

第 38 回 M テクノロジー学会大会 講演論文集



メインテーマ 真のPOMRの実現に向けて — 意味理解とその応用 —

- 主催 日本 M テクノロジー学会
- 大会長 山本皓二（鈴鹿医療科学大学）

- 開催日時 2011 年 8 月 5 日（金）～ 6 日（土）
- 会場 日間賀島公民館（愛知県知多郡南知多町日間賀島）

巻頭言

第 38 回 M テクノロジー学会大会

メインテーマ

真の POMR の実現に向けて
— 意味理解とその応用 —

鈴鹿医療科学大学
山本皓二



1969 年、L.L. Weed によって提唱された POS による診療録(POMR)の記載方法がその後の医学教育や診療録の書き方に大きな変革をもたらし、診療録はより科学的な臨床の経過記録へと変貌しました。この時、L.L. Weed が本当にやりたかったことは、診療録を臨床の科学的な実験記録と位置付け、その中から臨床にとって有益な知見を見つけだして医学・医療の発展を図ろうということであったかと思えます。その後開発された多くのいわゆる「電子カルテ」と称される電子版診療録も POS を理想形と考え、入力画面も固定的な SOAP 形式を実装してきました。

無論、医療は起承転結が直線で結びつけられたような 2 次元平面で記述できる単純世界ではなく、固定的な SOAP 形式が L.L. Weed が求めたであろう「真の POMR」の入力系に適していない可能性が大きいのですが、結果として診療で発生する多くの情報が電子化されるようになり、医学・医療に役立つ知見を見つけるための素データが揃いつつあります。そして今、入力の容易性や診療への適応性などの議論を一旦停止し、蓄積されたデータから医学・医療に役立つ知見を見出す研究、即ち、意味理解とその応用の研究に軸足を移していく必要があるのではないのでしょうか。その先に、その研究結果を受けて改めて入力系を考える時代が来て、真の POMR が実現していくのだらうと思えます。診療データの多くは、文章などのテキストデータです。そこで、第 38 回の本学会では、テキストデータからの意味理解を一つの柱に据えて議論をしてみたいと考えています。

本学会が持つもう一つの重要な役割は、研究者の輪を広げることです。個々の先生が日頃考えていることや研究してきた内容を持ち寄り、それを皆で議論してより良い論文に仕上げている。この学会がそのような験算の場になれば、それだけで成功だと考えています。この場合、テーマに制限はありません。また、内容も未完成で構いません。ここに来て発表をし、議論を通じて何かを持って帰ってほしいと願っています。自由に討論をし、議論を深めていく場として、風光明媚な南知多半島の離島を会場に選び、民宿で合宿をしながら、皆でワイワイと議論をしたいと考えています。

多くの研究者の参加を期待しています。

**第38回Mテクノロジー学会大会 in 日間賀島
プログラム**

8月5日(金)	
14:00~14:10	開会の挨拶 第38回大会長 山本皓二(鈴鹿医療科学大学)
14:10~15:30	【一般講演1】 座長: 山本皓二(鈴鹿医療科学大学) 1. 四大学病院間の臨床研究を支援する分散データベース構築 糸 直人 (京都大学病院 医療情報部) 2. 閲覧履歴を利用した診療文書のランキング手法の検討 岡本 和也 (京都大学病院 医療情報部) 3. 地域医療連携システムにおけるインフラとしてのwebサービス導入の試み 竹村 匡正 (京都大学病院 医療情報部)
休 憩	
15:40~17:00	【一般講演2】 座長: 嶋芳成(日本ダイナシステム株式会社) 4. フリーデータベース「Globals」 澤田 潔 (MTAオープンソースソフトウェア研究会) 5. 情報システム開発から医学研究利用へ 大楯 陽一 (東海大学 医学部・基礎医学系) 6. 獲得した知識の有効的格納について 紀ノ定 保臣 (岐阜大学大学院 医用情報学分野)
18:30~	【懇親会】 山本皓二(鈴鹿医療科学大学), 川中 普晴(三重大学)

8月6日(土)	
9:30~10:50	【一般講演3】 座長: 鈴木隆弘(千葉大学) 7. 院内紙文書を利活用するための文書画像処理～紙文書の電子化, そして検索への試み～ 川中普晴 (三重大学大学院工学研究科) 8. 診療文書作成プログラムの試作(Cache, JQueryを使用して) 土屋 喬義 (土屋小児病院) 9. ComeJisyoの退院時サマリへの適用 鈴木隆弘 (千葉大学)
休 憩	
11:00~11:50	【一般講演4】 座長: 河村徹郎(鈴鹿医療科学大学) 10. 同じ位置で撮影された過去と現在の歯科口内法X線画像の同定法 檜作彰良 (三重大学大学院工学研究科) 11. オープンソフトウェア GT.M 2011 澤田 潔 (MTAオープンソースソフトウェア研究会)
11:50~12:00	閉会の挨拶 第38回大会長 山本皓二(鈴鹿医療科学大学)
休 憩	
13:00~14:00	日本Mテクノロジー学会 総会

四大学病院間の臨床研究を支援する分散データベース構築

糸 直人¹ 真鍋 史朗² 松村 泰志² 荒木 賢二³ 高林 克日己⁴ 嶋 芳成⁵
岡本 和也¹ 竹村 匡正¹ 黒田 知宏¹ 吉原 博幸¹
¹京都大学医学部附属病院 医療情報部 ²大阪大学 医学部附属病院 医療情報部
³宮崎大学 医学部附属病院 医療情報部 ⁴千葉大学 医学部附属病院 医療情報部
⁵日本ダイナシシステム株式会社

【はじめに】

Evidence Based Medicine (EBM) で求められている“根拠に基づく医療”を実現するための一つの方法として、臨床研究において統計的なデータを収集・蓄積し、新たな知見をえる方法がある。また本質的に臨床研究は症例をより多くあつめて実施することに意味があり、大規模なデータセットを用意できることが望ましい。とくに疫学研究の場合には、ある病院、診療機関の枠を越えて地域を横断した統計的なデータが必須である。

【目的】

本研究は、大学病院をまたいだ診療情報の横断的な検索を可能とする大規模分散データベース (DB) の構築により、各施設の臨床研究を促進することを目的とする。特に一つの施設だけでは症例数がすくなく臨床研究が成立しないような難病に関しても、多施設間で症例数を確保することにより臨床研究を促進できることが期待される。

【四大学間の分散臨床研究 DB】

本研究では、医事情報と臨床情報を組み合わせた臨床研究 DB を構築し、施設をまたいだ横断検索を実現する。医事情報としては、患者基本情報、診断群分類、主病名、処方などの情報を利用できる。また、臨床情報としては、検査データ（血液検査、尿検査）を利用する。

施設を横断して情報を利用する場合、各施設によってセキュリティポリシーが異なるため、全データを一元的に集約するような

DB モデルはそぐわない。そこで、本研究ではデータセンター (DC) から各施設 DB へのインデックスだけを保持し、データは各施設に分散して保存することとした。DC から整形された検索クエリーだけが各施設に配布される。各施設では検索の実行結果を確認して公開しても問題無いことを確認してから、DC への検索結果の共有を許可できる仕組みを構築する。その際、分散された DB を仮想的に一つのローカル DB としてアクセスできる *Cache* (InterSystems 社) を採用する。また検索速度を向上させるために *Bitmap* 検索を実装する。臨床研究でよく利用される検索クエリーとして相対時間検索を実装し、あるイベントの前後数日を範囲とした検索を可能とする。具体的には、『ある薬 A を投与した前後 7 日間に検査項目 B の結果が C であった患者』を検索できる。

【結果と考察】

分散 DB を実装し、パイロットスタディとして 5 つの臨床研究を実施した。結果として、臨床情報や患者基本情報の不足のため、検索結果をそのまま臨床研究にもちいることは困難であるが、症例の *Case finding* には有効であることが確認できた。施設間のデータ突合を容易にするためには、検査マスターの標準化などが課題になることがわかった。

【まとめ】

臨床研究を目的とした四大学病院間の分散臨床研究データベースを構築した。

閲覧履歴を利用した診療文書のランキング手法の検討

岡本 和也^{1,2} 鶴岡 誠子² 竹村 匡正^{1,2} 糸 直人^{1,2} 黒田 知宏^{1,2} 吉原 博幸^{1,2}

¹京都大学医学部附属病院 医療情報部 ²京都大学 大学院情報学研究科

【はじめに】

現在、多くの病院において病院情報システムが稼動している。そして、病院情報システムにおいて、処方・処方情報、検査情報、病名情報、経過情報、退院時サマリなどの各種診療情報の蓄積が進んでいる。診療情報の電子的な蓄積の利点の1つに医療従事者間でのスムーズな情報の共有が挙げられる。しかし、診療情報の共有はそれほど活発とはいえず、それゆえに診療情報の質の向上が期待できない現状があると推測される。そこで、本研究では診療情報の共有を推進するため、良質な診療文書にアクセスするための環境の構築を検討する。

【目的】

良質な診療文書にアクセスするための環境の構築を目標とする。特に、診療文書のランキングに注目し、ランキング手法の検討を行う。

【診療文書と閲覧履歴】

診療文書をランキングする方法として、診療文書の内容を解析してランキングを行う方法と診療文書に付随する他の情報を利用してランキングを行う方法がある。文書の内容を解析することは一般的に意味解析などを用いるが、成功事例はほとんどなく困難と思われる。そこで、本研究では診療文書に付随する他の情報を利用したランキング手法を検討する。診療文書の質と最も関連が深い要素の1つとして医療従事者が挙げられる。なぜなら、「良い医療従事者の書いた診療文書は良い診療文書である」、「多数の医療従事者に閲覧されている診療文書は良い診療文書である」といった仮説が立てられるからである。これらの仮説はリンク情報を利用したウェブページのラン

キングアルゴリズムが利用している仮説と類似している。そこで、本研究では、ウェブページのランキングアルゴリズムとして広く知られている HITS アルゴリズムを適応することを検討する。

【HITS アルゴリズムの適応】

HITS アルゴリズムではリンクを張っている節点（ハブ）とリンクが張られている節点（権威）を考える。そして、HITS アルゴリズムは高いハブ得点を持つハブが高い権威得点を持つ権威にリンクを張っており、高い権威得点を持つ権威が高いハブ得点を持つハブからリンクを張られているといった関係を満たすように、各ハブにハブ得点を与え、各権威に権威得点を与える。そのため、医療従事者をハブ、診療文書を権威と考え、リンクを閲覧履歴に応じて張ることで HITS アルゴリズムを適応することとする。

【考察】

閲覧履歴を利用したランキング手法を適応する際に、重要となると考えられる要素として、閲覧履歴がある。つまり、現状で、ランキングを行うのに十分な閲覧履歴が得られるかということであり、今後分析を行っていく必要がある。また、ウェブ検索において、HITS アルゴリズムが上手く機能しなくなる原因としてリンクスパムが挙げられるが、実際に稼動した後に同様な問題が発生する可能性があるものとして留意する必要がある。その他、実稼動を考える際には計算時間やデータ量といったものを考慮する必要がある。

【まとめ】

閲覧履歴を利用した診療文書のランキング手法の検討を行った。

地域医療連携システムにおけるインフラとしてのウェブサービス構築の試み

徳永 達也¹⁾、竹村 匡正¹⁾²⁾、桑 直人¹⁾²⁾、岡本 和也¹⁾²⁾、黒田 知宏¹⁾²⁾

吉原 博幸¹⁾²⁾

¹⁾京都大学大学院情報学研究科、²⁾京都大学医学部附属病院 医療情報部

【はじめに】

多くの地域医療連携システム（EHR）が日本国内で稼働しつつあり、国の政策においても、IT 基本戦略におけるターゲットの一つに医療が挙げられ、その中でも EHR の促進が述べられている。その利点としては重複検査の削減等によるコスト削減等が取り沙汰されるが、EHR の本来の目的は医療をイノベーションすることであり、そのためのインフラとなることが重要である。今回、我々は、EHR のための病院間の情報共有の基盤である iDolphin に対して、Service Oriented Architecture を踏まえたウェブサービスを導入し、EHR 上のデータを簡易に、かつ多角的に利用できる仕組みを構築したので報告する。

【iDolphin】

iDolphin とは、Dolphin Project において開発された地域内で情報を共有するための基盤である。地域医療連携には、様々な機能が必要であるが、閲覧者の患者情報の表示、アクセス制御、データアップロードの仕組み等がパッケージ化されている。しかし、これまでの EHR に必要とされた機能を中心に実装されていることから、多角的なデータの利用を想定されたものではなかった。

【iDolphin のウェブサービス化】

iDolphin の仕組みを変更すること無く、かつ認証等の仕組みを利用しつつ、外部からデータの多角的な利用を行うことができれば、運用ノウハウ等を含めた資産を有効に活用することができる。そのために、iDolphin の各モジュールをサービスから利用するサービスを構築した。

iDolphin は、EHR におけるセッション管理、ユーザー管理、ドキュメントへのアクセス管理、またドキュメント本体の取得をモジュール単位で別々に行っている。特に診療文章の取得については、xml によって汎用化された問い合わせ方法（MMLQuery）によって cache' DB にアクセスし、同じく汎用化された結果（MMLResult）を得ることができる。この仕組みを利用して、iDolphin にアクセスするため一連の手続きをウェブサービス化した。

【結果・考察】

ウェブサーバー上にサービスを構築した。REST 形式による問い合わせに対して、患者情報を xml 形式のデータをセキュアに、かつ簡便に返すことが可能になった。例えば、これまで困難であった「任意の患者群の処方一覧」等の獲得が可能になった。また、これによって、EHR 上のデータの新たな利用が可能になると思われる。

フリーデータベース「Globals」

澤田 潔¹ CAMTA-JP¹

はじめに

Globals は、インターシステムズ社(以下 IS 社) Cache' の多次元配列・ストレージエンジンの部分を「フリー」のライセンスで利用できるコミュニティベースのデータベースである。DB エンジン部の利用においては「フリー」だが、オープンソースではない。また Globals に関しては IS 社のサポートは一切ない。Globals の最初の Version(2011/03/24)は、"Java 開発者のための Globals" で、Java API が公開されている。

Globals と Cache' の相違点

Globals は、多次元配列のストレージと JavaAPI だけを提供する。GlobalsDB のグローバル変数へアクセスするための M 言語は利用できない。GlobalsDB は純粋なデータベースエンジンである。Cache' は豊富な機能のセットを提供しているが、一方、Globals が提供するものは多次元配列ストレージがすべてである。Globals には、SQL、Cache' オブジェクトは含まれてない。また、ECP, Web サービス(CSP), セキュリティ, Cache' スクリプト, Studio, Cache' アプリケーション開発とデプロイするための環境など Cache' を構成する機能はない。

Globals と NoSQL

クラウドデータベースの要件のひとつとして、分散化処理がある。商用のライセン

スの Cache' では、かるく数 1000 を超えるユーザ数を必要とする負荷分散のための NoSQL-DB はコスト的に実現が難しい。Globals 本体はオブジェクトは利用できないが、Java 環境とフリーの DB エンジンとしてのメリットを生かし NoSQL のムーブメントを生み出そうとする狙いがある。

Globals のライセンス

IS 社は、Globals を "そのままの状態 (as is)" で提供する。ソフトウェアを再配布する場合は、出典を明記(attribution)しなければならないが、ソフトウェアの使用においてはエンドースは不要である。いかなる場合においても、IS 社はソフトウェアの使用において生じる損害賠償をする責任を負わない。また IS 社は、技術的または製品サポート、アップデートを提供しない。変更の有無にかかわらずバイナリ形式でのソフトウェアで配布し使用しなければならない。

Globals のコミュニティ

IS 社のサポートはないが、Globals には有志によるコミュニティがある。コミュニティには誰もが参加でき、Globals について学習ができ、メンバー間で情報やソースコードなどが共有できる。

参考文献(URL)

1. GlobalsDB.org <http://globalsdb.org/>
2. Globals 日本語サイト <http://sites.google.com/site/globals>

第33回日本Mテクノロジー学会大会抄録

演題名：情報システム開発から医学研究利用へ

発表者：大榎陽一

所属：東海大学医学部 基礎医学系 医学教育・情報学

1. はじめに

日本では1970年頃から医療分野への情報システムの利用が始まった。多くの医療情報が開発され、その中に貴重な医療データが蓄積されてきたが、思うようには活用されていない。「業務システムを開発して動かせば、自然に医療情報が蓄積されて、いずれ医学的利用が可能になる」という訳にはいかなかった。私が現在までに開発してきた医療保健福祉情報システムは次の通りである。示されている年次はシステム稼働年である。

1) 呼吸機能検査システム⁽¹⁾ (RIMCOS-11) - 1971年

無水式の呼吸機能検査回路を開発した。呼吸回路弁の切替えを自動化し、測定から報告書作成までをコンピュータ化した。特に、呼吸機能検査精度を左右する呼吸努力チェックも自動化した。測定機能は、Forced Vital Capacity, Closing Volume, N₂ Washout, DL_{co}, 呼気・血液ガス計算である。

2) 呼吸器疾患集中治療室24時間モニタリング・システム⁽²⁾ (IRCU) - 1973年

ベッドサイド・モニターはヒューレット・パッカートの端末を使用するが、その出力を1分単位でサーバーコンピュータに取り込んで記憶する。ナースステーションで、患者基本情報、血液ガス測定結果などをマークシートで入力する。これらのデータを統合して、定時または随時、報告書を印刷する。

3) 大阪府立羽曳野病院情報システム⁽³⁾ (オーダーリングシステム) - 1975年

日本では、東京駒込病院について、高知医科大学病院と同時期にオーダーリングシステムを開発した。これらの病院が大型コンピュータを使用したのに対して、汎用ミニコンピュータ(DEC社 PDP-11)を使用して、経済的なシステムであった。開発言語はMUMPS(現在のM言語)で、24時間365日ノンストップで継続開発と保守が可能である。外来診察室と受付、病棟ナースステーション、医事会計課、臨床検査室と採血室、薬局、給食係、用度課、経理課、総看護師長室などの業務を開発した。

4) 東海大学付属病院情報システム⁽⁴⁾ (オーダーリングシステム) - 1990年

大学病院としては初めてアプリケーションサーバーとしてUNIXマシーンを採用した。この開発言語とデータベースはU-MUMPSであった。この部分は住友電工が担当し、医事会計はNECが担当した。ジョイントプロジェクトでは各社の特徴を生かすことができたが、協力関係を維持するのは困難であった。

5) 伊勢原市光カードシステム&ISDNネットワーク⁽⁵⁾ (すこやかカード) - 1990年

保健医療の情報開示と地域でのサービス連携を目的として、光カードに検査結果、処方内容、医療画像を書き込んで、住民個人が管理・携帯するシステムである。当時はカルテを患者に開示する習慣はなく、薬局では薬剤名を消して患者に渡していたので、医療情報開示という考え方は画期的であった。世界各国や日本全国から見学者が訪れた。

6) 衛星通信による災害支援と医師生涯教育支援システム⁽⁶⁾ 1997年

電電公社時代に打ち上げたN-STARをNTT化後で公益的に利用する方法の一つとして、災害時の医療通信システムを開発した。埼玉県と神奈川県で防災の日にシミュレーションを行った。災害時だけ使うことは効率が悪いことと、普段の利用により操作の習得につながることから、医師の生涯教育用システムも開発した。教材と講師の動画は衛星通信で配信され、受講者の質問は地上回線で講師に伝えるという二重通信方式を用いた。3年間の実証実験で幕を閉じた。

7) 全国開業医向け電子カルテ+光カード⁽⁷⁾ (光メディカル) - 1998年

熊本県、東京都、神奈川県、埼玉県、福島県の医師会有志が核になり、大学や産業界が協力して、開業医向けのパソコン電子カルテを開発した。患者情報の共有ツールとして、光カードを採用した。この理由は、当時としてはインターネットのインフラおよびセキュリティが十分ではないこと、医師主導の医療から患者主体の医療とすること、光カードの容量が大きいことなどであった。約100台近くまで広がったが、医師会組織としての普及には至らなかった。

8) G8国際遠隔医療ネットワーク⁽⁸⁾ (日本、カナダ、ドイツなど) - 1998年

サミット参加国の経済活性化プロジェクトの一環として、医療分野では国際テレビ会議システムの開発を行った。当時は、国際間インターネットの通信が安定しておらず、その速度も十分ではなかったため、ISDNを6本束ねる技術を用いてテレビ会議を行った。最新の手術を中継したり、学生や研修医の討論を行ったり、症例検討会を行った。国際テレビ会議のマニュアル化を成果して、この事業を終了した。

9) 医師会ネットワーク⁽⁹⁾ (足柄上、茅ヶ崎、三浦、福島県郡山の医師会) - 2000年

光メディカードが患者主体のネットワークであるのに対して、医師間のネットワークも必要ということで開発した。インターネットを用いて、非再現暗号システムを独自に考案し、個人認証、暗号化通信、リアルタイム暗号化データベースに用いた。医師会で共通した機能は、スケジュール管理、医師会資料共有、掲示板、生涯教育、患者情報の共有などである。開発にはCacheのWebLinkを用いた。このツールではデータベースをアクセスしている間のみライセンスがカウントされるため、医師は接続したままで診療することができる。現在も使われている。

10) 在宅医療福祉ネットワーク⁽¹⁰⁻¹¹⁾ (西川町-1998年、朝日町-2007年)

西川町と朝日町は共に雪深い町で、積雪期間中での在宅医療福祉の支援として、テレビ電話を用いたシステムを開発した。1998年にはアナログ方式のテレビ電話を用いた。この方式では、特別な通信インフラを必要としない。このテレビ電話は自動受信機能や外部カメラ接続とコントロールの機能を有していた。医療保険制度上や交通の関係で医師が訪問する回数は限られているが、このシステムを使うと自由に在宅患者の様子を観察できる。朝日町では、2007年には光通信インフラが整備されたので、町役場にサーバーを設置して、医療機関や各家庭にはパソコンを設置して、光ケーブルで接続した。CacheのWebLinkを用いて、患者情報の共有を行った。テレビ電話機能は、Skypeを用いた。

11) 市町村保健情報システム⁽¹²⁻¹⁴⁾ (西川町-1996年、伊勢原市-2001年、郡山市-2001年、朝日町-2003年)

市町村では保健事業として、成人健診（特定健診、後期高齢者健診、国保人間ドック）、がん検診、予防接種、母子保健（健診、訪問、相談）、畜犬管理などを行っている。システム導入時には補助金が出るが多いが、システム更新時には独自財源となる。この時点で開発を依頼されることが多い。西川町では住友電工から、伊勢原市は NTT データから、郡山市は地元企業から、それぞれ引き継いだ。朝日町は新規導入であった。郡山市と朝日町は、特定健診を機会に企業パッケージに移行したが、西川町、伊勢原市では現在も現役のシステムである。

12) 保健所情報システム⁽¹⁵⁾（上越保健所）－1998年

上越保健所管内 18 市町村の老人基本健診における保健指導システムを開発した。また、老人基本健診と事業所健診を統合したデータベースを作成した。厚労科学研究費により、システムの充実とデータ解析を行った。

13) 事業所健診システム⁽¹⁶⁾（郡山市健康振興財団）－1991年

事業所健診業務システムを開発した。その後、健康増進業務、老人基本健診／特定健診業務システムを開発した。これらの結果をデータベース化し、保健指導、統計処理、電子健康手帳に活用するシステムも開発した。現在もフル稼働中である。

14) 学会データベース⁽¹⁷⁻¹⁸⁾（日本総合健診医学会－2002年、日本脳卒中学会－2003、2010年）

総合健診医学会に所属する全国 45 施設から 2002 年度の健診結果約 70 万件を集めて、検査法とデータのチェックを行い、全国で検査件数が大きく偏らないように抽出したデータベースを作成した。このデータを男女別・5 才ごとに区分し、非線形最適化法により正規分布を抽出し、24 項目の基準範囲を求めた。また、脳卒中学会と脳卒中協会により、脳卒中患者の急性期の治療カルテシステムを開発し、毎年症例を増やして、2010 年現在で約 46,000 症例に達して、世界最大規模の脳卒中治療データベースとなっている。現在までに、大規模解析を二回行って、その成果を出版した。

2. 医学研究に役立ったデータ

・神奈川県伊勢原市健康管理課－24年間

健診、がん検診、健康教育・訪問・相談、母子保健、予防接種、畜犬管理

住民基本台帳による生死データ、厚労省に個別死亡票の目的外利用申請で得た死亡原因

・福島県郡山市地域保健課－10年間

健診と問診、がん検診、住民基本台帳による生死データ

・福島県郡山市健康振興財団－22年間

事業所健診と問診、がん検診、健康増進、健康教育・訪問・相談、フォローアップデータ

・日本総合健診医学会－約80万件

全国45施設の健診結果

・上越保健所管内－約15万件

市町村健診、事業所健診、人間ドック、問診と生活習慣データ

・日本脳卒中学会－約4万6千症例

脳卒中急性期治療データベース

使用データベースと開発言語：インターシステムズ社－Cache、Msm

統計解析ソフト：IBM社－SPSS

3. 医学研究から得られたエビデンス (19-23)

1. 脂質関係

- ・LDL-Cが高い群の方が総死亡率は低い。中性脂肪が高い群の方が総死亡率は低い。
伊勢原市住民健診結果をベースラインデータとして、アウトカムは住民基本情報および厚労省人口動態情報を使用したコホート研究
- ・卵の摂取量とLDL-Cは関係しない。
上越保健所管内の住民健診、事業所健診、人間ドックの間診と結果を統合して使用した横断的研究
- ・日本人では中性脂肪に対する米国薬物治療基準(1,000mg/dL)以上の人は極めて素こない。
総合健診医学会が集めた健診結果から分布を求めた横断的研究
- ・日本版高脂血症の人は脳卒中の発症リスクが低い。
日本脳卒中学会の患者データと、郡山市住民データから、性と年齢をマッチングして、非高脂血症、高脂血症非治療、高脂血症治療の比率を比較した症例対象研究
- ・日本版高脂血症の人は脳卒中になっても症状が軽く、入院しても死亡率は低い。
脳卒中学会の患者データで、脳梗塞（心原性、アテローム性、ラクナ）、脳内出血、くも膜下出血ごとに、高脂血症と非高脂血症で臨床指標（modified Rankin Scale, Japanese Stroke Scale, NIH-Stroke Scale）と入院中死亡率をMann-Whitney Uで検定した。

2. 肥満とメタボリックシンドローム

- ・BMIが国際基準（30以上）の人は約2.5%である。
日本肥満学会の基準（25以上）では約25.0%と10倍になってしまう。
総合健診医学会のデータで計算した横断的研究
- ・BMIが25.0～26.9の群が最も死亡率は低い。
伊勢原市と郡山市を統合した住民コホート研究
- ・日本人は男女とも、どの年齢層でも、肥満（BMI \geq 30.0）よりヤセ（BMI $<$ 18.5）が多い。
総合健診医学会のデータで計算した横断的研究
- ・男性で胴回り85.0cm以上の人は約50%である。
郡山市住民健診データによる横断的研究
- ・国際基準（BMI \geq 30）に従うと、ヘソの位置での胴回りの基準は、男性100cm、女性98cmとなる。
郡山市住民健診データによる横断的研究

3. 血圧

- ・加齢に伴い血圧が上昇するのは、正常な集団である。

総合健診医学会のデータで計算した横断的研究

- ・降圧剤で 20mmHg 以上下げた群では、非服薬者と比較すると 5 年間の死亡率は 10 倍になっていた。

郡山市住民健診と問診データおよび住基情報によるコホート研究

4. 糖尿病

- ・糖尿病発症者の 55%は、BMI が 25.0 未満であった。

郡山市事業所健診と問診データによる横断的研究

4. 医学研究に役立つ条件

1. アウトカムが登録されている。

例) 健診結果+住民異動情報

2. 長期間の一貫したデータベースを維持する。

例) 制度に依存しないシステム

3. しっかりした契約関係

例) 大学との研究委託契約

4. 開発者と運用者の信頼関係

例) マニュアル整備、障害対応、制度変更対応

5. 現場の優秀な人材とその育成

例) 毎年、学会発表をしてもらう。

5. まとめ

すべての医療保健情報システム開発が医学研究へとつながるわけではない。設計時から、医学研究に必要なデータの登録を行うこと、アウトカムを持つシステムとの連携を行うこと、個人情報の利用に対する理解と協力を得る努力をすることなどが大切である。

文献

- (1) A Pulmonary Function Analysing System. MEDINFO 80, Proceedings of the Third World Conference on Medical Informatics 1980:524-528
- (2) IRCU のミニコンによる情報処理. 第 13 回日本 ME 学会大会資料集 1974;510-511
- (3) オーダエントリを特徴とした総合病院情報システムの開発と評価に関する研究. 医用電子と生体工学 1984;22:9-16
- (4) 病院情報システムから地域医療システムまで—東海大学医学部・医学情報部—. 医療とコンピュータ 1992;5:1-6
- (5) Health and welfare data on optical memory cards in Isehara city. Medical Informatics 1996;21:69-80
- (6) Disaster Drills and Continuous Medical Education using Satellite-based Internet. Methods of Information in Medicine(Schattauer) 2000;39(4-5):343-347

- (7) 地域における保健医療情報の共有ークリニック用電子カルテと患者用マイ電子カルテ。
病院管理 2004;41:103-104
- (8) International concerted action on collaboration in telemedicine: Recommendations of the G-8 Global Healthcare Applications Subproject-4, **TELEMEDICINE JOURNAL and e-HEALTH** 2002;8:149-157
- (9) 地域医師会活動を支援するネットワークシステムの構築とその評価。医療情報学 2004; 24:125-131
- (10) テレビ電話による在宅医療福祉支援。医療とコンピュータ 1998;9(10):11-15
- (11) 朝日町在宅医療ネットワーク. 第34回日本エムテクノロジー学会大会論文集 2007:33-38
- (12) 西川町における携帯型保健指導システムの開発。第23回日本Mテクノロジー学会大会論文集 1996:68-71
- (13) 市町村における統合保健情報システムの開発。第32回日本エム・テクノロジー学会大会論文集 2005:21-29
- (14) 伊勢原市と郡山市健診センターにおける特定健診・特定保健指導システム。Mumps 2008;24:3-8
- (15) 職域と市町村を統合した健診データベースの開発と解析。医療情報学 2000; 20(Suppl. 2):678-679
- (16) 事業所健診システムの再開発。Proceedings2001-M Technology Association of Japan 2001:1-4
- (17) 年齢別基準値の意義と地域および年次比較。総合健診 2004;31:95-105
- (18) 脳卒中データバンク、中山書店、東京、2005
- (19) 検査値と病気 間違いだらけの診断基準。太田出版、東京、2006
- (20) メタボの罫 「病人」にされる健康な人々。角川 SSC 新書、東京、2007
- (21) 「ちょいメタ」でも大丈夫。PHP 研究所、東京、2008
- (22) コレステロールと中性脂肪で、薬は飲むな。詳伝社新書、東京、2008
- (23) 長寿のためのコレステロールガイドライン 2010年版。中日出版、名古屋、2010

獲得した知識の有効的格納について -獲得した知識を有効に使うための方法論-

Key Words

Massive Data (Clinical Data & Administrative Data)
 Longitudinal Data
 Require Expertness and Know-How
 Persons working at Professional Organization

紀ノ定 保臣

岐阜大学大学院
 医学系研究科 医療情報学分野
 連合創薬医療情報研究科 医療情報学専攻

研究の進め方

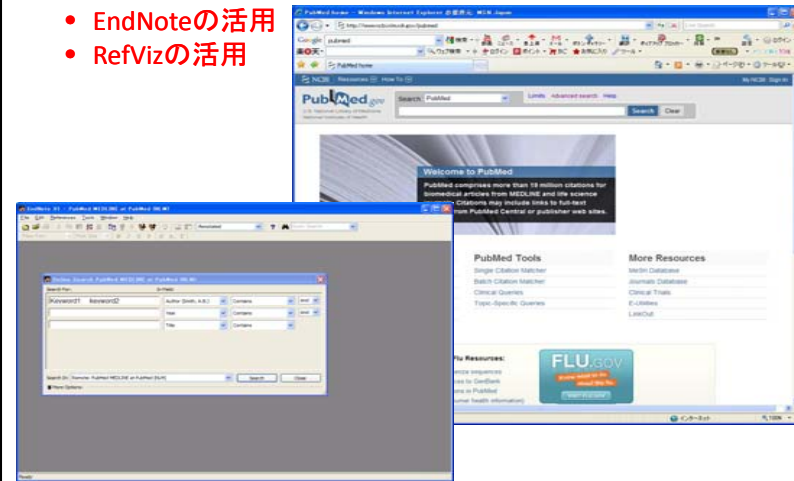
1. 研究の課題を決める
2. 先行研究(論文等)を調べる
3. 研究課題の絞り込み・範囲を決める
4. 研究(実験、調査等)を進める
5. データを収集する
6. データを分析する
7. データをまとめる
8. 知的財産を確保する
9. 研究成果を発表する

PubMedを活用する

1. キーワードで検索
2. アブストラクトのダウンロード
3. アブストラクトを分析する
 - EndNoteの活用
 - RefVizの活用

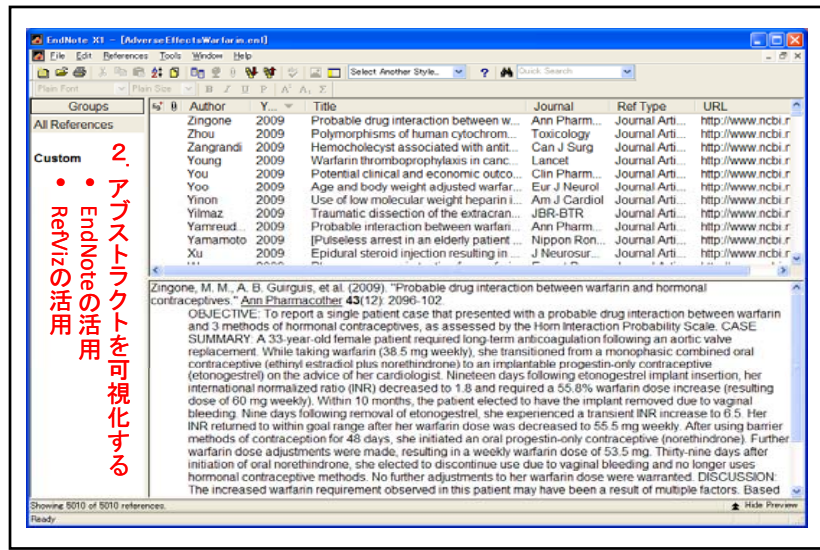
1. 先行研究を調査する

- EndNoteの活用
- RefVizの活用



2. アブストラクトを可視化する

- EndNoteの活用
- RefVizの活用



3. アブストラクトを分類する

- EndNoteの活用
- RefVizの活用

Title: Tranexamic acid mouthwash--a prospective randomized study of a 2-day regimen vs 5-day regimen to prevent postoperative bleeding in anticoagulated patients requiring dental extractions
Authors: Carter, G., Goss, A.

4. 対象を絞り込む

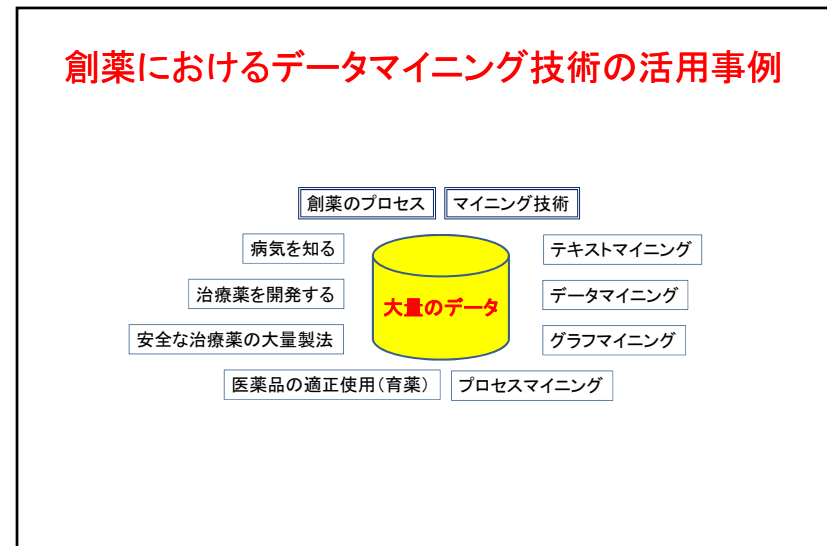
- EndNoteの活用
- RefVizの活用

Title: Safety of discontinuation of anticoagulation in patients with intracranial hemorrhage at high thromboembolic risk
Authors: Phan, T. G.; Koh, M.; Wijdicks, E. F.
Pub Date: 2000

4. 対象を絞り込む

- EndNoteの活用
- RefVizの活用

Title: Warfarin and aspirin as thromboprophylaxis in atrial fibrillation
Authors: Lip, G. Y.; Lowe, G. D.
Pub Date: 1996



創薬においてデータマイニング技術を活用するために

医薬品の開発

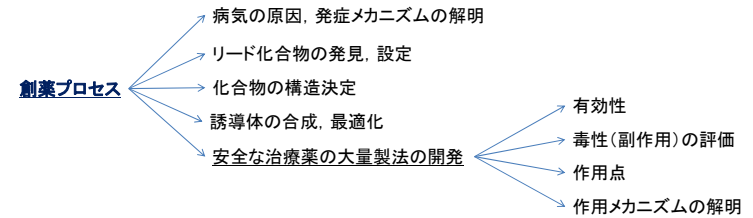
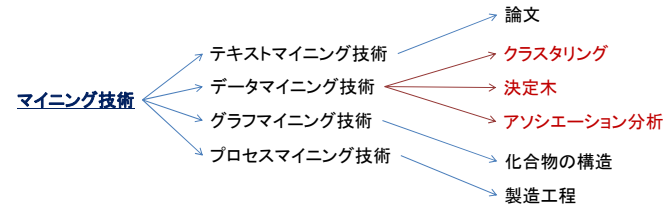
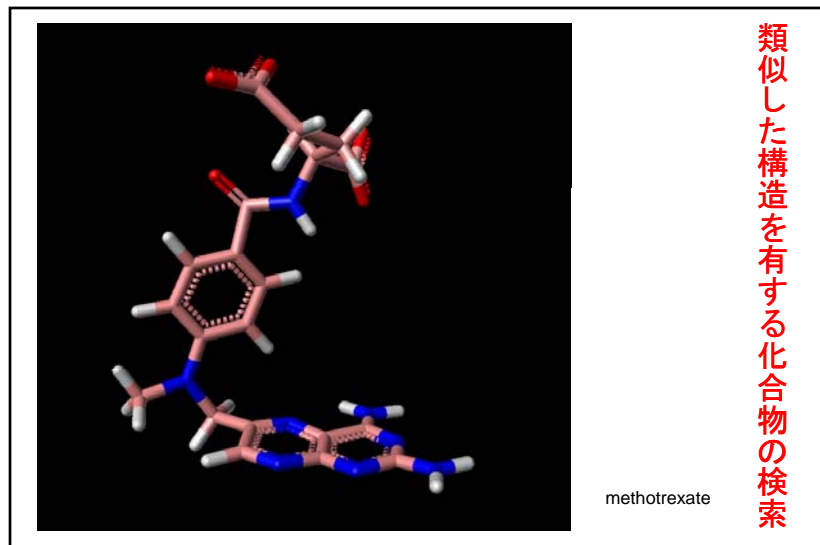
1. 病気の発症メカニズムを知る：(例)疾患関連遺伝子の発見やその発現機序を解明
2. 治療薬を開発する：(例)リード化合物の発見と構造決定、最適化
3. 安全な治療薬の大量製法を開発する：(例)有効性及び毒性の評価、作用点・作用機序の解明

マイニング技術

1. 大量に蓄積されたデータを解析し、その中に潜む項目間の関係やパターンを探し出す手法
2. データベース技術、データマイニング技術、テキストマイニング技術の知識が必要
3. データマイニング技術：(例)クラスタリング、アソシエーション分析、など
4. テキストマイニング：(例)抽出語リストの作成、抽出語検索、クラスタリング、共起ネットワークの作成、など

医薬品の適正使用

1. 有害事象分析からの新たな知見の発見

類似した構造を有する化合物の検索

類似構造の検索用画面

類似構造の検索結果

知見のデータベース化

●形質転換成長因子β受容体

文献の管理

Category	Property	Value
Common Properties	Journal	Cell Phys
Local Properties	Journal Reference	v286 i6 pC1410 (2004)
All References	Journal Link	http://ajpcell.physiology.org/cgi/content/full/286/6/...
Reference 1	MedLine Reference	14960415:1025
Reference 2	Sentence	ERK5 and p38 MAPK stimulate MEF2 activity by direc...
Reference 3	Source	Ariadne Signaling Pathways

Don't show properties with empty values for references

Linked Entities: MEF2, ERK5, p38, MAPK, etc.

Relation Type: ProtModification

Property Value Table:

Property	Value
# of References	39
Connectivity	2
Effect	positive
Mechanism	phosphorylation

リン酸化反応

MEF2
白血球遊走促進因子

大規模医薬品有害事象DBの構築と原因候補医薬品の抽出機能開発

統合データベース管理システム My SQL

- Drug Bank Release 1.0
- AHFS PHARMACOLOGY THERAPEUTIC CLASSIFICATION
- U.S. FDA AERS
- MedDRA/J
- JAPIC 市販医薬品届付文書情報
- AERS 報告薬品名・一般薬品名・薬効分類変換テーブル
- 有害事象副作用分類テーブル
- 健康データおよび代謝パスデータベース参照結果

データ管理 My SQL

医薬品候補物質分子構造データ

整形済み統合バイオバンク教育用データベース

データ抽出

教育用データベース管理システム My SQL

対象データの出力 csv xls

データ分析

R / Visual Mining Studio / MATLABネットワーク解析 / ChemBioOffice Ultra 2010

可視化

アソシエーションルール

ネットワーク図

代謝パスデータベース (外部接続)

データマイニングで薬害予防

AERSや薬剤情報
大量蓄積データ活用
新薬開発や人材育成も寄与

医薬品の適正使用に向けた取り組み
安心安全な医療提供
システム構築

SEEDS
イノベーション

紀ノ定 保臣教授

創薬・医療情報環境を活用した高度医療専門職 育成のための実践的教育手法の確立

岐阜大学
GFU UNIVERSITY
大学院連合創薬医療情報研究科

創薬・医療情報環境を活用し、次世代の創薬・個別化医療を推進することができる高度な技術と専門知識を備えた人材を効果的かつ実践的に育成する教育手法を確立する。

創薬・医療情報環境を活用し、リアルな情報を教育研究材料として扱うことにより、より実践的な教育効果の向上を図る。

社会のニーズに応える人材育成

- 各種医療情報等を活用した個別化医療へ対応できる技術者
- ゲノム創薬研究の推進が図れる研究者

・データバンクから得られる新しい知見を利用した疾病の早期発見、発症・重症化防止手法の開発等応用力養成のための知識活用型実践教育の実施
・マネジメント能力養成のためのプロジェクト指向型実践教育の実施

創薬・医療情報環境を活用した高度な教育

データマイニング手法の開発により得られる新しい知見等有用な情報

データ活用手法の教育

加齢や疾病状態への生理的な経時変化の可視化

情報と知識の活用

創薬医療マネジメント (平成22年度)

(平成20年度) (平成21年度) (平成22年度)

創薬・医療情報環境を活用した実践教育をシステム化し、効果的・実践的教育手法として確立

今後取り組むべき課題

Key Word

IT ↔ Data, Intelligence Analysis ↔ Optimization

何が強みか

Examples of "Optimization"

- Control to something
- Reduce the cost
- Speed up the process
- Improve the Quality

教育、人材育成 研究

どのように効果的なのか

何が身につくのか

基礎力 (学問、技術) 応用する力

社会への貢献

医療情報学を通じた新しい知の再構築

Better Value

Better Life

Better Health

Better Health

知恵

知識

情報

データ

蓄積した知識の活用

ルールの発見と知識の表現・蓄積

データの価値増分

データの収集

→ 構造化された知識の「作成」・「構築」・「活用」

医学・医療・生命科学分野における新しい知識学構築への挑戦

獲得した知識の有効的格納について -獲得した知識を有効に使うための方法論-

Key Words

Massive Data (Clinical Data & Administrative Data)
Longitudinal Data
Require Expertness and Know-How
Persons working at Professional Organization

紀ノ定 保臣

岐阜大学大学院
医学系研究科 医療情報学分野
連合創薬医療情報研究科 医療情報学専攻

研究の進め方

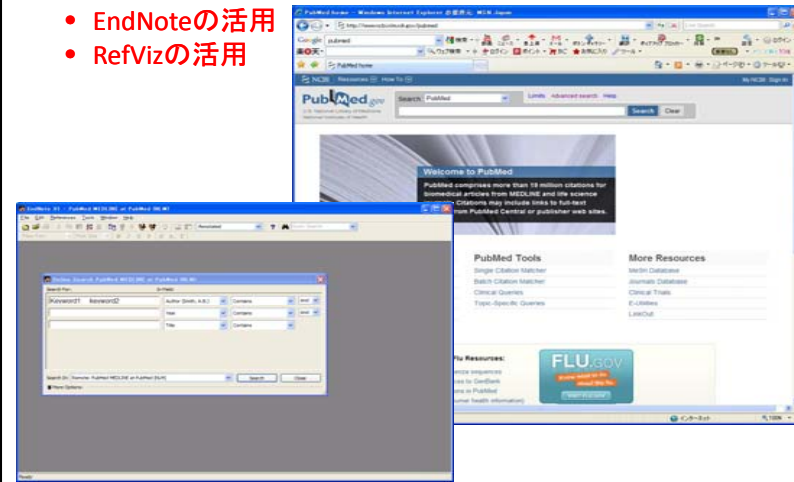
1. 研究の課題を決める
2. 先行研究(論文等)を調べる
3. 研究課題の絞り込み・範囲を決める
4. 研究(実験、調査等)を進める
5. データを収集する
6. データを分析する
7. データをまとめる
8. 知的財産を確保する
9. 研究成果を発表する

PubMedを活用する

1. キーワードで検索
2. アブストラクトのダウンロード
3. アブストラクトを分析する
 - EndNoteの活用
 - RefVizの活用

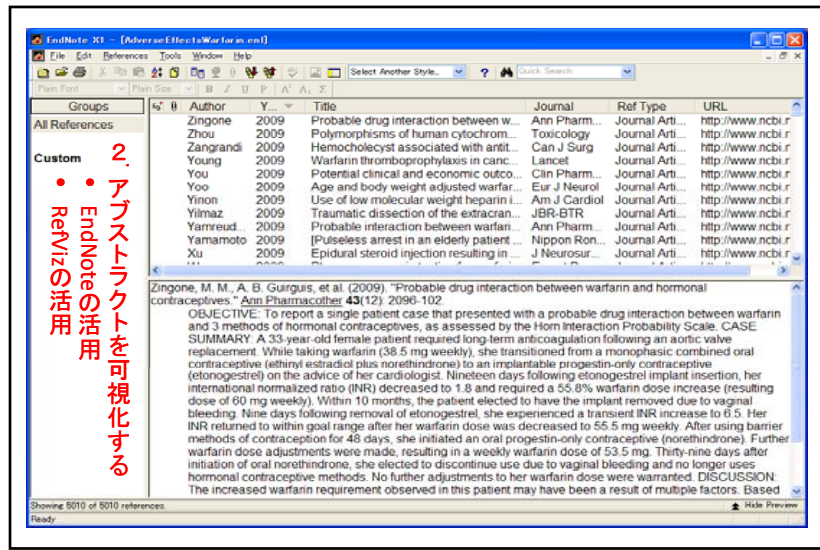
1. 先行研究を調査する

- EndNoteの活用
- RefVizの活用



2. アブストラクトを可視化する

- EndNoteの活用
- RefVizの活用



3. アブストラクトを分類する

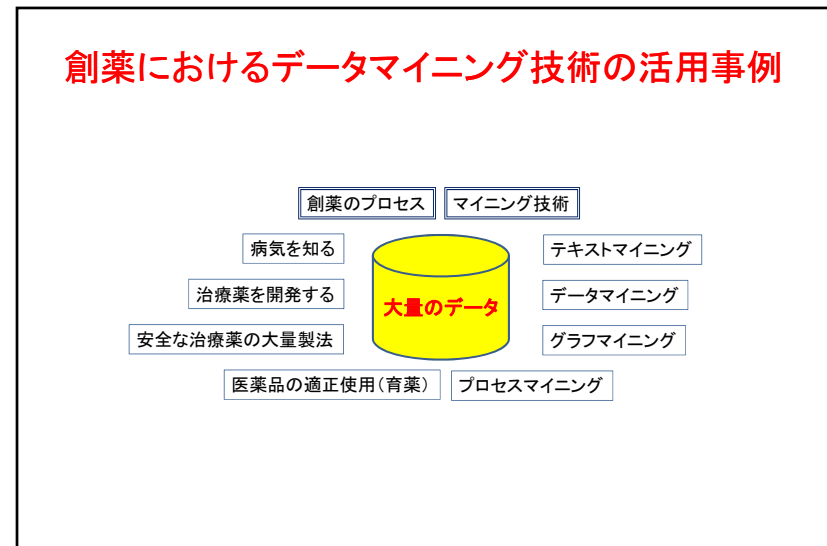
- EndNoteの活用
- RefVizの活用

4. 対象を絞り込む

- EndNoteの活用
- RefVizの活用

4. 対象を絞り込む

- EndNoteの活用
- RefVizの活用



創薬においてデータマイニング技術を活用するために

医薬品の開発

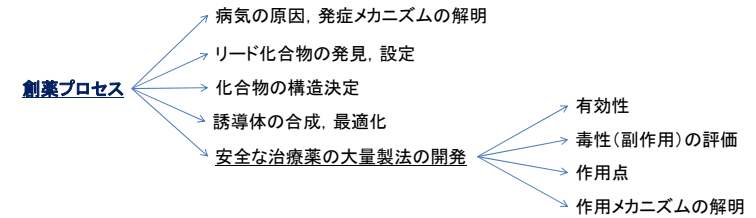
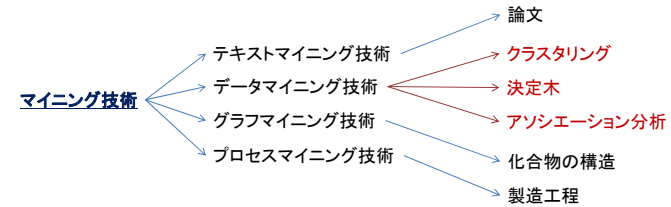
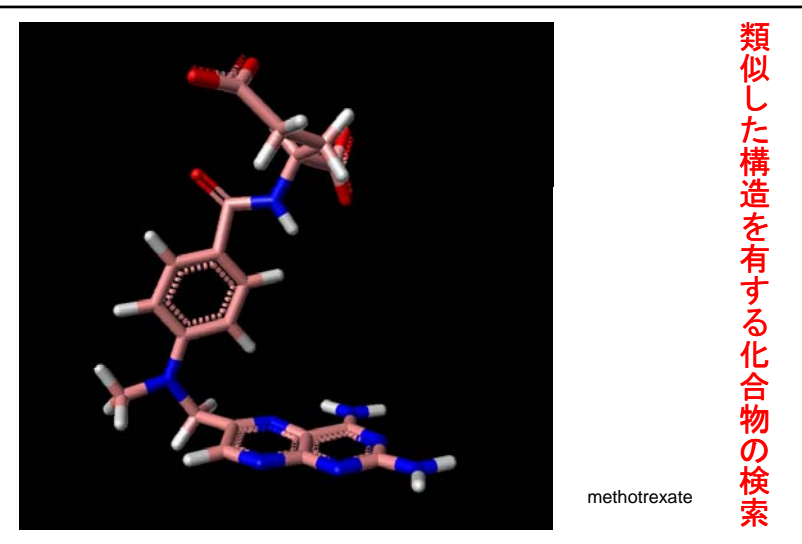
1. 病気の発症メカニズムを知る：(例)疾患関連遺伝子の発見やその発現機序を解明
2. 治療薬を開発する：(例)リード化合物の発見と構造決定、最適化
3. 安全な治療薬の大量製法を開発する：(例)有効性及び毒性の評価、作用点・作用機序の解明

マイニング技術

1. 大量に蓄積されたデータを解析し、その中に潜む項目間の関係やパターンを探し出す手法
2. データベース技術、データマイニング技術、テキストマイニング技術の知識が必要
3. データマイニング技術：(例)クラスタリング、アソシエーション分析、など
4. テキストマイニング：(例)抽出語リストの作成、抽出語検索、クラスタリング、共起ネットワークの作成、など

医薬品の適正使用

1. 有害事象分析からの新たな知見の発見

類似した構造を有する化合物の検索

類似構造の検索用画面

類似構造の検索結果

知見のデータベース化

transforming growth factor-beta receptor
●形質転換成長因子β受容体

MEF2 activation in differentiated primary human skeletal muscle cultures requires coordinated involvement of parallel pathways

白血球遊走促進因子

文献の管理

Category	Property	Value
Common Properties	Journal	Cell Phys
Local Properties	Journal Reference	v286 i6 pC1410 (2004)
All References	Journal Link	http://ajpcell.physiology.org/cgi/content/full/286/6/...
Reference 1	MedLine Reference	14960415:1025
Reference 2	Sentence	ERK5 and p38 MAPK stimulate MEF2 activity by direc...
Reference 3	Source	Ariadne Signaling Pathways

Linked Entities

Relation Type: ProtModification

Property	Value
# of References	39
Connectivity	2
Effect	positive

リン酸化反応

MEF
白血球遊走促進因子

原因候補医薬品の抽出機能開発

大規模医薬品有害事象DBの構築と

統合データベース管理システム
My SQL

- Drug Bank Release 1.0
- AHFS PHARMACOLOGY THERAPEUTIC CLASSIFICATION
- U.S. FDA AERS
- MedDRA/J
- JAPIC 市販医薬品届付文書情報
- AERS 報告薬品名・一般薬品名・薬効分類変換テーブル
- 有害事象副作用分類テーブル
- 健康データおよび代謝パスデータベース参照結果

データ管理 My SQL

医薬品候補物質分子構造データ

整形済み統合バイオバンク教育用データベース

データ抽出

対象データの出力 csv xls

データ分析

R / Visual Mining Studio / MATLABネットワーク解析 / ChemBioOffice Ultra 2010

可視化

アソシエーションルール

ネットワーク図

代謝パスデータベース (外部接続)

データマイニングで薬害予防

AERSや薬剤情報
大量蓄積データ活用
新薬開発や人材育成も寄与

医薬品の適正使用に向けた取り組み
安心安全な医療提供
システム構築

SEEDS
イノベーション

紀ノ定 保臣教授

創薬・医療情報環境を活用した高度医療専門職 育成のための実践的教育手法の確立

岐阜大学
GFU UNIVERSITY
大学院連合創薬医療情報研究科

創薬・医療情報環境を活用し、次世代の創薬・個別化医療を推進することができる高度な技術と専門知識を備えた人材を効果的かつ実践的に育成する教育手法を確立する。

創薬・医療情報環境を活用し、リアルな情報を教育研究材料として扱うことにより、より実践的な教育効果の向上を図る。

社会のニーズに応える人材育成

- 各種医療情報等を活用した個別化医療へ対応できる技術者
- ゲノム創薬研究の推進が図れる研究者

・データバンクから得られる新しい知見を利用した疾病の早期発見、発症・重症化防止手法の開発等応用力養成のための知識活用型実践教育の実施
・マネジメント能力養成のためのプロジェクト指向型実践教育の実施

創薬・医療情報環境を活用した高度な教育

データマイニング手法の開発により得られる新しい知見等有用な情報

データ活用法の教育

加齢や疾病状態への生理的な経時変化の可視化

情報と知識の活用

創薬医療マネジメント (平成22年度)

(平成20年度) (平成21年度) (平成22年度)

創薬・医療情報環境を活用した実践教育をシステム化し、効果的・実践的教育手法として確立

今後取り組むべき課題

Key Word

IT ↔ Data, Intelligence Analysis ↔ Optimization

何が強みか

Examples of "Optimization"

- Control to something
- Reduce the cost
- Speed up the process
- Improve the Quality

教育、人材育成 研究

どのように効果的なのか

何が身につくのか

基礎力 (学問、技術) 応用する力

社会への貢献

医療情報学を通じた新しい知の再構築

Better Value

Better Life

Better Health

Better Health

知恵

知識

情報

データ

蓄積した知識の活用

ルールの発見と知識の表現・蓄積

データの価値増分

データの収集

→ 構造化された知識の「作成」・「構築」・「活用」

医学・医療・生命科学分野における新しい知識学構築への挑戦

院内紙文書を利活用するための文書画像処理 ～紙文書の電子化，そして検索への試み～

川中 普晴¹，山本 皓二²，高瀬 治彦¹，鶴岡 信治³
 三重大学大学院工学研究科¹ 鈴鹿医療科学大学²
 三重大学大学院地域イノベーション学研究科⁴

1. はじめに

近年，HIS の普及に伴い，様々な診療データが電子化されつつある．一方，HIS 導入前に作成された診療文書は現在，紙文書として保管されている．これらの紙文書は，過去の症例や治療内容やその結果等が詳細に記録されており，臨床研究においては貴重な知識データベースである．しかしながら，これらを電子データに変換する際，手作業での入力では時間と費用がかかるといった理由等により，現在のところ有効活用されていないのが実情である．

このような背景から，これまで筆者らは，院内にて蓄積されている紙文書（主に表形式の紙文書）を類似症例検索に活用するための文書画像処理に関する研究を進めてきた[1][2]．また現在，これらの成果を活用して，院内運用するためのシステムの構築と，実証実験のための準備・検討を進めている．

今回の発表では，筆者らがこれまでに取り組んできた研究の概要を述べるとともに，研究成果を活用して作成しているシステムを紹介する．本発表の内容をもとに，院内での紙文書の利活用方法，今後の展開等について議論できれば幸いである．

2. 文書画像処理とその医療への応用

2.1 文書構造の認識と電子データ化

図 1 に文書の電子化に関する概要を示す．本研究において紙文書の電子化プロセスは，

- ①文書種・文書構造の認識
- ②内容認識，検索用データの生成

という 2 つから構成される．

まず，処理対象となる紙文書は光学スキャナを用いて e 文書法に適合する規格の文書画像に変換される．次に，これらの画像に対して黒画素の直線性を用いて文書中の罫線とそれらの交点（本研究では節点と呼んでいる）を抽出する．これら抽出された情報は，予め登録されているマスターと照合され，入力された文書の種類が決定される．次に，得られた各節点をトレースすることにより，文書中の各セル部分が切り出される．

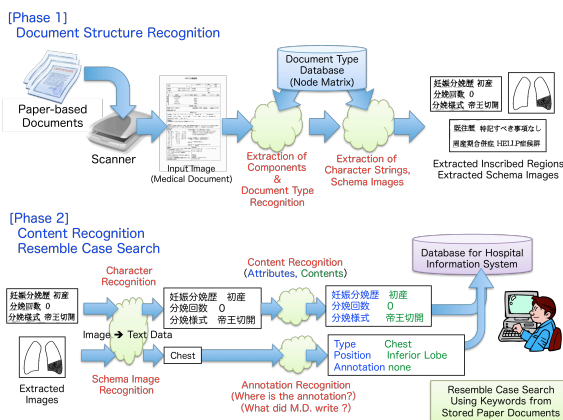


図 1 紙文書の電子化プロセス

次に、これら切り出された部分の内容は、文字認識エンジンを用いてテキストデータ化される。なお、本システムでは未記入の文書をスキャンした場合、予め文書内の各セルに書かれていた項目がテキスト化され、その情報が各セルの項目タグとなる。すなわち、未記入の文書をスキャンさせればマスター情報が、また記入済みの文書を入力すればマスターとの差分（記入部分）がテキストデータ化される。

2.2 シェーマ認識とその利活用

一般的に、文書中に書かれるシェーマや記入されるアノテーションは、文書の内容を反映していると考えられる。そこでここでは、切り出されたシェーマ画像についても認識し、検索タグの一つとして活用することとした（図 2）。

本システムでは、使用される各シェーマによってその形状の特徴が大きくことから、手書き文字認識エンジンとして用いられている「加重方向指数ヒストグラム法」と「疑似ベイズ識別関数」を用いた認識エンジンを用いることとした。また、認識に成功したシェーマについては、辞書に登録されている未記入のシェーマ画像との差分を得ることによりアノテーションが書かれた部分を検出し、その部位名も検索キーとして利用できるようにした。

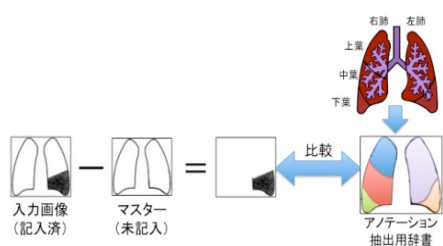


図 2 シェーマとアノテーションの認識

これにより、スキャンされた文書中の各セルの内容だけでなく、これまでは検索キーとならなかったシェーマそのものについても、検索キーとして利用可能となる。

3. 認識精度と出力されたファイル

3.1 文書構造認識・文字認識の精度

三重大学医学部附属病院病歴室にて保管されている表形式の退院サマリ文書ならびに学内で使用されている表形式文書を用いた評価実験を行った。実験の結果、表構造の認識に失敗した（各セルが切り出しできなかった）ケースは無かったが、セル内の文字を誤認識するケースが見受けられた。また、入力する文書によって若干の変化はあるものの、OCR の認識精度は概ね 90% であった。この結果については、大部分が OCR エンジンに由来するものと考えられる。しかしながら、セルの切り出し処理や記入部分の抽出処理において罫線や文字の一部が残っていると誤認識を誘発する可能性が高い。今後は、セルの切り出し方法やノイズ除去処理の見直しと改良、さらには医学辞書やオントロジーを取り入れて認識結果を修正する方法等により、（感覚的にはあと数%程度）認識率を向上させることができると思われる。

3.2 シェーマ認識の精度

シェーマの認識については、98% の認識精度が得られた。これは先に示した文字認識の精度と比較して高い値となっている。シェーマ認識の場合、文字認識と比較してカテゴリ数（すなわちシェーマの種類）が少ないこと、また認識用辞書作成の際に、様々なパターンのアノテーションが加えら

れた画像をしたことにより、多くのケースに対応できたと考えられる。現在、シェーマ認識に関しては、さらなる精度向上と誤認識の検出に関する研究を進めている段階である。

3.3 出力された XML ファイルの例

図 3 に、入力文書画像と出力された XML ファイルの一部を示す。図は、文書中にシェーマが複数個存在する場合の例である。この場合においては、以下のようなタグが XML ファイル中に 2 カ所追加されていた。

```
<schema>
  <name>腹部</name>
  <annotation-part>--</annotation-part>
</schema>
```

図からも明らかであるが、複数のシェーマが文書中に存在する場合、その数だけ schema タグが追加されていることが分かる。なお、—印部分はアノテーション部位が辞書に追加されていなかったためである。この部分については、辞書に部位名を追加することにより、対応可能である。

4. 作成しているシステム

現在、筆者らが開発しているシステムの一部を示す。本システムでは、できる限りユーザビリティを考慮し、各種設定部分はウィザードを採用している。また、通常使用時にはシンプルなウィンドウのみを表示し、使用者の誤操作を防ぐような GUI も現在作成している段階である。

なお、図中左側の入力画像から右側の XML を作成するまでの処理時間は数秒程度である。これは、1 日あたり 2000 枚以上の文書をスキャンする必要が出てきた場

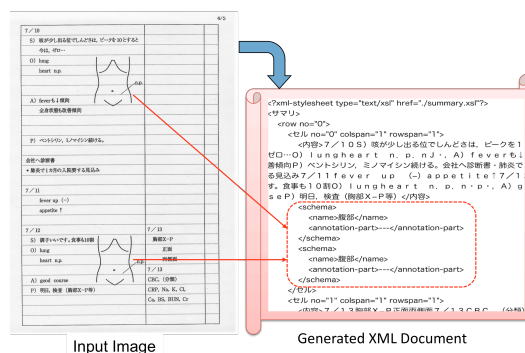


図 3 入力画像と出力された XML (例)

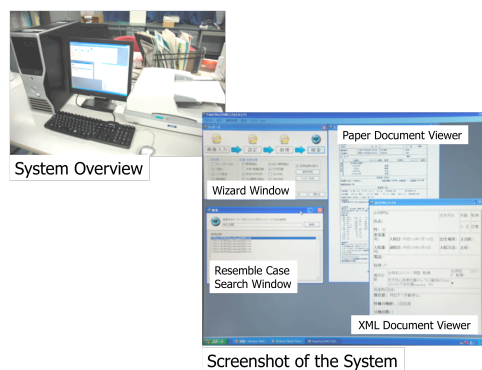


図 4 作成しているシステムの一部

合においても十分に対応できる処理速度であり、実運用においても問題ない速度であると考えられる。

5. おわりに

本稿では、文書画像処理についてその概要を述べるとともに、これまでの研究成果を用いて作成してきたシステムを簡単に紹介した。今後は現在の開発を進めるとともに、複数の施設での実験を進めていきたい。

参考文献

[1] H. Kawanaka *et. al*, “Document Recognition and XML Generation of Tabular Form Discharge Summaries for Analogous Case Search System”, *Methods of Information in Medicine (Schattauer)*, 46,6, 700-708, 2007
 [2] 林裕樹, 他: 文書種ならびに文書構造認識機能を備えたスキャンングシステムの開発とその評価, 第 30 回医療情報学連合大会, pp.895-898, 201

診療文書作成プログラムの試作(Caché, jQueryを使用して)

土屋喬義¹⁾²⁾、鈴木智之¹⁾、田中千恵子¹⁾、木村一元³⁾

1) 土屋小児病院、2) 獨協医科大学小児科

3) 獨協医科大学病院医療情報センター

土屋小児病院では 1991 年より U-MUMPS 上の医事システムと M 言語システムである MSM 上で開発した検査、及び医療情報表示システムを DDP 接続したネットワーク上で運用してきた。この間当院の使用したユーザーインターフェースはシリアル接続のダム端末、Telnet 端末、Windows 端末より VB で MSM への接続、Web ブラウザを介しての接続と変遷して来た。当院の主力として使用している MSM は既に 1998 年に InterSystems 社に吸収合併された時点で開発が停止され、サポート OS は Windows NT のままである。しかしながらその安定した性能と MSM PDQWeb という優れた Web インターフェースを持っており、これらを用いた動的な Web ページを作成することが出来る。

近年 Ajax を使用した綺麗で画面遷移を伴わないインタラクティブな Web 画面を多く見る様になった。Ajax でのアプリケーション開発は複雑でコード量が多く理解するのに困難そうに見えたが、jQuery という Ajax 用アプリケーションフレームワークを使用することにより簡便にリッチな画面を作る事が出来、これらの技術を取り込んだ画面を作る事を試みた。Ajax で通信を行う場合 UTF-8 で通信する必要があるが、当院で使用している MSM がシングルバイトサーバーで使用している文字コードが S-JIS のため Ajax を単純に適応するのは困難と考えられた。既に採用している医事システムが Caché 上で動作していたことと Caché と MSM 間の DDP 通信速度が遅く大量のデータ参照に適さなくなって来たため、開発を Caché で開発する事とした。

MSM での開発に慣れていたこともあったため MSM PDQWeb と非常に近い仕組みで動く WebLink を使用した。実際、設定ファイルの変更のみで MSM のコードはそのまま動作した。また Caché は UTF-8 が標準であり、また文字コードの変換機能も備えているため Ajax, jQuery とも問題なく WebLink 上で動作した。今回開発に使用した Caché は 5.0 である。

試作したプログラムは入院計画書と退院サマリーの作成ページを作成した。今まで使用していた固定した画面を編集するタイプに比べるとインタラクティブに格段に使用感が良く、また表示する文書、オブジェクトを Web 画面の div 要素等の中に他の Web ページやデータを画面遷移せずに置けるためより自由なデザインとコードや Web ページの再利用が有効に行えるようになった。現在実際の現場で使用できるよう改良を続けている。

ComeJisyo の退院時サマリへの適用

鈴木隆弘¹ 土井俊祐¹ 相良かおる² 高林克日己¹
 千葉大学医学部附属病院企画情報部¹ 西南女学院大学²

目的

近年の電子カルテの普及によって蓄積されたテキストから、診療に役立つ情報を抽出することが期待されている。しかしながら医療テキストには略語を含む専門用語が多く使われているため、一般文書用に辞書で精度の高い解析を行うことは困難である。

我々は医療施設で使われる用語を集めた、分かち書き用の辞書である ComeJisyo を用いて退院サマリの解析を行い、効果を比較したので報告する。

対象と方法

千葉大学医学部附属病院情報システムに保存された退院サマリ文書のうち、症例数が年間 20 例以上の 147 種類の DPC を対象とし平成 20 年退院分の約 9000 件の退院サマリを DPC 別の文書ベクトル空間の構築に、21 年の約 8900 件のサマリを DPC 自動判定の検証に用いて解析を行った。

ベクトル空間モデルを作成するための形態素解析には Mecab の ver. 0.98 を用い、辞書には一般的用語辞書の IPADIC に加えて、当院で使われている用語を登録したオリジナル辞書と、ComeJisyo 辞書を使用した。ComeJisyo は教科書や国家試験問題、Web 上で公開されている辞書などから用語を抽出して作成され、Ver.3 からは 3 つの病院の看護サマリから抽出した用語を加え、録語数は 41,592 語となっている。

形態素解析したテキストから名詞のみを抽出し、TF-IDF 法を用いて重要度を算出し

た。次いで文書ごとに単語重要度ベクトルを作成し、モデル作成用サマリのベクトルと、検証用サマリのベクトルとを比較して類似度を算出した。類似度の検証には DPC コードを用い、辞書による相違を検証した。これらの解析には Cache を用いた。

結果

形態素解析では、IPADIC のみ:低/髄/圧/症候群/で/硬/膜/外/自己/血/注入/、ComeJisyo:低/髄/圧/症候群/で/硬膜外/自己血/注入/、千葉大辞書:低髄圧症候群/で/硬膜外/自己血/注入/のように切り分け位置の違いが認められた。DPC 判定についても、IPADIC のみの時の 14 桁一致率が 51%であったのに対して、ComeJisyo との併用では 55%に、千葉大オリジナル辞書との併用では 57%に向上がみられた。(図 1)

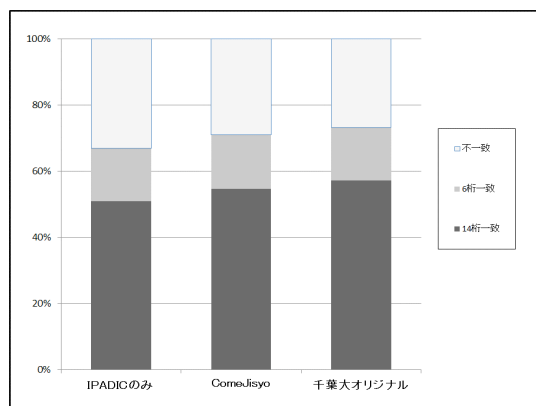


図 1. 辞書による一致率の違い

考察

ComeJisyo の適用によって医師が記載する退院時サマリの解析精度の改善が認められた。看護系のテキスト解析では更に有用であると期待される。

同じ位置で撮影された過去と現在の歯科口内法 X 線画像の同定法

檜作彰良^{*1} 中山良平^{*2} 鶴岡信治^{*3}

^{*1} 三重大学大学院工学研究科 ^{*2} 三重大学医学部附属病院 ^{*3} 三重大学大学院地域イノベーション学研究所

Computerized Identification Method for Previous and Current Dental Radiographs Taken at the Same Positions

Akiyoshi Hizukuri^{*1} Ryohei Nakayama^{*2} Shinji Tsuruoka^{*3}

^{*1} Information Processing Laboratory, Division of Electrical and Electronic Engineering

E-mail: ^{*1} hidukuri@ip.elec.mie-u.ac.jp

1. まえがき

歯科院では、虫歯や歯周病の進行度合い、治療効果を評価するため、定期的に口内法 X 線画像(図 1(a))が撮影されている。これらの画像ヘッダには撮影された位置の情報が保存されていないため、歯科医が医療用画像管理システム(PACS: Picture Archiving and Communication System)上で診断する際に、手動で画像を並べ、位置情報を登録する必要がある。位置情報が登録された過去画像と同じ位置で撮影された現在画像を自動で同定することが出来れば、現在画像を自動で並べることも可能である。そこで本研究では、同じ位置で撮影された過去画像と現在画像の自動同定法を開発する。

2. 実験試料

実験試料は、4 患者から得られた過去と現在、それぞれ口内法 X 線画像 56 枚で構成された。これらの画像は、全顎域が 14 枚の画像で構成される 14 枚で撮影された画像であり、空間分解能 0.05[mm/pixel]、濃度分解能 8[bits]である。

3. 方法

口内法 X 線画像は、複数本の歯とその歯周組織が映るように入手でフィルムを口腔内に挿入し撮影される。したがって、画像ごとに、歯の軸方向が若干異なる。そこで、歯の先端の近似直線を最小 2 乗法[1]で求め(図 1(b))、近似直線が水平になるように画像全体をアフィン変換[1]を用いて回転することにより、歯軸方向を垂直に補正する(図 1(c))。次に、歯間の画素値が低いことに着目し、各 x 座標において垂直方向の平均画素値を求める。そして、平均画素値が極小値となる x 座標において、近似直線の垂線を引くことにより、各歯領域を分割する(図 1(d))。各患者の現在と過去のすべての画像ペア(14*14 ペア)において、歯の分割線ごとに画像を移動させながら画素値に基づく相関係数を求め、相関係数が最も高くなったペアを同じ位置で撮影された画像と同定する。

4. 結果及び考察

本手法を実験試料に適用した結果、80.4%(45/56)の同定率が得られた。正しく同定できなかった口内法 X 線画像として、異なる画像に映っている数本の歯のうち、1 本または 2 本の同じ歯が重複して写っていたことが挙げられる。この問題の解決策として、3 本以上の歯の領域で相関係数を求めたり、

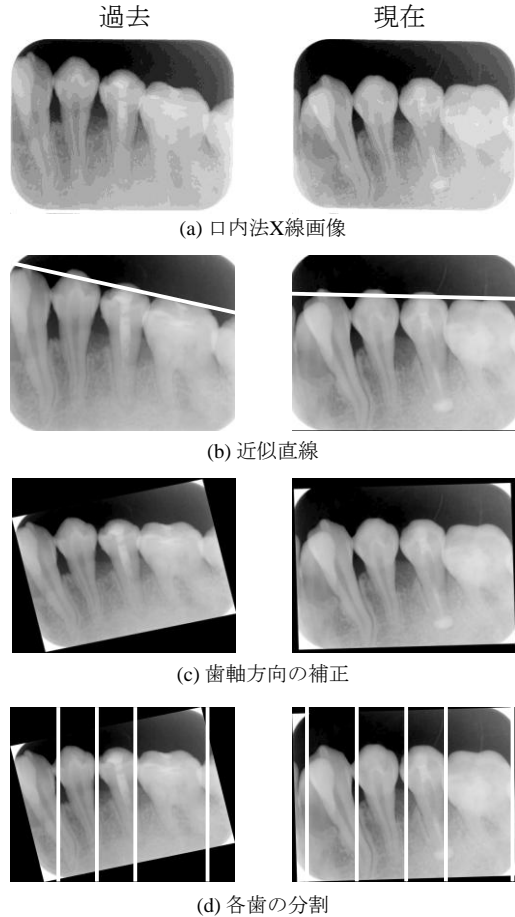


図1 提案手法の処理手順

画像上に映っている歯の位置を考慮したアルゴリズムに変更したりすることにより対処できると考えている。これについては今後検討していく予定である。

5. 結語

本研究では、同じ位置で撮影された過去と現在の口内法 X 線画像の同定法を開発した。本手法を実験試料に適用した結果、80.4%(45/56)の同定率が得られた。今後、同定率が向上すれば、本手法が歯科医の診断業務の手助けとなる可能性がある。

文献

[1] R. C. Gonzales, R. E. Woods, Digital Image Processing, 2nd ed, MA: Addison-Wesley, (1992).

オープンソースソフトウェア GT.M 2011

澤田 潔¹ CAMTA-JP¹

はじめに

GT.M(Greystone Technology MUMPS)は、AGPL3 に準拠したオープンソースソフトウェア(OSS)である。ANSI/ISO/JIS 国際標準 M 言語で、64bit Linux プラットフォーム対応、ACID トランザクションをフルサポートし、Unicode 対応などの特徴を持つハイパフォーマンスな M テクノロジー・データベース・エンジンである。

GT.M の近況

GT.M プロダクトは「Rock Solid. Lightning Fast. Secure. No compromises.」なるキャッチコピーの元、2010 年 2 月にメジャーバージョンである Ver5.4 がリリースされ、数々の改良が加えられ最新版は V5.4-002A(2011/05/17 現在)である。

一方、GT.M と SQL(JDBC)とのインターフェース・ソリューションである PIP は、2010 年 4 月に Ver0.2 がリリースされた。

GT.M Ver5.4 リリース概要

トリガ(Trigger)は、Ver5.4 で追加された最大の機能である。トリガは特定のグローバル・ノードにマッチする更新が発生した場合、あらかじめ定義されたアクションを自動実行しその応答を返す。一般的な RDB では、トリガ関数(PL/sql)、手続き型呼び出しなどと呼ばれている機能と同等である。

今後 PIP にて SQL ライクな問い合わせに対応するための重要な機能追加であるとともに、定型的な処理においてコード記述量を少なくすることが期待できる。

Ver5.4 では、他にソケット通信モジュールの改良や 64bit 版バイナリー配布などのアナウンスされている。

PIP Ver 0.2 の概要

PIP は従来、F. I. S 社の FIS ProfileTM 銀行向けリアルタイム・コア処理システムの商用ベースのインフラであった。Ver0.2 から FIS ProfileTMはプロファイルのコードベースから分離され、より GPL ライクなソフトウェアとして生まれ変わった。

日本における GT.M コミュニティ活動

筆者は GT.M 研究会(Google Group)にて GT.M およびその周辺技術についてコミュニティ活動を行っている。現在、GT.M 研究会にて GT.M プログラマーズガイドと GT.M 管理および操作ガイドの翻訳に注力している。翻訳作業の協力者を募集している。

参考文献(URL)

1. GT.M Database Engine with Extreme Scalability and Robustness.<http://www.fisglobal.com/Products/TechnologyPlatforms/GTM/index.htm>
2. GT.M 研究会 <http://groups.google.com/group/GTMstudy/>