

医療分野におけるオープンソースソフトウェアについての現状および病院コホートを利用した医療情報と外部データ突合用データマイニングシステムの紹介

星野隆之[†] 小林慎治¹⁾ 植田栄子²⁾ 平山雄一³⁾ 松下由実

IRYO Vol. 71 No. 11 (437-442) 2017

要旨

情報通信技術の発達にともない、医療情報の電子化と集積が急速に進み、電子カルテには日々の診療や保険請求業務を通じて膨大な医療情報が蓄積されている。これらの利用は電子処方箋や画像データの関連医療施設同士での共有、患者による診療履歴の閲覧等の形で還元されはじめている。さらに、これらの基盤を集約し、個々人の生涯にわたる健康記録を個人情報に配慮した形で抽出、二次的に活用する EHR (Electronic Health Record) の構築も進められつつある。EHR に集約された情報はビッグデータとして臨床疫学的に個々の中核医療機関が主体的に探求することができれば、医学的知見を新たに還元することに役立つと考えられる。本稿では、まず、EHR 構築において注目されている医療分野におけるオープンソースソフトウェアの一つである FLOSS (Free/Libre/Open Source Software) の現状と問題点について概説し、次いで、国立国際医療研究センターで開発、稼働しはじめた、病院コホートを利用した医療情報と外部データ突合用データマイニングシステムについて紹介する。

キーワード 電子カルテ, オープンソースソフトウェア, 臨床疫学

はじめに

情報通信技術の発達にともない、医療情報の電子化が急速に進み、電子カルテには日々の診療や保険請求業務を通じて膨大な医療情報が蓄積されつつあ

り、電子化されたデータを国家規模で収集する EHR (Electronic Health Record) の構築も試みられつつある。さらにこの EHR に蓄積された医療ビッグデータを個人情報に配慮した形で抽出し、臨床疫学的な探求を個々の中核医療機関までもが主体的に行う

国立国際医療研究センター 内科系診療部心療内科臨床研究推進部 教育研修室, 1) 京都大学 医学研究科, 2) 青森公立大学 経営経済学部, 3) 株式会社システム情報パートナー †医師
著者連絡先: 星野隆之 五百山クリニック 〒370-0867 群馬県高崎市乗附町1526-2
e-mail: hoshino1133@hotmail.co.jp

(平成29年2月17日受付, 平成29年6月16日受理)

Using FLOSS (Free/Libre/Open Source Software) for Medical Informatics and Data-Mining, Including Hospital Cohort, External Data, and Other Environmental Factors

Takayuki Hoshino, Shinji Kobayashi¹⁾, Teruko Ueda²⁾, Yuichi Hirayama³⁾ and Yumi Matsushita, National Center for Global Health and Medicine, 1) Kyoto University, 2) Aomori Public University, 3) System Information Partner, Inc.

(Received Feb. 17, 2017, Accepted Jun. 16, 2017)

Key Words: electronic health record, FLOSS (free/libre/open source software), clinical epidemiology

ことができれば、医学的な知見をさらに臨床へ還元することに役立つと考えられる。本稿では、まず、EHR構築において注目されているオープンソースソフトウェア（FLOSS: Free/Libre/Open Source Software）について概説し、医療分野におけるFLOSSの応用として日本のORCA（Online Receipt Computer Advantage）Project、アジアにおけるFLOSS、なかでも近年ジャマイカ、ラオス、フィリピンで戦略的に導入されつつあるGNU Healthについて概説する。また、今後のFLOSSとその問題点についても最新の知見を交えて説明する。次いで、医療ビッグデータの国立病院機構などの中核病院での二次利用や相互連携を視野に、国立国際医療研究センターで開発、稼働しはじめた、病院コホートを利用した医療情報と外部データ突合用データマイニングシステムと、ソフトウェアの無償提供について紹介する。

医療分野における FLOSS について

本項では、まず FLOSS とはなにかについて概説し、その医療分野における応用と最近の問題点について本邦および諸外国の動向を含め概説する。

1. FLOSS とは

コンピュータソフトウェアは、コンピュータに処理させる動作をコンピュータに理解されるように設計されたプログラム言語で記述された後に、コンピュータが理解し実行可能な形式に翻訳され、製品として提供される。コンピュータ言語で記述された文書はソースコードと呼ばれ、知的財産として著作権法で保護される。

コンピュータが誕生した1950年代から1960年代はソースコードを含めたソフトウェアはコンピュータの付属物と考えられており、ソースコードは製品とともにユーザーに提供されることが一般的であった。しかし、ソフトウェアの商用利用が進められた1970年代よりソースコードがユーザーに提供されなくなった¹⁾。

コンピュータプログラムの動作や考え方を理解する上でソースコードは非常に有用なものであるため、ソフトウェア技術の発展のためにはソースコードを自由に利用できるようにすべきであると Richard Stallman らが呼びかけ、Free software 運動が1980年代に始まった¹⁾。

Free software 運動の中心となった GNU Project では著作権をもとに、ソースコードの再配布や利用に制限をかけないという GNU General Public License を提案し Copyleft と呼ばれた。さらに、1990年代には商用利用も視野に入れて、Free software 運動の教条的側面をやや緩和した Open Source Software 運動が Eric Raymond らによって提唱された (Eric R. Goodbye, “free software”; hello, “open source”. <http://www.catb.org/~esr/open-source.html>, 1998)。同時期に世界的に広まったインターネットを通じて、ソースコードの流通が容易になると同時にこの Open Source Software の開発は盛んになり、商用利用も進められている。

Free software と Open source software は前者が開発者の「自由」を求め、後者がソースコードの流通と利用を促進するという点で立場が少し異なる。また、“Free”という言葉のために「無償」であることが条件であるかのように誤解されることもある。ソースコードが自由に利用できる結果として「無償」でソフトウェアが利用できるが、ソースコード配布に掛かる実費を要求することやソフトウェアをインストールしたり、追加の開発を行う場合の手数を請求したりすることは禁じられていない。「自由」であることを強調して、Libre software という用語が使われることもあり、これらはまとめて FLOSS (Free/Libre/Open Source Software) と呼ばれている。

2. 医療分野での FLOSS について

FLOSS が成熟するにともない、商用利用も進んでいる。当初はボランティアベースで開発が進められていたが、近年では企業を中心となって開発することも珍しくない。金融などの信頼性が求められる分野でも現在では FLOSS が利用されている。

当然、医療分野でも FLOSS を利用していく試みがこれまで行われてきた。医療分野でのソフトウェア開発には、開発資源が限られることや専門性が高いことなどの問題があり、FLOSS には以下のような期待がある。

- i) 開発コストの低減
- ii) 標準化の推進

世界的に医療費の高騰が経済学上の問題となっており、限られた資源を有効活用するために医療 IT (information technology) に期待が寄せられてきた。FLOSS はこの開発コストを下げるものと期待

されている。

以下に代表的な医療分野での FLOSS を挙げる。

- ・ OpenVistA : アメリカ退役軍人病院で開発された電子カルテシステム
- ・ GNU Health : Tryton (ERP : enterprise resource planning) をベースとして開発された無償の電子カルテシステム
- ・ OpenMRS : 後天性免疫不全症候群や結核の管理のために開発された低開発国向けの電子カルテシステム
- ・ DHIS 2 (district health information software) : 疾病管理レポジトリ

3. 日本における医療分野での FLOSS 活用

日本の医療分野での FLOSS として、ORCA Project による成果物があげられる。これは、日本医師会と日本医師会総合研究機構が会員向けの情報基盤としてレセプトコンピュータ (レセコン) を FLOSS として開発すると2000年に宣言した、日医 IT 化宣言に基づき、レセコン以外にも介護保険用のソフトウェアも開発し、データを収集する情報基盤とすることで、膨大な医療情報を収集・分析し、政策に対して根拠のある反論を行うことを目的としていた。レセコンの普及率は当時でも90%程度あり、多くの医療機関で受け入れやすい状況であった。

現状 (2016年10月14日現在)、日医標準レセプトソフト採用状況を見ると、導入医療機関数は15,868施設であり、シェアは約17%となり、導入は一部医療機関にとどまるが、ORCA Project の前後では、レセコンの初期導入コストに大きな差があり、ORCA 以前200万-300万円程度と高額であったものが、ORCA 以後、50万-150万円で導入可能となり、医療の電子化についての自由化が進み、医療機関に応じてソフトウェアをカスタマイズできるようになったことがある。法整備やガイドライン整備も行われた。その他、日本で開発された医療 FLOSS の主なものには、下記がある。

- ・ OpenDolphin : 地域医療連携システムのクライアントとして開発され、汎用電子カルテシステムとして独立している
- ・ NOA (Neat Online Assistant) : 大橋産婦人科で作成された PHP (Hypertext Preprocessor) ベースの電子カルテ
- ・ tkDerm : 犬塚医師による皮膚科専門画像データベース

4. アジアにおける医療分野の FLOSS

アジアは、約40億人 (中国13億人、インド13億人、ASEAN (Association of South-East Asian Nations) 6億人、日本1.2億人、韓国・北朝鮮7,500万人) の巨大な人口を抱えた、新興経済圏である。開発途上国も多く伝染性疾患を問題として抱える一方で、急激な経済成長・発展にともない、肥満、糖尿病、生活習慣病などの疾病増加による疾病構造の変化もみられている。さらに、出生率の急速な低下にともなう少子高齢化社会の進行が特徴的でもある。このような状況下、医療情報分野には、経済的、宗教的、政治的な格差をこえた、Universal healthcare support が期待され、多くの国では経済状況がいまだに脆弱であるため、開発および導入コストの削減、実装に基づいた標準化が求められている。この観点から、アジアにおいても、医療 FLOSS は積極的に導入が試みられている ASEAN においては、AeHIN (ASEAN eHealth Information Network) は積極的に FLOSS を使っていくと宣言しており、事例として、下記がある。

・ DHIS 2 : 国境なき医師団で統一的に利用されている患者登録レポジトリであり、患者個人の疾患登録をベースに地域での疾病状況などの統計データ地図データを図示することができる。

・ OpenHIE (Health Information Exchange) : IHE (ihe.net) 標準の FLOSS 実装で、病院間でのデータ連携を容易にするソフトウェアである。

そのほか、近年ジャマイカでの国家戦略としての導入のほか、ラオス、フィリピンの病院施設等350カ所で採用されている GNU Health は、電子カルテのみならず、医療および福祉の情報を包括的に扱い、疫学的な解析も視野に入れたシステムとなっている。GNU Health は、Luis Falcon により2008年より開発開始され、Tryton という ERP (Enterprise Resource Planning) パッケージをベースとしている。病院単体あるいは複数の病院を連携して利用可能であり、個人および集団の健康管理、患者属性情報、家族単位、地域ごとなど、患者管理として、社会的経済的状況、生活歴、外来診察、入院、検査だけでなく、保健所業務として、予算、資産、検査施設、病院などの管理が可能であり、情報管理では、報告書作成、人口動態や疫学統計が可能である。GNU Health 日本語翻訳は、ボランティアで運営されており、一度は100%翻訳された時期があるが、このたび GNU Health 3.0 にアップデートされている。今後、多く

の施設で利用される場合は、翻訳の監修などの作業が必要になることが考えられる。GNU Health 3.0では、さらに、スマートフォンやWebからもアクセス可能になるなど、新規機能が追加された。

5. 医療分野での FLOSS 活用の問題点

このように、FLOSS は医療分野においても世界中で積極的に利用され、データの標準化などにも役に立っている。一方で、いくつか問題点も指摘されている。

1つは医療データを二次利用するための国際的コンセンサスがまだ不十分であることが挙げられる。現在、本邦においても改正個人情報保護法を基にした匿名化データの二次利用について国会で議論が進められている段階である臨床研究においては、国際的な取り組みが必要となることも多いため、EU (European Union) 圏域やアメリカなどの動向も含めてデータの相互利用も視野に入れた法制度整備が必要とされている。

2つ目は医療分野でのソフトウェアに対して、安全性の確保のため平成27年に薬事法が改正されたことが挙げられる。適用される範囲は限られており、国際的に同様の法整備が進まれている中、本邦での法制度の趣旨も正当ではあるが、医療用画像管理ソフトウェアがその対象となるとされており、その分野の FLOSS をどう評価し認可を受けるかが問題となっている。医療分野の FLOSS 開発はボランティアベースのコミュニティが主体であることが多く、認証を受けるためのコストをどう負担していくかが喫緊の課題となっている。

病院情報システムからの臨床データと外部データの突合解析が可能なデータマイニング用「共通システム」開発とこれを用いた疾病構造に関わる「環境モニタリングシステム」の紹介

前項で紹介されたように、世界的に医療情報の二次利用を目指し、医療情報をデータマイニングするソフトウェアが求められ、FLOSS の形で提供されつつある。一方で、医療 FLOSS が抱える諸問題、とくに医薬品医療機器等法の適用や運用・改定コスト、ソフトウェアの医療分野での利用や責任も大きな問題である。

このたび、われわれ国立国際医療研究センターでは、環境省地球環境保全等試験研究事業の、「環境

因子の変化にともなう疾病構造変化モニタリングと中長期環境モニタリングおよび暴露調査結果を用いた環境がヒトへ与える影響の解析を行う病院コホートを利用したデータマイニングシステムの研究」により、病院情報システムからの臨床データと外部データの突合解析が可能なデータマイニング用「共通システム」を開発した。このシステムを用いた、電子カルテビッグデータと環境ビッグデータ突合解析による、疾病構造に関わる「環境モニタリングシステム」の具体例を図1に示す。

われわれが開発した「共通システム」では、病院電子カルテやレセプトデータなどの病院情報ビッグデータから情報抽出する際、抽出と同時に対応表を作成しない匿名化を行い、突合用臨床データベースに再構築するため、個人情報非安全エリアに持ち出さずに、疫学研究が可能となる。国立国際医療研究センターでは、病院電子カルテシステム「Mega-Oak™ NEC Corporation, Tokyo, Japan」から臨床データを抽出している。Mega-Oak などの大病院向け電子カルテシステムの多くは電子カルテメーカーが独自に開発し、商品化したものである。このため、データベース化して外部データ、本研究の具体例では外部の公開環境データとの突合を行うためには、データ形式の標準化処理が必要となる。今回われわれは医療情報の「標準化ストレージ」という概念に着目し、すべての医療機関を対象とした医療情報の交換・共有による医療の質の向上を目的とした「厚生労働省電子的診療情報交換推進事業」に準拠し、標準化の方法として、SS-MIX (SS-MIX: Standardized Structured Medical Information eXchange software version 2, Consortium for SS-MIX Dissemination and Promotion, Kanagawa, Japan) を用いてデータベースの作成を行った。本臨床データベースから解析用に条件づけて抽出可能な臨床データ項目は、全ケースでの年齢、性別、郡市までの居所、保険種別、初受診日、初診日、再診日、受診診療科、病名、発症日、発症日を含む前7日移動平均データ、病名終了日、転帰(治癒、治療継続、中断、死亡)、一部入力があったケースの診療計測データ(収縮期血圧、拡張期血圧、脈拍、体重など)、検査項目、検査結果、検査日時、投薬内容、投薬日時、処置内容、処置日時などである。たとえば、このシステムで抽出した臨床データと、環境データの公開されたデータソースとして、環境省大気汚染物質広域監視システム「そらまめ君」(<http://soramame>。

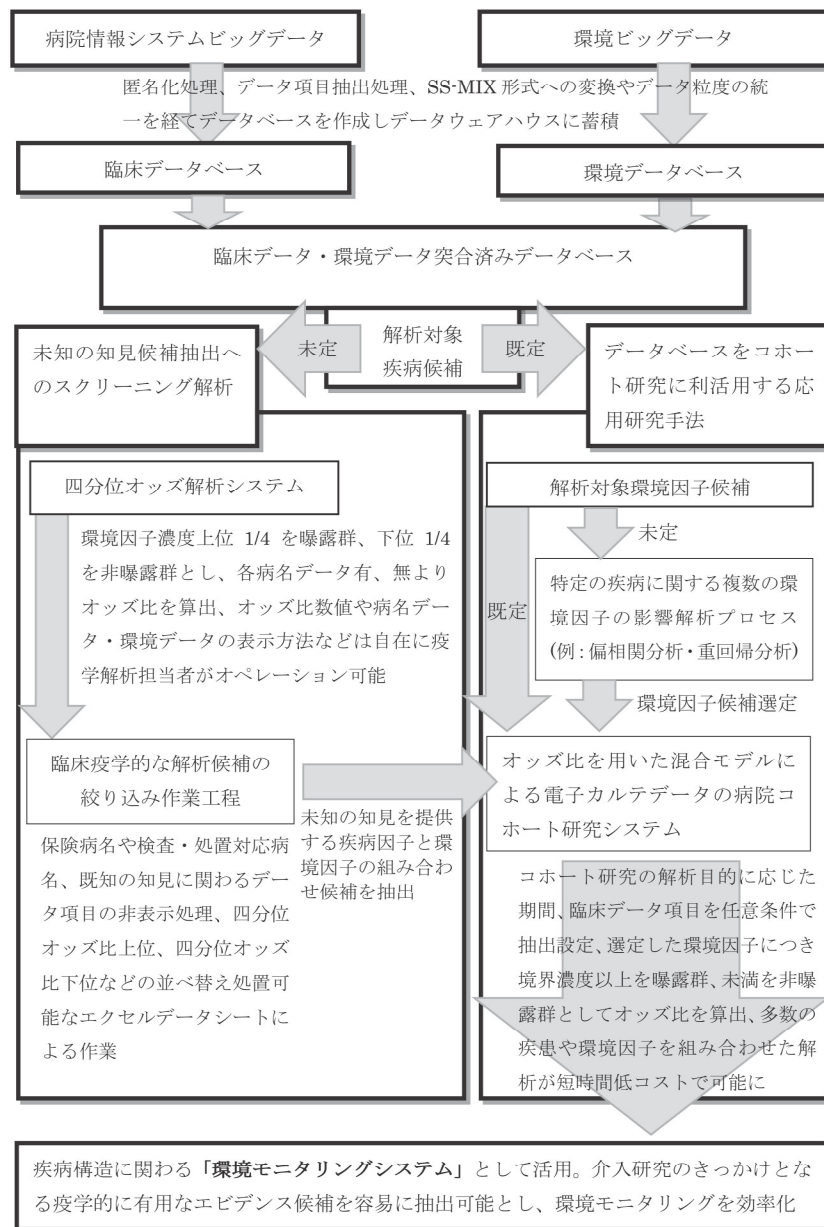


図 1 病院情報システムと環境ビッグデータ突合解析のための「共通システム」の概要

taiki.go.jp/環境省), 環境省花粉観測システム「はなこさん」(<http://kafun.taiki.go.jp/環境省>) および気象庁気象観測データ (<http://www.jma.go.jp/jma/menu/menureport.html> 気象庁) を利用し、臨床データと環境データの突合解析を行うことが可能であり、われわれ研究班でも、実際に本システムを稼働させて、既存研究で着目されつつある、大気中の PM2.5 濃度と虚血発作の発現との関連性につき、電子カルテから抽出・構築した臨床データベースと、環境データベースを突合解析し、得られた知見を発表している²⁾。

われわれが開発したシステムは、ソフトウェアとしての特許取得は目指さず、希望される他施設への

提供および、今後の疾病構造と環境因子などの外部データとの突合解析研究の基盤となることを目指している。医療分野のソフトウェアの宿命として、各導入施設において実装後、さらに改善点をあげて改良・カスタマイズが不可欠であり、いわば使用しながら次々に改善していくソフトウェア開発手法である。スパイラルメソッドの一過程というべきものである。本ソフトウェアをこの「医療」にご紹介する意義として、類似の病院情報システムを稼働させ、せっかく、病院コホート研究が可能な臨床データをかかえた国立病院機構の医療情報ご担当に情報提供する目的がある。臨床ビッグデータを用いた解析を行う際、症例数も多く集約され、地域に根差してお

り、高度な医療と医療情報システム、その担当者を擁したわれわれ国立病院機構の医療機関は、ちょうどよいセンターと成り得る。その際、臨床疫学的な解析には、今後本ソフトウェアのような、電子カルテシステムと、外部データをつなぐ、突合解析ソフトウェアが不可欠と考える。

おわりに

本報告では、まず、医療分野におけるオープンソースソフトウェアの現状と問題点について概説した。次いで、国立国際医療研究センター病院にて開発、稼働しはじめた、病院コホートを利用した医療情報と外部データ突合用データマイニングシステムについて紹介した。本システムは、環境省地球環境保全等試験研究事業の、「環境因子の変化にともなう疾病構造変化モニタリングと中長期環境モニタリングおよび暴露調査結果を用いた環境がヒトへ与える影響の解析を行う病院コホートを利用したデータマイニングシステムの研究」により研究開発されたものであり、その実施要項に基づき、作成した共通システムを公開、頒布用情報提供することで、将来、協力施設とも連携してデータ解析を行い得るシステム環境が整い、温暖化等の地球規模の環境因子の変動

が健康におよぼす影響をはじめ、外部因子が疾病構造に与える影響を臨床疫学の視点で明らかにするための、一手段となることを目指している。このため、われわれが開発し得たソフトウェアの範囲内において、日本語版マニュアル一式とともに、ソフトウェアの無償での提供を行っている。実装には個々の施設での医療情報担当者によるカスタマイズや施設個々の電子カルテとの接続作業が必要で、この際にはコストが発生する可能性が考えられる。提供希望施設のお問い合わせをお待ちする。

著者の利益相反：本論文発表内容に関連して申告なし。

【文献】

- 1) Williams S. Free As in Freedom: Richard Stallman's Crusade for Free Software. O'Reilly & Associates Inc. , Sebastopol, Calif, 2002.
- 2) Hoshino T, Hoshino A, Nishino J. Assessment of associations between ischaemic attacks in patients with type2 diabetes mellitus and air concentrations of particulate matter <2.5 μm . J Int Med Res 2016 ; 44 : 639-55.