Contract Name: Characterize HL7 V2 Vocabulary Tables and V3 Value Sets

Summary:
Characterize and prioritize HL7 tables and value sets that are referenced in HL7 Version 2 and Version 3 messages. The characterization will be used as a guide for ordering the work of replacing HL7 table contents and value set contents with CHI approved terminologies.

The intent of the work is to accomplish the following two tasks as specified in the SOW from the NLM:

1) As a first step towards designating the specific parts of CHI standard vocabularies that are valid for particular HL7 message segments, ***develop a list of the HL7 message fields that can use or carry CHI standard vocabularies and assign priorities to the message fields based on such factors as current use, importance to federal programs, and other appropriate criteria.***

2) As a first step towards comparing appropriate HL7 lists of coded values to CHI standard vocabularies to determine if some of the HL7 lists can be replaced by subsets of a CHI standard vocabulary***, identify the HL7 code sets most likely to overlap with CHI standard vocabularies and assign priorities for the comparison work.***

These two tasks need to be completed for both Version 2 and Version 3 messages.

Task Deliverables:

Deliverables: One set of tables for V2 fields and one set of tables for V3 fields as described in detail below.

Produce a final report describing the project. Including process description, queries, measures and metrics developed, descriptions of all files created, methodologies and assumption used in making determinations of importance, issues with the CHI vocabularies and specifications, recommendations to NLM to aid in integrating content.

In addition, contractor will:

- participate in weekly NLM Vocabulary conference calls,
- keep a log of anamolies found while reviewing the value sets and report these findings to the HL7 vocabulary TC.

Required Skills: Expert knowledge of HL7 Version 2 and Version 3 messages, of Frank Oemig's database, and the V2 and V3 HL7 vocabulary tables. Skilled in use of MS Access databases and SQL.

Desired Skills: Familiarity with common use of HL7 messages. Familiarity with UMLS semantic types.

Expected start: End of November 2004
Estimated Hours: 100
Expected Duration: 1 month

Tools, queries, programs, and methods used will be explicitly documented. Intellectual property will enter the public domain under the ownership of HL7.

Payment will be dependent on approval of the deliverable and payment by NLM to HL7.

Task Description Detail for Version 2 Tables

Using the message database (Frank Oemig's database) for Version 2.5 of the HL7 standard as a starting point, create the following tables, adding in the values for the columns that evaluate the importance of each object.

Table 1: Importance of message

There will be one row for each Version 2 message. The table will contain the following columns:

1) Message Abbreviation
2) Message Description
3) Importance to EHR data exchange project
4) Importance to exchange of public health data
5) Importance in clinical data exchange

Each of the importance columns will use a numeric ranking system, as follows:

2 – Essential, very important, or in nearly universal use
1 – Needed, important, or in common use
0 – Nice to have, niche or uncommon use

Contractor will document the rational and method used to determine message importance.

Table 1 Example

Msg Abrev	Message Description	EHR	PH	Clin
ACK	General acknowledgment message	2	1	2
ADR	ADT response	0	0	1
ADT	ADT message	2	2	2
BAR	Add/change billing account	0	0	1
BPS	Blood product dispense status message	0	0	1

Table 2: Importance of a field in a segment

There will be one row for each coded field or subfield in each version 2 message segment. The table will contain the following columns:

1) Identifier
2) Name of field or subfield
3) Segment abbreviation
4) Sequence number in segment
5) Table number
6) Chapter reference
7) UMLS semantic type
8) UMLS semantic type description
9) Overlap with CHI approved terminology
10) Candidate CHI terminology
11) Importance of the field in a segment

The semantic type column will contain the abbreviation of a semantic type as defined in the UMLS Metathesaurus. The most specific semantic type that covers the content of the table will be used. For items that do not correspond to an existing semantic type will have one of the following values assigned

HL7S – HL7 Structural code table
HL7M – HL7 Messaging use only code table
HL7C – HL7 Clinical code table
HL7U – Unknown or unclassified semantic type

The overlap with CHI column will use a numeric ranking system, as follows:

2 – Known overlap
1 – Possible overlap
0 – No expected overlap

In the CHI terminology column, SCT is an abbreviation for SNOMED CT and HL7 means that the HL7 table itself is a CHI approved terminology. Abbreviations for other CHI terminologies will need to be established for use in the table.

The importance column will use a numeric ranking system, as follows:

2 – Essential, very important, or in nearly universal use
1 – Needed, important, or in common use
0 – Nice to have, niche or uncommon use

Table 2 Example

Id	Name	Seg	Seq	DT	Tbl	Chapter	ST	Descript	Ovrlp	CHI	Imp
576	Abnormal Flags	OBX	8	IS	78	7.4.2.8	T080	Qualitative Concept	2	HL7	2
639	Abnormal Text/Codes for Categorical Observatio	OM3	5	CE	9999	8.8.10.5	T080	Qualitative Concept	2	HL7	1

	ns										
1449	Academic Degree	EDU	2	IS	360	15.4.3.2	T065	Educational Activity	2	SCT	1
528	Accident Code	ACC	2	CE	506	6.5.9.2	T080	Qualitative Concept	2	HL7	2
814	Accident Death Indicator	ACC	6	ID	136	6.5.9.6	T080	Qualitative Concept	2	HL7	2
813	Accident Job Related Indicator	ACC	5	ID	136	6.5.9.5	T080	Qualitative Concept	2	HL7	1

Table 3: Overall Rank of Table

The information in Table 3 will be calculated from the values in Tables 1 and 2. There will be one row for each Version 2 vocabulary table. The table will contain the following columns:

1) Table number
2) Table type
3) Table description
4) UMLS Semantic type of table
5) Importance of table for EHR data exchange project
6) Importance of table for exchange of public health data
7) Importance of table for clinical data exchange
8) Evaluation of overlap with CHI
9) Abbreviation for possible CHI terminology replacement
10) Overall rank

The EHR, PH, and Clin columns represent the overall importance of the table within the particular context of use. The importance columns will be calculated based on the number of segments that use the table, the importance of the field in the segment, the number of messages that use the segment, and the importance of the message within the given context.

The rank column will contain an integer number starting at 1 and increasing to the number of tables. A rank of 1 means that this table is the most important table to begin comparing to CHI terminologies for potential replacement.

Table 3 Example

Type	Tbl #	Name	ST	EHR	PH	Clin	Ovrlp	CHI	Rnk
User	78	Abnormal flags	HL7C	2.00	2.00	2.00	2	HL7	35
User	129	Accommodation code	T083	1.12	0.00	1.87	2	HL7	125

User	117	Account Status	T051	0.58	0.00	1.16	1	SCT	219
HL7	183	Active/Inactive	T051	1.80	1.80	1.95	0	None	355
HL7	371	Additive/Preservative	T123	1.32	1.97	1.66	2	SCT	10
HL7	190	Address type	HL7C	1.46	1.78	1.72	2	HL7	77
HL7	164	Administration Device	T203	1.90	1.63	2.00	2	SCT	8

Task Description Detail for Version 3 Value Sets

Using the results of the "Vocabulary_HMD_rows" query from the RIM database provided by Woody Beeler, BAL8_RPComposite.mdb, which contains all the HMDs of all balloted messages, create the following tables, adding in the values for the columns that evaluate the importance of each object.

Table 1: Importance of message

There will be one row for each Version 3 message and CMET. The table will contain the following columns:

1) hmdId – hierarchical message definition identifier
2) hmdName – hierarchical message definition identifier
3) msgeIdentifier – message identifier
4) name – message name
5) Importance to EHR data exchange project
6) Importance to exchange of public health data
7) Importance in clinical data exchange

Each of the importance columns will use a numeric ranking system, as follows:

2 – Essential, very important, or in nearly universal use
1 – Needed, important, or in common use
0 – Nice to have, niche or uncommon use

Table 1 Example

hmdId	hmdName	msgeIdentifier	name	EHR	PH	Clin
COCT_HD010000	A_Encounter	COCT_MT010000	A_Encounter	2	1	2
COCT_HD020000	A_Appointment	COCT_MT020000	A_Appointment	1	0	2
COCT_HD030000	E_LivingSubject	COCT_MT030000	E_LivingSubject	2	2	2
COCT_HD030200	E_Person	COCT_MT030200	E_Person	2	2	2

FIAB_HD010100	Account Management	FIAB_MT010101	Patient Billing Account Event Activate	2	1	2
FICR_HD710100	Payment Advice Detail	FICR_MT710101	Payment Advice Detail	0	0	2
MCAI_HD700200	Trigger Event Control Act	MCAI_MT700201	Trigger Event Control Act	0	0	2
POBB_HD001000	Blood Bank Product Storage Event	POBB_MT001000	Blood Bank Product Storage Event	0	0	1
POBB_HD001021	BloodBankStockInput	POBB_MT001021	BloodBankStockInput	0	0	1
POLB_HD002100	Placer Order	POLB_MT002100	Placer Order	2	2	2
PORR_HD100001	CaseReportHMD	PORR_MT100011	CreateCaseReportMT	0	2	0

Table 2: Importance of a message domain in a RIM class

There will be one row for each rimClassName-rimAttributeName-msgDomain tuple. The table will contain the following columns:

1) rimClassName
2) rimAttributeName
3) rimDomain
4) msgDomain
5) rimStrength
6) UMLS semantic type
7) UMLS semantic type description
8) Candidate CHI terminology
9) Importance of message domain in class

The semantic type column will contain the abbreviation of a semantic type as defined in the UMLS Metathesaurus. The most specific semantic type that covers the content of the table will be used. For items that do not correspond to an existing semantic type will have one of the following values assigne

HL7S – HL7 Structural code table
HL7M – HL7 Messaging use only code table
HL7C – HL7 Clinical code table
HL7U – Unknown or unclassified semantic type

The overlap with CHI column will use a numeric ranking system, as follows:

2 – Known overlap
1 – Possible overlap
0 – No expected overlap

In the CHI terminology column, SCT is an abbreviation for SNOMED CT, LNC is an abbreviation for LOINC, and HL7 means that the HL7 table itself is a CHI approved terminology. Abbreviations for other CHI terminologies will need to be established for use in the table.

The importance column will use a numeric ranking system, as follows:

2 – Essential, very important, or in nearly universal use
1 – Needed, important, or in common use
0 – Nice to have, niche or uncommon use

Table 2 Example

rimClsNam	rimAttrNam	rimDomain	msgDomain	rimStr	ST	Descript	Ovrlp	CHI	Imp
Act	code	ActCode	ActCode	CWE	T051	Event	2	HL7	2
Act	priorityCode	ActPriority	ActPriority	CWE	T080	Qualitative Concept	1	HL7	2
Observation	targetSiteCode	ActSite	ActSite	CWE	T017	Anatomic Structure	2	SCT	2
Act	statusCode	ActStatus	ActStatus	CNE	HL7C	HL7 Clinical Code	2	HL7	2
Act	confidentialityCode	Confidentiality	Confidentiality	CWE	T080	Qualitative Concept	2	SCT	1
Observation	methodCode	ObservationMethod	Observation Method	CWE	T060	Diagnostic Procedure	2	SCT	2
Act	code	ActCode	Observation Type	CWE	T034	Laboratory or Test Result	2	LNC	2
SubstanceAdministration	routeCode	RouteOfAdministration	RouteOfAdministration	CWE	HL7C	HL7 Clinical Code	2	SCT	2
Act	code	ActCode	SubstanceAdministration ActCode	CWE	HL7C	HL7 Clinical Code	2	SCT	1

Table 3: Overall Rank of a message Domain

The data in the columns in this table are created mainly from the information in tables 1 and 2, with the addition of a rank column. There will be one row for each Version 3 message domain. The table will contain the following columns:

1) msgDomain
2) UMLS Semantic type of domain
3) Importance of table for EHR data exchange project
4) Importance of table for exchange of public health data
5) Importance of table for clinical data exchange
6) Evaluation of overlap with CHI

7) Abbreviation for possible CHI terminology replacement
8) Overall rank

The EHR, PH, and Clin columns represent the overall importance of the domain within the particular context of use. The importance columns will be calculated based on the number of classes that use the domain, the importance of the attribute within the class, the number of messages that use the class, and the importance of the message within the given context.

The rank column will contain an integer number starting at 1 and increasing to the number of tables. A rank of 1 means that this table is the most important table to begin comparing to CHI terminologies for potential replacement.

Table 3 Example

msgDomain	ST	EHR	PH	Clin	Overlap	CHI	Rank
ActCode	T051	2.00	2.00	2.00	2	HL7	5
ActPriority	T080	1.12	0.00	1.87	1	HL7	125
ActSite	T017	2.00	2.00	2.00	2	SCT	10
ActStatus	HL7C	1.46	1.78	1.72	2	HL7	77
Confidentiality	T080	1.90	1.63	2.00	2	SCT	8
ObservationMethod	T060	1.80	1.22	1.76	2	SCT	21
ObservationType	T034	2.00	2.00	2.00	2	LNC	2
ParticipationMode	HL7C	0.84	0.00	1.67	0	HL7	232
RouteOfAdministration	HL7C	2.00	2.00	2.00	2	SCT	4
SubstanceAdministrationActCode	HL7C	1.00	1.23	1.45	1	SCT	99