

Programmatic Generation of c_metadataxml Values

Hubert Hickman
hhickman@essexmanagement.com



Introduction

Programmatic generation of `c_metadataxml` entries can greatly speed up the tailoring of laboratory results and vitals/findings data for deployment at an i2b2 site. Combining the information available for LOINC results along with data analysis of results in the source EMR system allows the creation of these entries in an automated fashion.

Information from LOINC

One key piece of information is available based upon the LOINC code of the result. The *scale type* defines what kind of result values are expected. Quantitative implies a measured result. Codes with an ordinal or nominal scale imply a list of choices. This scale drives which type of `c_metadataxml` result is created.

This information is available from the LOINC downloads from <http://www.loinc.org> and may also be contained within the EMR system (as is the case with Epic®).

Information from the EMR

The EMR of a given site contains the data needed to perform an analysis of the data for a given LOINC code.

For a given ordinal or nominal scale type, the list of choices can be analyzed from the data for the LOINC code.

For a given quantitative type, the analysis is more complex. If a site has normalized data in i2b2 then the units for a given LOINC code could be homogeneous. If a site has many different units for data for a given LOINC code, and chooses to not normalize data at ETL time via the use of the i2b2 server side data conversion (setting `CRC_ENABLE_UNITCD_CONVERSION = ON` for a given project), then many different units codes may exist for a LOINC code.

i2b2 supports several concepts with the `<UnitValues>` data in the `c_metadataxml` field:

NormalUnits This specifies the normally expected units. In the case of dynamic analysis for heterogenous units, the most common units code from the source data is heuristically chosen as the NormalUnits.

EqualUnits These are equivalent units to the normal units.

ExcludingUnits These units are not convertible from the NormalUnits

ConvertingUnits These units can be converted from the NormalUnits by multiplying the matching units by the specified `<MultiplyingFactor>`.

The application maintains a list of known equivalent and convertible units.

An analysis is also performed for the `Flagstouse` section of the `c_metadataxml`. Namely, the EMR data is analyzed for flag content. If high and low flags are noted, then the HL option is used for flags. If this is not the case and abnormal values are noted, then the A is chosen.

The `DataType` choice is made based upon an analysis of the data – whether decimal points occur in the numeric data or if the data is string data. Similarly, the range data is is heuristically chosen from the most common low and high ranges for the most common units noted in the EMR.

Architecture and current state

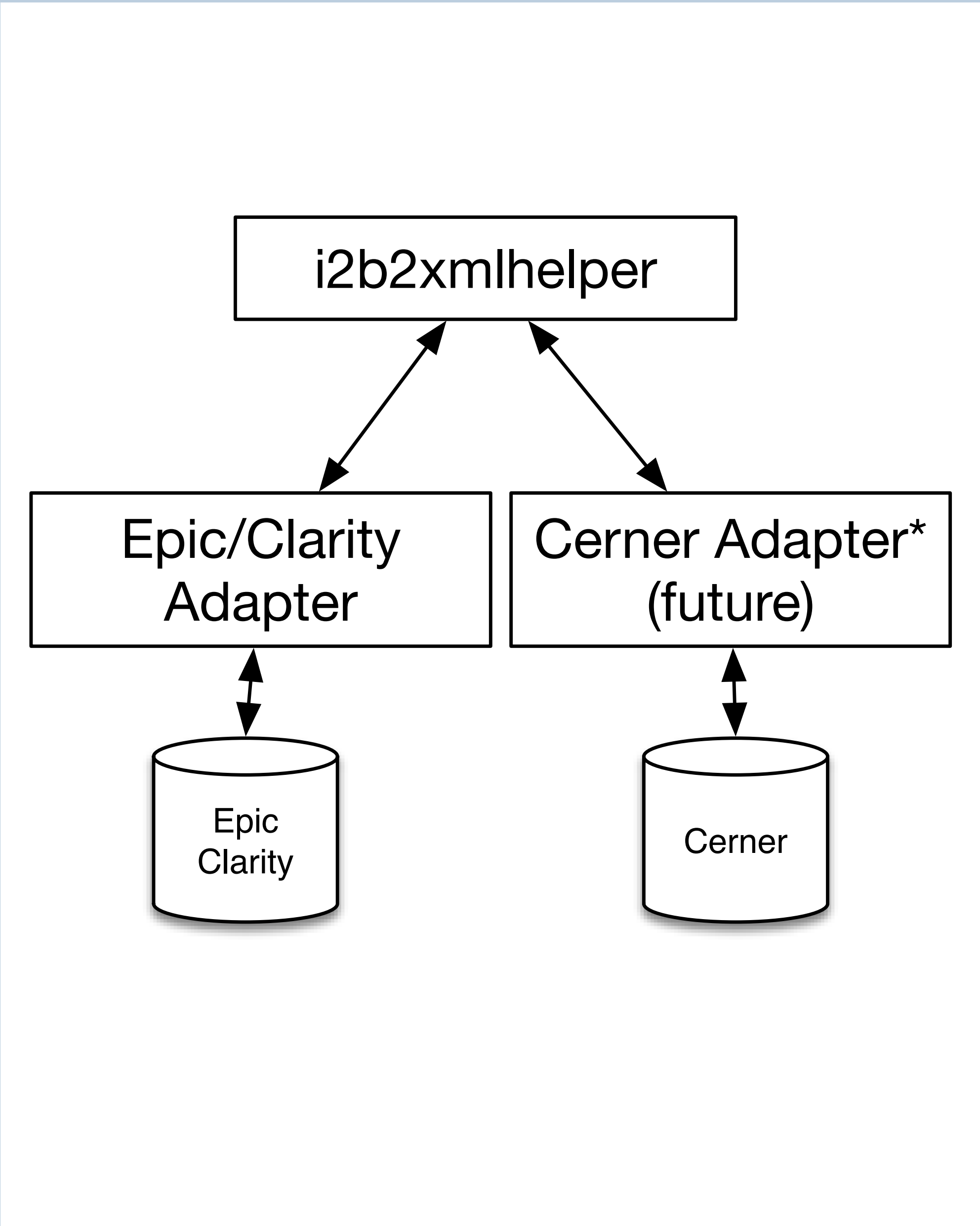


Figure 1: Application architecture

The application is architected in such a manner that will allow the separation of non-proprietary parts of the code from the vendor specific EMR adapters.

Results

The application has been used to generate over 1,200 `c_metadataxml` entries at a selected site.

If one estimates that to construct a given `c_metadataxml` entry manually takes at a minimum 15 minutes, then the time savings garnered from the user of an automated generation tool are quite significant.

Future work

The application needs to have adapters written for other EMR systems to allow a full spectrum of use across the i2b2 community. Continual refinement of the generated `c_metadataxml` entries will be made as experience is gained from its use.

It is anticipated that the code will be placed for use by the community under an open source license.