

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 11 日現在

機関番号：14301

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24590614

研究課題名(和文) 診療情報の意味論的解析を行う自律エージェントアルゴリズムの研究

研究課題名(英文) Autonomic agent algorithm to analyze semantics on clinical information

研究代表者

小林 慎治 (KOBAYASHI, SHINJI)

京都大学・情報学研究科・特定講師

研究者番号：20423519

交付決定額(研究期間全体)：(直接経費) 4,200,000円

研究成果の概要(和文)：研究全体を通じて医療分野での標準規格およびターミノロジーについて精査を行い、相互運用性について検討を行った。現在、国際的協調として実施されている臨床情報モデルの統一的構築の基盤となっている ISO 13606規格を軸に多言語対応、複数ターミノロジー対応を行う中間情報モデルを開発した。この中間モデルを軸に自然言語処理の手法を用いて項目名、用語間の近さを判定して大学病院間の検査マスタを突合せさせることを試みた。

研究成果の概要(英文)：We designed interim information archetype by ISO 13606 methodology and implemented an autonomic agent to bind each item name to the interim archetype by score of natural language processing.
Terminology mapping: We examined this agent to bind laboratory test terminology with the other hospitals. Although most of them were binded with JLAC10 terminology, the agent found multiple bindings to same code and not efficient to bind them. After cleansing data, the performance was improved. Standard mapping: The agent could bind HL7 and MML address/name information. The efficiency was almost perfect. However, MML formats and HL7 formats were targeted to different use cases, the mapping itself was not semantically incorrect.

研究分野：医療情報学

キーワード：医療標準規格 ターミノロジー

1. 研究開始当初の背景

電子カルテなどの診療情報システムの普及に伴い、電子化された診療情報が各医療機関で蓄積されつつある。これらの診療情報を共有する試みは 2000 年代より積極的に行われてきた。世界的にも診療情報の電子化が進む一方で、そのデータを国家レベルで集積するという統合 EHR(Integrated Electronic Health Record)プロジェクトが検討されつつある。ISO/TR 20514 では、統合 EHR を下記のような目標を実現するためのものと定義している。

- 診療情報の広域における共有
- 診療の質の評価と向上
- 匿名化された診療データを利用した臨床研究の推進
- 公衆衛生向上のための行政データの収集

統合 EHR を実現するためには、医療機関に保存された電子データを統合的に管理する必要がある。そのためにはデータの標準化は必須であり、HL7 や MML といった標準規格への対応も進みつつあるが、現在に至っても十分に普及・活用されているとはいえない状況にあった。

図 1 に示すように通信するメッセージは標準化されていても、各医療機関に保存されている電子データは統一されたスキーマで保存されておらず、標準規格との間で個別に手作業でデータマッピングする必要がある。このようなマッピングルールを生成するためには多大なコストが必要となり、標準化のための障害となっている。医学の進歩や行政上の変化により、診療情報の扱う範囲や定義も変化し続けているため、診療データや標準規格もしばしば改訂を加える必要があり、そのたびに新しいマッピングルールを定義しなければならなかった。

一方で、近年ではオントロジー工学に基づく

意味論的解析を行う自律エージェント技術についての研究が進められている。その成果として、インターネット上の情報から関連する情報を抽出して個人向けの広告へと反映することが商業的にも成功していた。そこで、本研究ではこのような技術を利用して、図 1 に示すように診療情報スキーマ間のマッピングを行うオートマトンを動的に生成する自律エージェントのアルゴリズムについて研究を行うこととした。

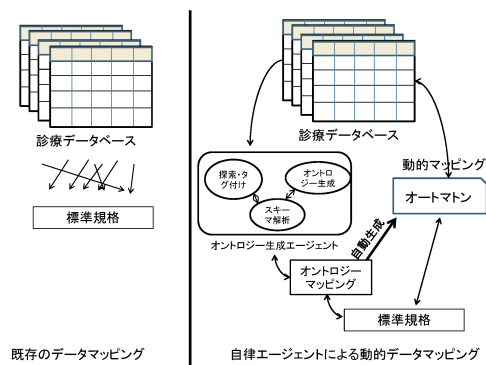


図 1 診療情報標準化とデータマッピング

2. 研究の目的

本研究の目的はデータスキーマの異なる診療情報間で相互に運用するためのマッピングルールを動的に生成する自律エージェントアルゴリズムを研究することであった。この研究を発展させることで、

- 1) 診療情報の意味論的解析技術の確立、
- 2) 診療情報連携のために必要な標準情報形式への動的変換、
- 3) クリニカルインジケータの元となるデータを診療データベースから自動的に収集することができるようになり、医療の質の向上に結びつけること、
- 4) 臨床研究の推進、
- 5) 公衆衛生の向上につなげることができることと考えられた。

3. 研究の方法

本研究の達成目標は前述の通り以下の3項目である。これらについて3年で解決するための研究を行った。

1. 診療データベースを探索エージェント
2. 自己組織化マップによるオントロジー生成
3. スキーマ間のマッピングを行うオートマトンの動的生成

診療データベーススキーマからデータ型や項目名といった情報を語彙として抽出する探索エージェントを開発し、同時に個々のデータに対しての情報を付加したタグでデータをマークアップするエージェントの開発を試みた。対象は愛媛大学医学部附属病院の電子カルテ内に保存された診療情報であり、具体的には参照系データベースに保存されているデータとデータスキーマであった。このエージェントを自律動作させるために、以下のような項目を開発することとした。

1. 語彙の探索(スキーマ解析)
データスキーマからデータ型や項目名を抽出し、語彙として分類する探索エージェントの開発を試みた。
2. データ型の分類
データスキーマから数字、文字列などのデータの種類あるいはデータの単位について情報を収集し、分類を試みた。
3. 意味論的タグ付け
得られた語彙とデータ型などの情報と既存の医学オントロジーやターミノロジーを比較して、内容について検証し、個々のデータに対して意味論的タグ付けを行うこととした。データはXML/OWL形式を用いて、保存して意味論的検討が可能にようにする。効率良くデータを集積し、データスキーマの解析を行うことができるような知識ルールを作成し、適切なタグ付けが行えるアルゴリズムについて検討した。

可能であれば、公開されている診療情報データベースのスキーマを利用して比較検討および検証も進めていくこととした。

4. 研究成果

研究全体を通じて医療分野での標準規格およびターミノロジーについて精査を行い、相互運用性について検討を行った。

現在、国際的協調として実施されている臨床情報モデルの統一的構築の基盤となっているISO 13606規格を軸に多言語対応、複数ターミノロジー対応を行う中間情報モデルを開発した。この中間モデルを軸に自然言語処理の手法を用いて項目名、用語間の近さを判定して大学病院間の検査マスタを突合させることを試みた。JLAC10をターミノロジーとして使用したが、複数項目に同じコードが出現しているなどの問題点があった。負荷試験のために時系列に沿ってデータを格納するための配列などをクレンジングする必要があった。遺伝子検査など各大学独自コードについては、名称からのマッピングを試みた。ウイルス感染症などの血清学的検査については比較的高い適合率が得られたが、HbA1cなどの計法が複数あり、時代により変遷を伴うものは自動で突合することはできず候補として列挙するにとどまった。さらに、複数データモデル間のマッピング試験としてMMLとHL7のマッピングを試みた。住所、患者氏名についてはほぼ確実に突合可能であったが、互いのユースケースが異なるために対応するスキーマが少なく突合作業自体が限定的なものとなった。これらの成果については一部を学会にて報告し、現在論文として公開するための準備を進めているところである。

5. 主な発表論文等
(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 5 件)

1. Kimura E, Kobayashi S, Ishihara K, HTML5 microdata as a semantic container for medical information exchange., Stud Health Technol Inform, 205, 418-422(2014)
2. Kobayashi S, Kimura E, Ishihara K, Archetype model-driven development framework for EHR web system, Healthc Inform Res. 4, 271-277, (2013)
3. Tapuria A, Kalra D, Kobayashi S, Contribution of Clinical Archetypes, and the Challenges, towards Achieving Semantic Inter-operability for EHRs, Healthc. Inform. Res.19, 286-292, (2013).
4. Yoshikawa T, Kimura E, Kobayashi S, Ishihara K, Verification system for postoperative autologous blood retransfusion., Studies in health technology and informatics 192, 977, (2013).
5. Kobayashi S, Tatsukawa A, Ruby implementation of the openEHR specifications, Journal of Advanced Computational Intelligence and Intelligent Informatics, 16, 42-4, (2012).

〔学会発表〕(計 6 件)

1. Kobayashi S, Kume N, Kuroda T, Yoshihara H, Designing clinical models of EHR (Electronic Health Records) for long-term care providers to elderly persons, AMIA2014, Nov 18, 2014, Washington D.C. USA
2. 小林慎治, 糸直人, 吉原博幸, 介護分野での情報規格の開発についての試み, 日本医療情報学連合大会, 2014年11月08日, 幕張メッセ, 千葉
3. Kobayashi S, Bosca D, Kume N, Yoshihara

H, MML (Medical Markup Language) Standard with Archetype Technology, APAMI 2014, Nov 2, 2014, New Delhi, India

4. Kobayashi S, Ueno T, A sustainable business model of Free/Libre Open Source Software for eHealth in Japan, IWEEE2014, May 29 2014, Las Palmas, Spain.
5. Kobayashi S, March A.D., Sandval E, Karlsson D, Nystrom M, Freire S, Berhees B, Corbal I, Anani N, Chen R, The openEHR developers' workshop, MEDINFO2013, Aug 23, 2013, Copenhagen Denmark
6. Kobayashi S, Kimura E, Ishihara K, Creating electronic health records within 15minutes with Ruby on Rails and ISO 13606/openEHR standardized clinical models, MEDINFO2013, Aug 20, 2013, Copenhagen, Denmark

〔図書〕(計 1 件)

1. Kobayashi S, Open Source Software on Medical Domain, Chapter 5, Modern Informationsystems 3, 87-102, (2012).(著作、分担執筆)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ等

The openEHR project Japan

<http://openehr.jp/>

openEHR Ruby implementation

<https://github.com/skoba/openehr-ruby>

MML developers

<https://github.com/skoba/mml>

6. 研究組織

(1) 研究代表者

小林 慎治 (KOBAYASHI SHINJI)
京都大学・大学院情報学研究科 EHR 共同研究講座・特定講師
研究者番号：20423519

(2) 研究分担者

木村 映善 (KIMURA EIZEN)
愛媛大学・医学系研究科(研究院)・准教授
研究者番号：20363244

上野 雄文 (UENO TAKEFMI)
九州大学・医学系研究科(研究院)・研究員
研究者番号：00441668