1.7.13 Release Notes

Important links

Downloads:

- i2b2 Download Page (binaries and virtual machine): https://www.i2b2.org/software/index.html
- i2b2 v1.7.13 demo Docker containers (with SAML IDP demo): https://github.com/kvb2univpitt/i2b2-demo/tree/i2b2-1.7.13
- i2b2 Github (source code): https://github.com/i2b2

Documentation:

- SAML Setup for i2b2: Chapter 8. SAML Setup for i2b2 (v1.7.13 release)
  - SAML Diagrams: Informational Diagrams for SAML Setup
  - SAML login button: i2b2 SAML: Customize Identity Provider (IdP) Login Button
- User registration tool: 6.5a i2b2 Webclient User Registration
- Synthea-i2b2: Synthea-i2b2 Community Project, Synthea-i2b2 scripts and 63k sample file
- Summary of new optional config parameters in webclient: 1.4.2 Domain Configuration.

Issue Loading CRC Stored Procedures on MSSQL

Note that we have discovered an issue loading the CRC Stored Procedures on SQL Server and have found a workaround. Please be sure to read the updated content of Step 3.4.4 Create Crcdata Stored Procedures before performing this step.

1.7.13 Release Notes

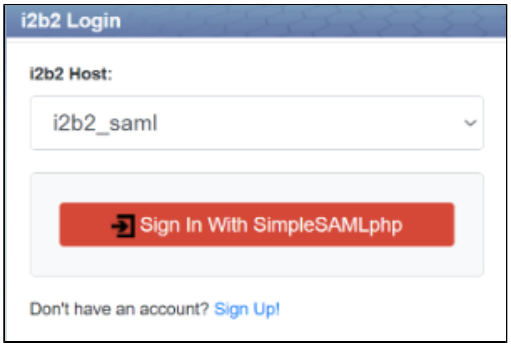
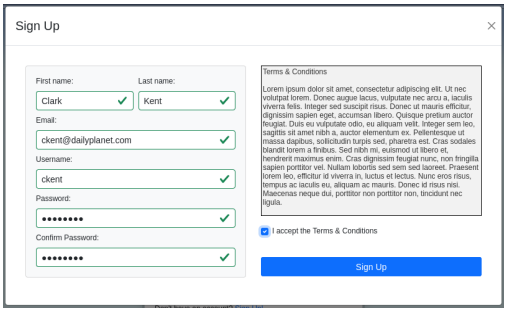
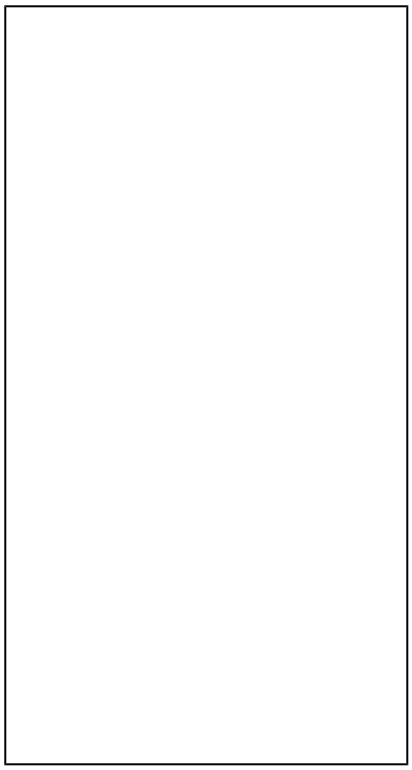
i2b2 1.7.13 offers support for SAML federated login, enhanced security due to improvements found via an internal Veracode scan, a client-based user registration tool, support for Synthea synthetic data for testing, and a variety of other bugfixes and performance improvements. i2b2 1.7.13 has been tested with SHRINE 3.2.1. SHRINE Approved

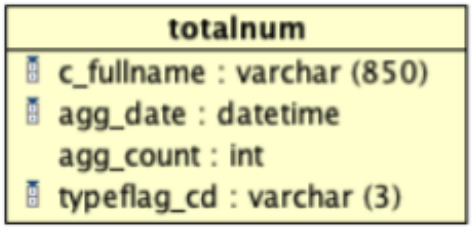
- 1.7.13 Release Notes
  - Highlight of Features
    - Top New Features
    - Community-Contributed Features

- Detailed Documentation on New Features
    - User Registration Tool
    - SAML Authentication
    - Improved Totalnum Scripts
        - Totalnum Scripts Setup
    - Additional New Stored Procedures
        - Age In Years Updater
        - Concept Dimension Updater
    - I2b2-Synthea data Load
        - Synthea Load Process
            - Loading Synthea data from scratch
    - ACT Version-4 Ontology data load
        - ACT4 data load process
    - Security Enhancements
    - Improved db Upgrade Process
- Changelog
    - Database Drivers
    - Supported Db Server versions
    - Supported software versions
- Supported Operating Systems
- i2b2 Server and Client Changes
    - New Features and Improvements
- i2b2 Database Changes
    - New Features and Improvements
    - Bug Fixes
- Notes for Developers

Highlight of Features

Top New Features

Description	
SAML Authentication	
User Account Registration Tool	
ACT Ontology v4	

Improved patient counting scripts ("totalnum")	
Synthea SyntheticMass dataset in i2b2 format	
Simplified database upgrade method	
log4J upgrade (to address security concerns)	
Code changes to address security vulnerabilities	
Bugfixes	

Community-Contributed Features

Contribution	Contributor	

SAML Authentication	Kevin Bui (lead developer) Michelle Morris **University of Pittsburgh** Justin Prosser (security expert) **University of Washington** Mike Mendis Jeff Klann Reeta Metta **Mass General Brigham**	i2b2 now includes support for SAML-based enterprise authentication via an institutional Identity Provider. See more information below.
Ability to specify user parameter defaults	Michael Horvath **Wake Forest University**	This change is meant to allow user params to take precedence over hive params. Currently, it's the other way around. Particularly, if you have the situation where you have a large number of users who use an authentication method other than the default basic, but your service account (AGG_SERVICE_ACCOUNT) is using basic then you need to specify a user param for each of your users. With this change, you can set default authentication params in pm_hive_params for all users, and then set your service account specifically to be authentication_method = 'BASIC'.
LDAP UPN Support	Michael Horvath **Wake Forest University**	Active Directory enables other methods of binding which are more flexible besides just using the distinguished name. https://docs.microsoft.com/en-us/openspecs/windows_protocols/ms-adts/6a5891b8-928e-4b75-a4a5-0e3b77eaca52. This change is to enable binding the the User Principle Name form, which is very convenient when the distinguished names for users is not easily available (OU by department, etc.).
API to get all children of an ontology node	Kevin Bui Michelle Morris **University of Pittsburgh**	The metadata GetChildren API call, which returns information on the children of an ontology node, can now be configured to return multiple levels of children (e.g., children, children's children, etc.). This is done by specifying the numLevel parameters. By default, the function assumes numLevel = 1 and will return the direct descendants of the concept, which is one level of children. When the numLevel = -1 the function will return ALL descendants of the concept, otherwise the function will return up to and including the number of levels specified by numLevel (eg. numlevel=2 returns two levels of descendants, numLevel=4 returns four levels of descendants).

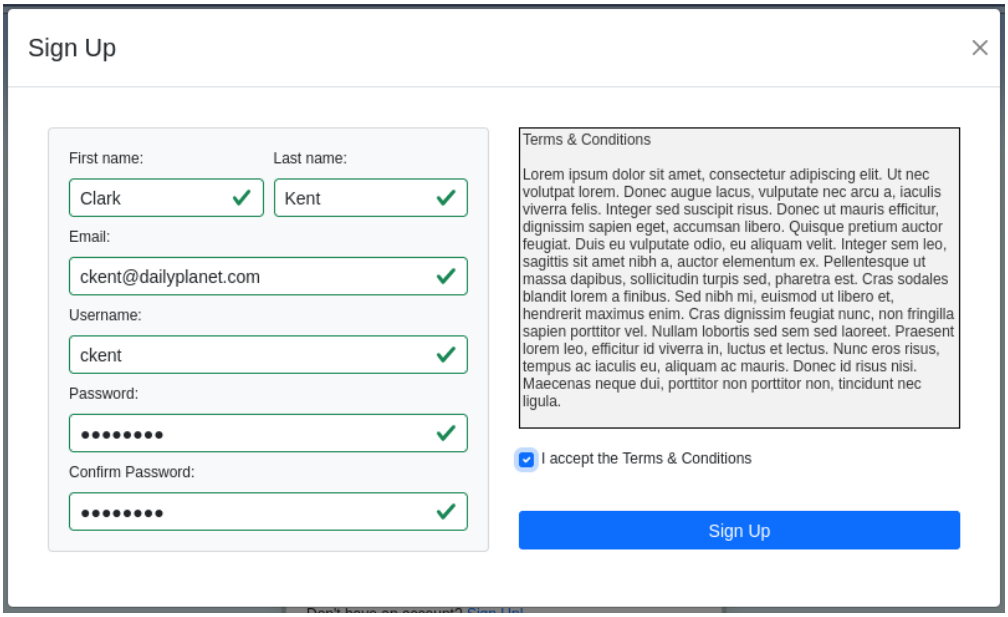
Totalnum Counter Performance Improvements	Darren Henderson **University of Kentucky**	Performance enhancements on SQL Server totalnum counting: stop unnecessarily recomputing temp tables.
ACT v4 Postgres bugfixes	**Ambreen Zaver** University of Colorado	Bugfixes in time interval calculation (for age and age-at-visit) in ACT v4 ontology for Postgres.

Detailed Documentation on New Features

User Registration Tool

There is a new user registration tool that can be enabled in the webclient. It allows users to request an i2b2 account that can then be activated by an administrator. It supports manual entry of user information through the form shown below, or automatic population of user information through SAML.

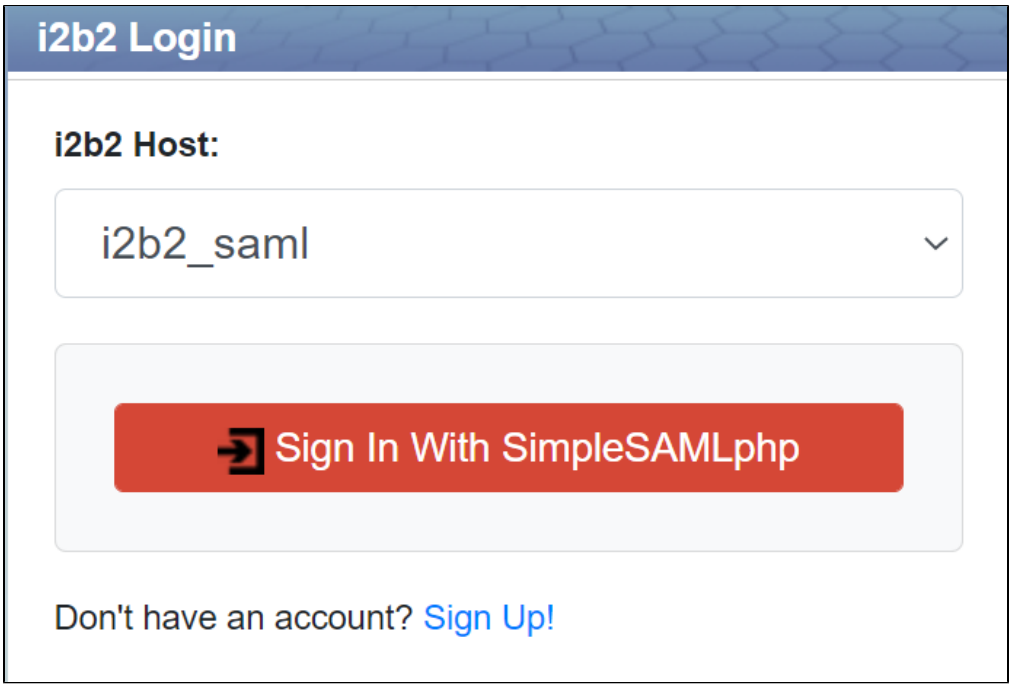
Documentation on this new feature is here: 6.5a i2b2 Webclient User Registration



SAML Authentication

i2b2 now includes support for SAML-based enterprise authentication via an institutional Identity Provider.

Detailed setup instructions are in Chapter 8 of the Installation Guide.



Improved Totalnum Scripts

Totalnum Scripts (patient counting scripts) have been updated to improve the counter's performance on both multiple ontology tables and very large(>1.5 million) ontologies such as ACT medications. Debug messages have also been added for troubleshooting and profiling purposes. Support for multiple fact tables has been added and bugfixes have been made.

Totalnum Scripts Setup

1. If upgrading, create the totalnum and totalnum_report tables. In Release_1-7/Upgrade/Metadata, run the ant upgrade script.
   ant -f data_build.xml upgrade_tables_release_1-7-12a
2. In the Release_1-7/NewInstall/Metadata/ run the ant script to create the stored procedures.
   ant -f data_build.xml create_metadata_procedures_release_1-7
3. Set privileges: If using multiple schemas, the stored procedure should be run from the *metadata* schema. Make sure the stored procedure can read the tables in the *crcdata* schema (observation_fact, visit_dimension, patient_dimension) and can both read an update ontology tables in the *metadata* schema (including table_access).
4. If using multiple fact tables, the recommended approach is to create a fact table view as the union of all your fact tables. (This is essentially going back to a single fact table, but it is only used for totalnum counting. This is needed to correctly count patients that mention multiple fact tables within a hierarchy.)

```
e.g.,
  create view observation_fact_view as
  select * from CONDITION_VIEW
  union all
  select * from drug_view

If running the counting script in SQL Server, add the wildcard flag, to ignore multifact references in
the ontology:
      e.g. exec RunTotalnum 'observation_fact_view','dbo','@','Y'
This is automatically accounted for in the other database platforms.
Note this approach does not work if you have conflicting concept_cds across fact tables.
```

5. Run the stored procedures on your database. This can be done in two ways:

   - Run the ant command to execute the data_build.xml file with below specified target
     POSTGRESQL : ant -f data_build.xml db_metadata_run_total_count_postgresql
     ORACLE : ant -f data_build.xml db_metadata_run_total_count_oracle
     SQL SERVER : ant -f data_build.xml db_metadata_run_total_count_sqlserver

   - Execute the RunTotalNum stored procedure manually against your database from a SQL Client. This can take several hours for large databases or large ontologies. Examples are below.

Oracle:	begin
	RUNTOTALNUM('observation_fact','i2b2demodata');
	end;
	
	You can optionally include a table name if you only want to count one ontology table (this IS case sensitive):
	begin
	runtotalnum('observation_fact','i2b2demodata','I2B2');
	end;
	
	Note: If you get the error as: ERROR at line 1: ORA-01031: insufficient privilege, then run the command:
	grant create table to (DB USER)

SQL server:	exec RunTotalnum 'observation_fact','dbo','@' Parameters are: 1) the observation table name (for multi-fact-table setups), 2) the schema name, 3) a single table name to run on a single ontology table or '@' to run on all, and 4) and a wildcard flag that will ignore multifact references in the ontology if 'Y'
PostgreSQL:	select RUNTOTALNUM('observation_fact','public') Replace 'public' by the schema name for the fact table If using a schema other than public for metadata, you might need to run "set search_path to 'i2b2metadata','public' " first as well

When finished, verify it is complete by checking that c_totalnum columns in your ontology tables contain numbers (not nulls). These total counts will be visible in the ontology browser in the web client.

Parent folders will get counts (of all patients with facts in the leaves) *except* for ontology folders derived from visit_dimension or patient_dimension. These cannot be rolled up because of the way these terms are defined in the ontology. They will have no count at all (not a zero).



i2b2 users must have the DATA_AGG user permission to view the counts through the web client.

Additional New Stored Procedures

Age In Years Updater

When the CRC data is installed via ant, a new SQL script updates the age_in_years_num in the patient dimension based on the birth dates of the sample patients. As a reminder, this load process can be triggered with ant -f data_build.xml db_demodata_load_data in the CRC directory of NewInstall.

Concept Dimension Updater

Insert_Concept_FROMTableAccess is designed to populate concept_dimenison table using the ontologies listed in table_access table records. The stored procedure loops through the table_access and inserts values from each metadata table (specified in the c_table_name column), when c_dimtablename is set to 'concept_dimension'
Example usage*: exec Insert_Concept_FROMTableAccess*

I2b2-Synthea data Load

Synthetic patient data generated by Synthea can be loaded into i2b2. The Synthea SyntheticMass sample files have been converted to i2b2-ACT format, and scripts to load Synthea data from scratch are available here: https://github.com/i2b2/i2b2-synthea

Synthea Load Process

1. Set up an i2b2 project with the ACT ontology.
2. Either download the SyntheticMass 63k sample in i2b2 format from https://github.com/i2b2/i2b2-synthea/blob/main/syntheamass_63K_sample.zip, or follow the instructions below to load any Synthea dataset from scratch. This information can also be found on the Synthea-i2b2 Community Project page.

Loading Synthea data from scratch

1. Download SyntheticMass Data, Version 2 (24 May, 2017)
    - All data sets (1k, COVID 10k, COVID 100k) have been verified to work EXCEPT the 100k patients in the large SyntheticMass Version 2 download.
    - The 100k patients in the large SyntheticMass Version 2 download needs an extra step to delete invalid records before import. In this case, download synthea_cleanup.pl to your disk, and then run "synthea_cleanup <directory-for-synthea-csv-files>" The fixed csv files will be in <directory-for-synthea-csv-files>/fixcsv.
2. Set up an i2b2 project with the ACT ontology.
3. Download the scripts from https://github.com/i2b2/i2b2-synthea
4. Run `create_synthea_table_<your dbServertype>.sql` in your project to create the Synthea tables.
5. Import the Synthea data you downloaded in step one into the Synthea tables in your project.
6. Load the i2b2-to-SNOMED table in this repository into your project. https://www.nlm.nih.gov/healthit/snomedct/us_edition.html
7. Click on the "Download SNOMED-CT to ICD-10-CM Mapping Resources" link to download. (You will need a UMLS account.)
8. Unzip the file
9. Import the TSV file into a table called SNOMED_to_ICD10 in your database.
10. In Postgres and Oracle, follow the additional instructions in the comments at the top of `synthea_to_i2b2_<your dbServerType>.sql` to clean up the date formatting.
11. Run `synthea_to_i2b2_<your dbServertype>.sql` to convert Synthea data into i2b2 tables (this will truncate your existing fact and dimension tables!)
12. Replace references to `i2b2metadata.dbo` in the script. Use the database and schema where your ACT ontology tables are.

ACT Version-4 Ontology data load

Metadata scripts are now available to load the latest ACT Version-4 ontology into your i2b2 db schema.

> ⚠ The CPT4 ontology table is not included with i2b2 due to AMA restrictions on redistribution of CPT code information. Contact the ACT team to get a copy if your institution is an AMA member.

ACT4 data load process

1. Download and extract the newinstall zip package from "Download Binary Distribution" in the top section of https://www.i2b2.org/software
2. Edit the edu.harvard.i2b2.data\Release_1-7\NewInstall\Metadata\db.properties file to update the project properties to 'ACT' ; db.project=ACT
3. From the edu.harvard.i2b2.data\Release_1-7\NewInstall\Metadata folder, run the ant command: ant -f data_build.xml db_metadata_load_data
    a. This will execute the SQL scripts from the edu.harvard.i2b2.data\Release_1-7\NewInstall\Metadata\act\scripts\<db type> folder and create and load ACT4 Ontology metadata tables
4. You can now verify the new Ontology by logging into the webclient.

Security Enhancements

1. i2b2 has been made more secure by addressing parameterization and other potential vulnerabilities found in an internal a Veracode scan.
2. Log4J has been upgraded to the latest version. The following jars are updated in lib/axis2.war/WEB-INF/lib the folder:
    - log4j-api-2.17.1.jar
    - log4j-core-2.17.1.jar
    - log4j-jcl-2.17.1.jar

Improved db Upgrade Process

Previously, i2b2 db upgrade was a multi-step process of running upgrade SQL scripts and stored procedures individually on the db instance. This release simplifies the process of running the table upgrade SQL scripts and stored procedures from data_build.xml files. Details are on the i2b2 Upgrade Page here.

Changelog

Database Drivers

The JDBC drivers were updated to the following versions.

Driver	ojdbc8.jar	postgresql-42.2.14.jar	mssql-jdbc-9.2.0.jre8.jar
New Version	Oracle 21.5	PostgreSQL 42.3.2	MS Sql Server 9.2

Supported Db Server versions

Server Type	SQL Server	Oracle	Postgres
Supported Version/s	2012+ (tested with up to 2019)	12g+ and 21c	9 to 14

Supported software versions

Application Type	Java	Wildfly	Apache HTD	Apache Ant	Apache Axis2	PHP
Supported Version/s	8 or 11	17.0.1Final	2.0 (RHEL 6) and 2.2 (RHEL 7)	1.9.6	1.7.1	7.2.27 or higher

Supported Operating Systems

CentOS versions 6 (deprecated) or 7 (highly recommended)

Windows 7-2019

Unofficially, MacOS and other flavors of Linux are likely to work.

i2b2 Server and Client Changes

New Features and Improvements

Core-server	webclient

CORE-399 Oracle index hints must use the table alias CORE-382 Username / password errors should not specify which had the problem CORE-402 Fix Veracode identified Security flaws in i2b2 Server-Side Code CORE-404 Adding support for JDK 11. Now including the gensrc due to jaxb has been removed. CORE-413 FetchAllChildren CORE-414 SAML (server side) CORE-415 Log4j upgraded to v2 CORE-416 User parameter precedence change (contributed by Michael Horvath) CORE-417 LDAP UPN support (contributed by Michael Horvath) CORE-405 Upgrade JDBC Drivers CORE-412 Disable login to agg service account	WEBCLIENT-334 Provide tabs for major plugins and temporal query WEBCLIENT-344 Cleanup Analysis Tools list of Plugins to only Supported Items WEBCLIENT-325 Wayne's improvements to hierarchical find-by-name WEBCLIENT-353 SAML and user registration tool (client side)

i2b2 Database Changes

New Features and Improvements

DATA-7 QT_PATIENT_SET_ENC_COLLECTION should be a bigint

DATA-14 Synthea i2b2 data

DATA-6 improve i2b2 db upgrade process

DATA-12 Stored procedure to update concept dimension

DATA-11 Age in years updater during demodata install

DATA-9 ACT v4 ontology

DATA-13 Postgres time interval corrections in ACT v4 demographics ontology

CORE-389 Totalnum performance improvements

CORE-394 Obfuscated totalnum reporting table

CORE-398 Multifact support for totalnums

CORE-400 Show totalnums in top level folders

Bug Fixes

Webclient	Core-server
WEBCLIENT-351 Obfuscated User Not Showing Graph	CORE-418 Local timestamps in sessions on Oracle - appserver /db server in diff time zones
WEBCLIENT-342 switch response status check from "OK" to 200 to handle lab value pop up in http/2 protocols	CORE-282 Error returned when obfuscated user is locked out
WEBCLIENT-335 temporal query in webclient with no anchoring events not running	CORE-281 Query continues to run after user receives lockout message
WEBCLIENT-350 Unable to drag items in workplace	CORE-399 use alias for index hint
WEBCLIENT-294 Webclient Reports "QUERY CANCELLED" While Query Is Still Running	
WEBCLIENT-354 Removed broken context menu in Find Previous Queries	

Notes for Developers

For Java 11 install, if you change the xsd (REST API message definitions), then you will need to regenerate gensrc via JAXB in Java 8. In the i2b2-core cell directory for which you're regenerating the XSD-Java, run the ant target "jaxb_gen" on Java 8 and then build as usual using Java 11.