

一般口演 | 第40回医療情報学連合大会（第21回日本医療情報学会学術大会） | 一般口演

一般口演10 標準化

2020年11月20日(金) 09:00 ～ 10:57 G会場 (イベントホール・特設会場2)

[3-G-1-05] HL7 FHIRと SS-MIX2標準化ストレージのマッピング検証

*田中 勝弥¹、山本 隆一²（1. 国立がん研究センター, 2. 医療情報システム開発センター）

*Katsuya Tanaka¹, Ryuichi Yamamoto²（1. 国立がん研究センター, 2. 医療情報システム開発センター）

キーワード：Secondary Use, Standardization, HL7 FHIR, SS-MIX2

【はじめに】 広く普及している SS-MIX2標準化ストレージは、膨大な数の HL7 v2 Messageで構成されていること、患者を起点としたディレクトリ構造となっていることから、臨床研究等を目的とした患者横断的なデータセットの抽出には不向きであり、また、二次利用において HL7 v2のメッセージ定義への予備知識が求められる。そこで、直感的な二次利用向けのユーザインターフェイスを実現するためのアプローチとして、HL7 FHIRを用いたインターフェイスを試作している。これまでに開発したデータ検索可能な SS-MIX2ストレージに連携したデータセット抽出機能の開発を試みる過程で、インターフェイスとデータソース間の規格違いからマッピングが必要となった。

【方法】 HL7 FHIR Resourceと SS-MIX2標準化ストレージとの互換性の程度を評価するために、10の HL7 FHIR Resourceと SS-MIX2間のマッピングを実施・検証した。検証に使用したバージョンは、HL7 FHIR：R 4.0.0、SS-MIX2標準化ストレージ：Ver 1.2c、とした。

【結果】 FHIRを起点として、Resource Propertyで評価すると、282件中138件(49%)と件数的にはマッピング率は低い。しかしながら、検査項目コード、結果値、処方薬剤コード、数量、などの主要なデータ項目についてはマッピング可能なことが確認できた。なお、HL7 FHIR、SS-MIX2双方で未使用項目として定義されている項目は、マッピング成績には影響しない。

【おわりに】 既存の標準化ストレージを活用可能な基盤が実証されれば、過去のリソースを低コストで有効利用できる。現時点では、マッピング率の低さから、HL7 FHIR Resource、SS-MIX2標準化ストレージ間の記述規格変換において欠落してしまう項目の拡充が今後の課題となる。

HL7 FHIR と SS-MIX2 標準化ストレージのマッピング検証

田中 勝弥^{*1}、山本 隆一^{*2}

^{*1} 国立がん研究センター、^{*2} 医療情報システム開発センター

An Investigation of Message Definition Mappings between HL7 FHIR and SS-MIX2 Standardized Storage

Katsuya Tanaka^{*1}, Ryuichi Yamamoto^{*2}

^{*1} National Cancer Center Japan, ^{*2} Medical Information System Development Center

A secure data collection infrastructure was implemented in which Health Level-7 (HL7) Fast Healthcare Interoperable Resources (FHIR) was applied according to standardized electronic medical record (EMR) storage. This infrastructure aims to facilitate rapid secondary use of EMR data in cross-institutional analyses on the basis of the Standardized Structured Medical Information eXchange (SS-MIX), Japan's domestic standard for EMR export. Existing EMR storage comprised enormous numbers of HL7 v2 messages; therefore, the user interface and database structure are basically defined according to the HL7 v2 data types and message structures, causing difficulty in rapid extraction for researches by clinicians. To solve this problem, we are trying to enhance some user interface with HL7 FHIR, which needs SS-MIX items to be mapped through HL7 FHIR resources. We investigated definition gaps between HL7 FHIR and SS-MIX messages, and report a developed user interface with HL7 FHIR to query against existing storage.

Keywords: Secondary Use, Standardization, HL7 FHIR, SS-MIX2

1. はじめに

広く普及している SS-MIX2 標準化ストレージは、膨大な数の HL7 v2 Message で構成されていること、患者を起点としたディレクトリ構造となっていることから、臨床研究等を目的とした患者横断的な EMR データセットの抽出には不向きであり、また、二次利用において HL7 v2 のメッセージ定義への予備知識が求められる。我々はこれまでに、SS-MIX2 標準化ストレージを対象とした、多施設横断・分散型の安全なデータ収集基盤を開発してきており、ユーザインターフェイスとして、横断検索を行う機能については検討が必要となっている。本稿では、SS-MIX2 標準化ストレージに対する直感的な二次利用向けのユーザインターフェイスを実現するためのアプローチとして、HL7 FHIR を用いたインターフェイスを試作を目的とした検証を行った。これまでに開発した多施設横断なデータ検索可能型の SS-MIX2 ストレージに連携し、拡張インデックスを利用しつつ、効率的かつ簡便に標準化ストレージからデータセットの抽出機能開発を試みているが、インターフェイスとデータソースの規格の違いからマッピングが必要となった。マッピング結果の概要と性能評価結果を報告する。

2. 方法

本システムでは、SS-MIX2 標準化ストレージ上のデータを一般的な NFS や CIFS 等のファイルシステムとして利用可能なインターフェイスは確保しつつ、同時に RDBMS (Relational Database Management System) 内のデータとして検索可能なインターフェイスを実現している。2つのインターフェイスは、FUSE (Filesystem in Userspace) を採用することで可能であり、現時点では、PostgreSQL と pgfuse の組み合わせにより実装を行っている¹⁾。PostgreSQL に格納するデータは、SS-MIX2 標準化ストレージで規定される HL7 v2 メッセージを対象としており、HL7 v2 で定義されたセグメント、フィールド単位にデータベース内のテーブル構造を構成した。このため、格納されたデータを検索するためには、HL7 v2 のメッセージ定義、

構造を意識したクエリを作成する必要があり、ユーザインターフェイスとしては専門的知識を要する。より、直感的な Term で検索を行うために、ユーザインターフェイスとして HL7 FHIR の Vocabulary を実装することを試みる。

図1にシステム構成の概要を示す。HL7 FHIR によるラッピングには、HL7 HAPI Library を採用した。HAPI Model Object によって、両規格のコンバートが可能であるが、必要とするクエリに応じて、コンバータ部分は実装する必要がある。

HL7 FHIR Resource と SS-MIX2 標準化ストレージ HL7 v2 Message との互換性の程度を評価するために、10 の HL7 FHIR Resource と HL7 v2 Message Field 間のマッピング作業を実施、検証した。検証に使用したバージョンは、それぞれ、検証時の最新版である HL7 FHIR: R 4.0.0、および、国立がん研究センターに導入されている SS-MIX2 標準化ストレージ: Ver 1.2c、とした。

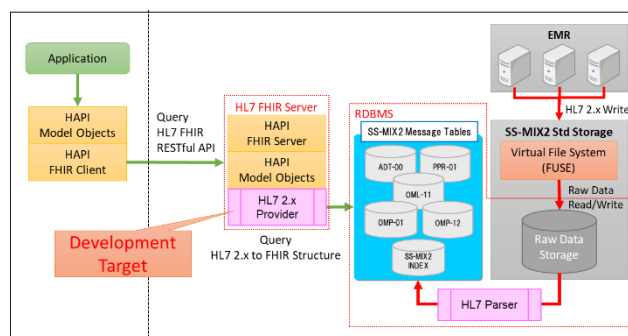


図 1 システムの概要(出展:1))

3. 結果

3.1. ユーザインターフェイス

図2に試作した SS-MIX2 標準化ストレージ用の検索画面

の概観を示す。上部は検索条件を HL7 FHIR のデータ項目で指定するエリアとなっており、下部は検索により抽出するデータ項目を同様に HL7 FHIR の Resource ごとのプロパティで指定するエリアとしている。

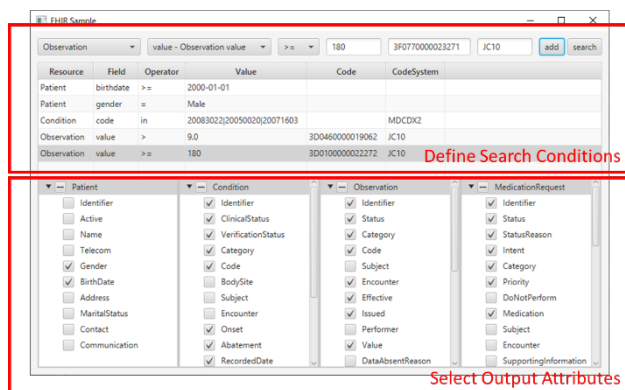


図 2 検索インターフェイスの概観

3.2. マッピング結果

FHIR を起点として、Resource Property で評価すると、282 件中 138 件(49%)と件数的にはマッピング率は低い。しかしながら、検査項目コード、検査結果値、処方薬剤コード、数量、などの主要なデータ項目についてはマッピング可能なことが確認できた。なお、HL7 FHIR で提供されているマッピング案として、Unused とされている項目が多数存在するが、各リソースのより詳細な情報は実際には、HL7 v2 Message で定義されているデータ項目も含まれており、マッピングは可能にみえる。ただし、多くの場合は SS-MIX2 標準化ストレージでは、未使用項目とされており、マッピング成績には影響していない。

表 1 HL7 FHIR Resource 起点でのマッピング結果

No.	FHIR Resource	Property	Mapped	Ratio(%)
1	Patient	24	14	58%
2	Encounter	40	11	25%
3	Condition	20	12	60%
4	AllergyIntolerance	22	13	59%
5	Observation	33	20	60%
6	Specimen	27	7	25%
7	ServiceRequest	33	17	51%
8	MedicationRequest	40	20	50%
9	MedicationDispense	28	16	57%
10	Dosage	15	8	53%

3.3. データマッピングの外部定義

HL7 FHIR の Resource 定義は継続的に検討がなされており、今後より精緻化されていくと考えられる。ラピッドプロトタイピングを実現するために、今回、HL7 v2 定義とのマッピング表については、ハードコーディングとせず、外部定義として取り込む構成が可能なよう設計を行った。

3.4. 検索性能

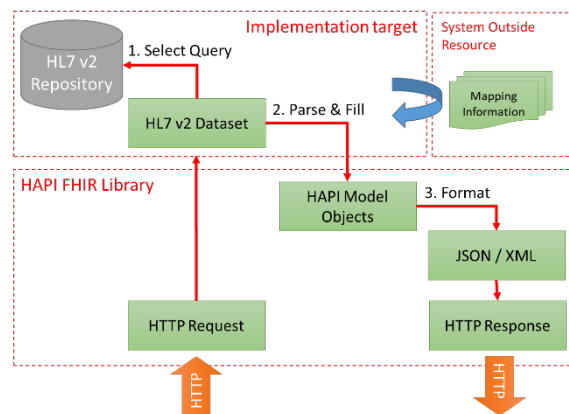


図 3 マッピング情報の動的利用

100 万レコードを越える規模の検索性能は悪化することが確認された。HL7 v2、HL7 FHIR はそれぞれ、Segment、Resource といった単位でのメッセージ構造が基本となっており、全件検索においては、複数の Segment、Resource をまたぐ検索は高コストとなる。性能改善においては今後の検討課題としたが、全件データベースから、目的ごとに特化した検索用サブテーブルを都度作成し、維持・管理する必要があると考えられる。

4. 考察

データ項目の両規格間のマッピング可能な割合としては、予想を下回ったが、処方、検体検査、病名などの主要な項目は対応がとれており、ユーザインターフェイスを変えてもデータ検索・抽出が可能なることが確認できた。従前からの課題でもあるが、内部構造として、HL7 v2 Segment および HL7 FHIR Resource をまたぐ検索はデータの肥大化とともに検索性能が悪化する。データ保持機構として、二十保持にならないような検索用スキーマの定義、メンテナンス機構を今後検討したい。

5. おわりに

既存の標準化ストレージを活用可能な基盤が実証されれば、過去のリソースを低コストで有効利用できる。現時点では、マッピング率の低さから、HL7 FHIR Resource、SS-MIX2 標準化ストレージ間の記述規格変換において欠落してしまう項目の拡充が今後の課題となる。

なお、本研究は、JST CREST グラント番号 JPMJCR1404 「ビッグデータ統合利活用促進のためのセキュリティ基盤技術の体系化」により実施した。

参考文献

- 1) Tanaka K, Yamamoto R, Nakasho K, Miyaji A. Development of a Secure Cross-Institutional Data Collection System Based on Distributed Standardized EMR Storage. Stud Health Technol Inform. 2018;255:35-9.