations, and eventually empty javascript error alerts start popping up endlessly.

Is the ability to query dimensions as I am attempting to do a feature of i2b2 indeed? What might I be doing wrong?

Brian O

---

#### Project User Roles [07-08-2010]

**On 07/08/10 9:44 AM, Mike Mendis wrote:**

What do you mean 'curve out project' If you mean a single star schema (obs\_fact, dim\_tables, ect...) and multiple projects accessing it with each one appearing to its own private space, than that is not possible. What is possible, is to have multiple projects, with a single PM cell/database AND each project has it OWN star schema database.  
You create your projects in the PM admin tool and than create entries in the lookup\_tables and create the corresponding --ds.xml files in the jboss deploy directory.  
mike

**On 07/07/10 4:38 PM, Arvinder Choudhary wrote:**

Mike:

Is there a way that we can curve out project from the complete database. Say we have database ABC and we have one PI who specializes in Hypertension and another in diabetes ..can we create a project where we can assign them and their group a project path within ABC say /ABC/HTN and /ABC/DBT

Arvinder

**On 07/06/10 2:28 PM, Mike Mendis wrote:**

You just need to redeploy the cells. All the data is fine and did not change from 1.4 to 1.4sp1

**On 07/06/10 3:25 PM, Peter Beninato wrote:**

Hi,

With v1.4 SP1 is there information on how to upgrade from v1.4?

It would seem that the cells need to be re-deployed via ant. But with regards to Data, what are the changes between v1.4 and v 1.4 SP1? I would like to apply the Service Pack, but not lose my existing data.

Thanks.

Peter Beninato

**On 06/30/10 8:24 AM, Mike Mendis wrote:**

It is up on the i2b2.org website now, it is referred to as Core Source Code Server v1.4 SP1  
mike

**On 06/30/10 10:40 AM, Arvinder Choudhary wrote:**

Dear Mike:

As you had explained earlier that there is a Flaw in Project User Roles. When are you releasing 1.4.1 ver with a fix.

Regards

Arvinder

---

#### **i2b2 v1.4 – Admin - Hive [07-07-2010]**

**On 07/08/10 9:07 AM, Shawn Murphy wrote:**

I agree Peter, we're working on this on 1.6 so that only admins and managers can see those. There is the ability for different cells to be associated with different projects, so the managers of a project may want some control.

Thanks,

Shawn.

**On 07/07/10 4:42 PM, Peter Beninato wrote:**

Hi,

When I log on as a user onto i2b2 v1.4 SP1 and click on Admin, the information under HIVE seems to be available in a READ mode.

I have some concerns about exposing the cell URLs. I think general users may be directed to a https URL, but the cells are then behind a firewall.

Not that a user could access the CELL URLs directly, but I'm wondering if there is a way from preventing end users from seeing the contents under Admin|HIVE?

Admin|HIVE?

Or the intent in allowing users access to Admin, Admin|Hive?

Peter Beninato

---

#### **i2b2 Sample Management – Demo Website [07-06-2010]**

**On 07/06/10 7:25 PM, Lynn Bry wrote:**

1) Regulatory cell

Purpose: Manage IRB or other types of protocols in i2b2. A user may register protocols in the webclient UI ("Authorities" and "Protocols" tabs in the "Sample Management" link in the webclient). Protocol registration would be restricted to 'managers' of the i2b2 project. Default fields capture what we believe to be common IRB data used across sites (IRB protocol#, assigning authority, etc.).

Sites may add fields to an EAV table that allows the IRB and Authority objects to be extended, as needed.

One or more i2b2 projects may also be mapped to a given IRB protocol so samples (collected under an IRB protocol) can be linked to the datasets in corresponding i2b2 projects.

To protect PHI data as it moves between systems, part of a project's PHI encryption key will be stored within the Regulatory Cell. When a user submits a sample, or cohort request, they provide the second half of the key 'on-the fly'. The Regulatory cell then combines the two parts of the key to decrypt/re-encrypt PHI as required.

The regulatory cell will leverage (and if necessary extend) existing ontology & CRC cell services (e.g ont:get\_child, add\_child, update\_crc\_concept) to maintain and populate the regulatory 'ontology'.

PDF from internal wiki page ([RegulatoryCellDesignSpec-7-6-2010.pdf](#))

(2) Cohort Map ("Cohorts" tab in the "Sample Management" link)

Purpose: Links primary patient identifiers (MRN + SOURCE – 2345667, BWH, e.g.) with a study or project-specific subjectID, where the subjectID is a research-specific identifier for the patient. For the purposes of prospective sample collection - the cohort map provides the primary patient identifiers (MRN) that will be used to find to-be-discarded clinical samples. Upon retrieval and accession for the associated study, a LIMS such as Crimson can

assign the study-specific subjectID, included in the cohort map, as the patient identifier, thus linking the sample with the appropriate i2b2 dataset. For the purposes of retrospective sample repositories - the full cohort can be used to interrogate an existing repository to generate a "picklist" of potential matching samples that could be broadcast back to the source LIMS. This would provide a first step in delivering samples to investigators. We envision that each bank would then implement its own operational procedures to handle subsequent communication with an investigator regarding sample release. A workbench or webclient plugin will be provided that will allow a user in any i2b2 project to drag and drop an existing patient set and 'broadcast' it to the EMSI/sample cell to request samples on the patients it describes (for retrospective repositories). The user will have the ability to explicitly append, replace or refresh the existing sample within their project, thereby preserving the temporal state of their project data.

(3) EMSI - Enterprise Master Specimen Index – stores instances of samples as broadcast from one or more source LIMS. It also maintains parent-child relationships between parent samples and any aliquots or derivatives. A prototype EMSI is in use at Partners and can receive sample messages from the Sunquest LIS and Crimson LIMS. A specification will be provided should sites wish to connect other LIMS, e.g. caTissue, others.

The source LIMS can broadcast any new samples or updates to samples (changes in availability, volume, etc.). The EMSI also tracks an instance of a sample if it is shipped/received across multiple LIMS, where each LIMS frequently assigns a different unique sampleID. Sample messages include the availability of the sample, whether it belongs to a "public" or "private"/study-specific repository, and whether it originated from a dedicated/consented or discarded sample – with flavors of anon, de-ID or consented.

The EMSI will be capable of broadcasting sample data into the CRC, so this information may be queried with other patient data (e.g. find patients with Lupus who have a DNA sample). Should time/resources on our grant permit, a webclient plugin would let an i2b2 user select a PDO to broadcast to the EMSI for the purposes of generating a "picklist" - namely a list of potentially matching sampleIDs, which could be forwarded to the corresponding repository as a request for samples.

As with the regulatory cell, existing ontology and CRC webservice will be exploited where possible to manage sample definitions. However, for performance a direct database import mechanism will also be considered for the 'bulk' transfer of large samples sets into the i2b2 hive – in this manner, sites could perform a "batch" upload/update of samples in the event source systems lack an appropriate webservice or other communication layer.

The ETL process for importing samples will include a mechanism to validate the raw metadata and provide a mean to map and transform the concepts provided by the external LIMS systems to a controlled sample metadata ontology maintained by the sample cell. When mapped a synonym will be created for the raw concept. This will ensure that samples can be located within the hive using either the raw metadata concept or via the preferred i2b2 hive term.

In the case of both cells, to get sample-related data into the CRC, and make it available as ontology concepts for use in the Query Tool, we are considering persisting a representation of the 'protocol' and 'sample' concepts within the c\_METADATAXML field of the ontology table, using a format that adheres to the ISO/IEC 11179-1 ([https://www.nescent.org/wg/dryad/images/c/c5/ISO\\_IEC\\_11179-1\\_2004\(E\)\\_Metadata\\_Registries\\_Framework.pdf](https://www.nescent.org/wg/dryad/images/c/c5/ISO_IEC_11179-1_2004(E)_Metadata_Registries_Framework.pdf)) standard, specifically the caBIG implementation of this standard - Common Data Element. This could potentially help foster closer integration between i2b2 and caBIG platforms.

**On 07/06/10 6:52 PM, Lynn Bry wrote:**

Website with "Example Scenarios" for the demo webclient UI: <http://i2b2crimson.partners.org/>

Demo webclient UI - <http://i2b2crimson.partners.org/i2b2crimson/index.php>

I. "Example Scenarios" goes through a test script to search for samples in the Query Tool.

II. To get to the IRB/sample management components in the webclient

(1) Click "Sample Management" in the upperRH menu.

(2) Page opens with "Authorities", "Protocols", "Cohorts" and "Samples" as options.

(3) Click any entry under a tab to see a detailed view.

(4) This is a demo pre-populated with dummy data, you won't be able to create/modify existing information.

Thoughts, feedback appreciated...

---

#### Linking samples within i2b2 [07-06-2010]

**On 07/06/10 6:49 PM, Lynn Bry wrote:**

Hi all

For those with interest in managing samples within i2b2, let me give an update from my group on the SAMP project

We received an ARRA grant to implement sample management features in i2b2, namely the ability to (1) broadcast existing sample inventories into i2b2, and (2) facilitate creation of cohort maps to enable prospective collection from Clinical Lab/Pathology Depts.

One thing I've learned is every academic center is different, having one or more preferred LIMS that often support multiple repositories. Our goal is to let sites leverage what they have. We'll provide a minimum data specification for sites to load samples from their favorite LIMS into i2b2, plus tools to let users generate a "picklist" of materials that could go back to the source LIMS/repositories that own them. For anyone who runs a biorepository, this is not an end-to-end solution, but it would allow use of i2b2 as a "store-front" to search for materials and start a process to request them.

Thanks go to Brian Wilson, the senior programmer and analyst on the BWH i2b2 team who architected some elegant solutions to maximize use of i2b2's existing infrastructure for sample queries.

I'll forward links to the demo website + more detailed description of the proposed cells in subsequent emails.

Please give us your feedback. In particular, we'd be interested in issues other sites face with sample handling, IRB oversight, + LIMS in use at various places.

Thanks!

Lynn