

ポスター

ポスター6

医療データ解析・地域医療

2017年11月22日(水) 08:45 ～ 09:35 L会場（ポスター会場2）（12F ホワイエ）

[3-L-1-PP6-3] SS-MIX2ストレージ上のデータ抽出プログラムの試作

乙竹 秀明¹, 中村 直毅², 中山 雅晴³（1.東北大学医学部, 2.東北大学大学院 医学系研究科 情報基盤室, 3.東北大学大学院 医学系研究科 医学情報学分野）

【目的】

標準化された形式で診療情報を蓄積・共有する SS-MIX2(Standardized Structured Medical Information exchange version.2)のストレージには、HL7(Health Level Seven)形式やXML形式からなるファイルが配置されている。SS-MIX2ストレージでは、診療情報がフォルダー構造の膨大なファイルから構成されており、SS-MIX2ストレージの中から必要な情報を簡単に抽出・検索することは容易でない。そこで、本稿では、SS-MIX2ストレージに蓄積されている膨大なファイルの中から、容易に診療情報を検索・抽出する仕組みの確立を目的とする。

【方法】

東北大学病院の SS-MIX2ストレージサーバに蓄積された SS-MIX2形式の診療情報を対象とし、標準化ストレージ上の HL7ファイルから、情報を抽出するプログラムを試作する。さらに、ウェブアプリケーションとして情報を抽出するプログラムを実装し、多数のファイルの中から情報を検索・可視化するプログラムを試作する。

【結果・考察】

SS-MIX2ストレージ中の HL7ファイルの抽出し、任意の患者の情報を、JSON形式で情報を管理するプログラムを実装した。JSON形式でタグを付与して情報を管理し、検索や処理が容易に行うことができることを確認した。また、ウェブアプリケーション上で複数の患者の診療情報を可視化するプログラムを実装した。MMWINの SS-MIX2ストレージ上のテストデータを用いて実装したプログラムの正常性を確認するとともに、蓄積された情報を表形式で抽出・一覧表示をすることができ、診療情報を容易に確認することができることを確認した。

【まとめ】

本稿では、SS-MIX2ストレージ中の HL7ファイルからデータを抽出し、可視化するアプリケーションを試作し、正常に動作することを確認した。現在は、必要な情報をオンデマンドで SS-MIX2ストレージからデータを抽出するシステムとなっているが、抽出した情報をデータベース等に蓄積し、効率的に情報を抽出・可視化するシステムの実現が今後の課題である。

SS-MIX2 ストレージ上のデータの抽出プログラムの試作

乙竹秀明^{*1}、中村直毅^{*2}、中山雅晴^{*3}

^{*1} 東北大学医学部医学科、^{*2} 東北大学大学院医学系研究科情報基盤室、

^{*3} 東北大学大学院医学系研究科医学情報学分野

Developing Python Programs to Get Medical Information from Standardized Storages of SS-MIX2

Hideaki Ototake^{*1}, Naoki Nakamura^{*2}, Masaharu Nakamura^{*3}

^{*1} Tohoku University School of Medicine,

^{*2} Tohoku University Graduate School of Medicine Information Infrastructure Office,

^{*3} Tohoku University Graduate School of Medicine Medical Informatics

It is difficult to get medical information from SS-MIX2 used to share medical information among medical institutions, because it is complexly standardized. Here we develop python programs to get medical information easily from it. HL7 files have hierarchical structure of fields, and standardized storage of SS-MIX2 has an enormous amount of HL7 files hierarchically, so these make it difficult to know what information each field have and where HL7 file you want to get is in the storage. We devise an algorithm to find HL7 files from the storage and to convert HL7 files to JSON files whose information is tagged. Then, we make two python programs, using these algorithms. One of them is able to get all valid HL7 files from one patient, and convert and gather to one JSON files. This program makes it easier to search and process patient's information. The other one is able to gather some information from many patients and put them into a table on WEB application. This program is useful when you compare patients and manage them. JSON files can be processed by database, however it is necessary to process HL7 more efficiently, because it takes more time. These two programs will help you to process and make use of SS-MIX2 and get information you want from a huge amount of complex information of SS-MIX2.

Keywords: SS-MIX2, HL7, Data Extraction, Data Display, Database Management System.

1. 緒論

医療分野における IT の発展に伴い、多くの医療機関において電子カルテシステムが導入され、膨大な医療情報が電子的に蓄積されるようになった。蓄積された医療情報を医療機関同士で共有・活用する地域医療連携への展開や臨床研究への応用などが期待されている。これらを円滑に推進するため、医療機関の保持している診療情報を、標準化された形式で蓄積・共有する、厚生労働省標準規格である SS-MIX2(Standardized Structured Medical Information eXchange version.2)ストレージの利用が推進されている。

2. 目的

SS-MIX2 のストレージには、標準化ストレージと拡張ストレージの 2 つが存在する。標準化ストレージでは、患者基本情報、病名、処方、注射、検体検査、生理検査などの限られた種別の診療情報が標準化された形式で蓄積され、拡張ストレージでは、様々な形式の診療情報が蓄積されている。SS-MIX2 ストレージでは、HL7(Health Level Seven)形式や XML 形式からなる診療情報が記録された膨大なファイルがフォルダー構造で配置されており、SS-MIX2 の詳細なルールを理解している IT 技術者以外の人では、SS-MIX2 ストレージから診療情報を簡単に抽出・検索することは困難である。本稿では、SS-MIX2 ストレージに蓄積されている膨大なファイルの中から、診療情報を容易に検索・抽出する仕組みを実現することを目的としている。

3. 方法

本稿では、東北大学附属病院の SS-MIX2 ストレージの標準化ストレージに蓄積された診療情報を対象とし、標準化ストレージの大部分を占める HL7 形式(図1)のファイルから、診

療情報を抽出・可視化するプログラムを Python というプログラミング言語を用いて試作した。

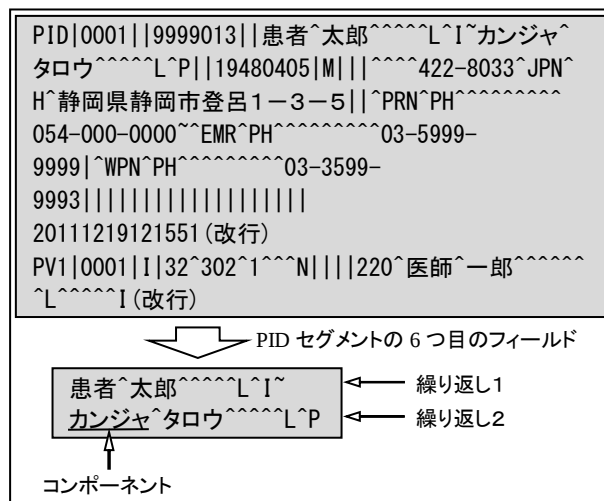


図1 HL7 ファイルのテキストの例

HL7 ファイルから診療情報を抽出するために、HL7 ファイルの構造を解析し、検索可能な形に変換するプログラムを作成した。

HL7 ファイルは、CSV ファイルのように特定の文字を区切り文字とする記述と、XML ファイルのように情報の複雑な階層構造を表現できる性質を合わせ持ったテキストファイルである。

HL7 ファイルはセグメント、フィールド、繰り返し、コンポーネントの 4 層の階層を構成し、それぞれ「改行」、「|」、「~」、「^」を区切り文字として区切られている。各セグメントの最初のフィールドにはセグメント名が付与されていて、フィールドの記録

内容を表す。各コンポーネントに記録されている情報の内容は、セグメント名ごとに決まっている。また、HL7 ファイルにどのセグメントが記録されるかは HL7 ファイルのメッセージ型によって決定する。メッセージ型は患者基本情報や入院予定、処方オーダーなど、診療情報の内容ごとに決められおり、SS-MIX2 標準化ストレージのデータ定義¹⁾と JAHIS (一般社団法人 保健医療福祉システム工業会) データ交換規約²⁾の定義に従っている。そこで、これらの定義に従って、ファイル名やセグメント名から各コンポーネントに記録されている内容に対応させ、SS-MIX2 標準化ストレージのフォルダー構造から任意の患者の診療情報を JSON(JavaScript Object Nation)形式で出力する。(図 2)

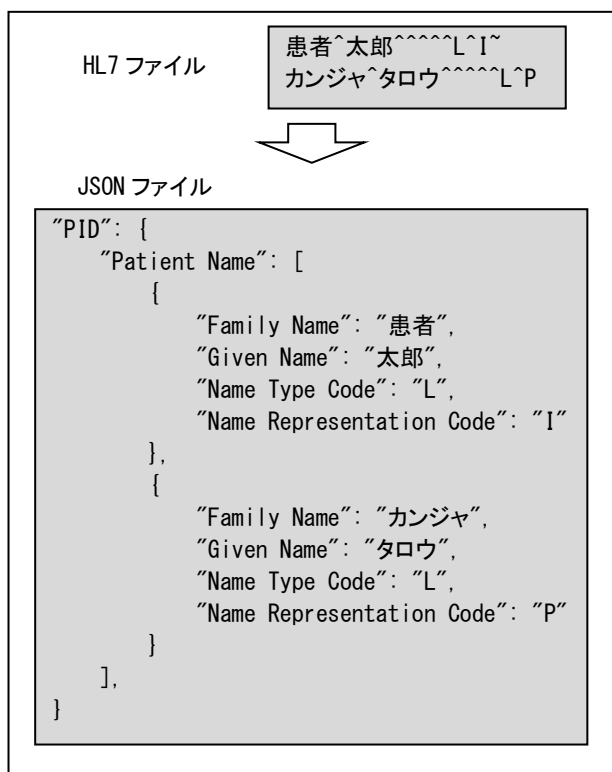


図 2 HL7 ファイルの JSON 出力例

抽出した情報を用いて、患者のすべての診療情報を JSON ファイルに出力するプログラムを作成した。

さらに、指定した患者 ID の範囲内のすべての患者を参照し、特定の診療情報を抽出し、表として出力するウェブアプリケーションとして実装した(図 3)。

4. 結果

作成したプログラムを使って、東北大学病院および宮城県地域医療連携システム MMWIN(Miyagi Medical and Welfare Information Network)に蓄積されている SS-MIX2 標準化ストレージ上のテストデータを用いて試作したプログラムの正常性を確認した。

最初に、HL7 ファイルを解析し JSON ファイルとして出力する仕組みは全てのテストデータで正常に機能した。HL7 ファイルに記録された全ての情報一つひとつに正しくタグを付与することができていることを確認した。さらに、HL7 ファイルを SS-MIX2 標準化ストレージのフォルダーの中から特定する仕

SS-MIX2 診療情報抽出ツール

患者IDの設定

start: 9990000 end: 9999999

項目の設定

項目01: データ種別: 患者基本情報 検査項目: 患者氏名(性)
 項目02: データ種別: 外来診療の受付 検査項目: 診療科コード
 項目03: データ種別: 外来診療の受付 検査項目: 患者氏名(性), 患者氏名(姓), 患者氏名(カナ), 患者氏名(カナ), 患者生年月日, 患者性別, 患者郵便番号, 患者住所, 患者電話番号, 診療科コード, 来院日時

セグメント設定

新規登録・変更

データ種別設定

新規登録・変更

登録データ種別一覧

No.	表示名	データ種別
01	患者基本情報	ADT-00
02	外来診療の受付	ADT-12
03	入院票	ADT-22
04	処方オーダー	OMP-01
05	検査検査結果	OMI-11
06	注射オーダー	OMP-02

SS-MIX2 診療情報抽出ツール

戻る

No.	ID	患者氏名(性)	診療科コード	来院日時
1	9990018	ヤマウチ	39	20160923134659
2	9990021	イシガミ	39	20160923135301
3	9990034	木村	39	20160923140343
4	9990047	ハヤシ	39	20151031112355
5	9990050	ホウジョウ	39	20160923135406
6	9990063	サクマ	39	20160923135557
7	9990076	北条	39	20160923135727
8	9990089	サトウ	39	20160923135717
9	9990092	佐藤	39	20160923225324
10	9990106	村田	39	20160923135027

図 3 ウェブアプリケーションのスクリーンショット
上の 3 枚が診療情報を抽出する設定画面、下の 1 枚がテストデータから診療情報を抽出した画面である。

組みもテストデータのストレージに対して正常に動作することを確認した。これによって、膨大な数のフォルダーとファイルが存在する SS-MIX2 標準化ストレージから必要な情報が記録されている HL7 ファイルを容易に抽出することができることが確認できた。

さらに、複数の患者の診療情報を表で可視化するウェブアプリケーションを通して、患者氏名や生年月日など設定によってさまざまな情報を抽出し一覧にすることができることを確認した。

以上より、東北大学病院および MMWIN の SS-MIX2 ストレージ上のテストデータを用いて実装したプログラムの正常性を確認するとともに、蓄積された情報を表形式で抽出・一覧表示をすることができ、診療情報を容易に確認ができることを確認した。

5. 考察

複数の患者の診療情報を表で可視化することで、患者間での診療情報の比較が可能となり、病院での患者の管理や

臨床研究などに応用できると考えられる。また、HL7 ファイルを JSON ファイルとして出力することが可能となることで人が直接見て意味が読み取れるだけでなく、MongoDBをはじめとする NoSQL を処理できるソフトウェアに取り込むことでさらに簡単に診療情報の検索や処理することが期待される。しかしながら、HL7 ファイルの複雑な階層構造は JSON 形式になったとしても扱いやすいものではないため、この階層構造を容易に処理するたる仕組みを引き続き検討している。また、患者一人の診療情報全てを JSON ファイルで出力することができたが、一人の患者の診療情報は通院が長くなるほど膨大になるため、ファイルの参照方法や HL7 ファイルの解析方法や処理の効率化が必要である。

6. 結論

本稿では、SS-MIX2 ストレージ中の HL7 ファイルからデータを抽出し、可視化するアプリケーションを試作し、正常に動作することを確認した。現在は、必要な情報をオンデマンドで SS-MIX2 ストレージからデータを抽出するシステムとなっているが、抽出した情報をデータベース等に蓄積し、効率的に情報を抽出・可視化するシステムの実現が今後の課題である。

参考文献

- 1) 日本医療情報学会, SS-MIX2 標準化ストレージ仕様書 Ver.1.2d データ格納方法およびデータ定義. 日本医療情報学会, 2016.
- 2) 一般社団法人 保健医療福祉システム工業会 医療システム部会 相互運用性委員会. JAHIS データ交換規約(共通編) Ver.1.1. 一般社団法人 保健医療福祉システム工業会, 2015.
- 3) 日本医療情報学会, SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2d. 日本医療情報学会, 2017.
- 4) みやぎ医療福祉情報ネットワーク(Miyagi Medical and Welfare Information Network) <http://www.mmwin.or.jp/>