

人が登録した症例の一覧表を CSV 形式のファイルで一括ダウンロードできる機能も付加した。こうしたサービスを「MyData」と呼び、検出内容を分析し診療にすぐに役立つ情報を表示し、報告に協力してくれた医師の労に還元できる工夫をした。

II. 電子メールによる情報還元

これまで述べたような集計が Web ページで随時閲覧できるが、深夜になると各都道府県の報告数をまとめたメールが自動的に配信される。医家向けにはメール本文に症例を登録する Web ページとそのログインアカウントが記載されている。

C. 結果

1. 報告状況と報告者数

方法で述べた ML で呼びかけたところ、運営期間中に情報提供者が 186 名、報告数は 6172 件あった(2012 年 2 月 25 日確認)。

都道府県別にみると今シーズンで情報提供者は滋賀県が最も多く 15 名であった。一方、岩手県、島根県、高知県、大分県には報告医が現れなかった。人口 10 万人当たりでは、全国平均で 0.14 名。滋賀県が最も高く 1.07 名であった。(図5)。

2. 感染症週報との相関

図6は、感染症週報(IDWR)の報告数を縦棒で、RSV-OS の報告数を点線で描いたグラフがである。図7は IDWR を縦軸に RSV-OS の報告数を横軸にした相関図である。y を IDWR の報告数、x を RSV-OS とした場合、線形近似式は「 $y = 48.657x$ 」で現され、決定係数(r^2)は 0.7243 となった。

D. 考察

1. 調査協力者の確保

感染症の流行サーベイランスであれば、本来、検出情報を報告する定点を人口に比して定めるべきである。しかし、INET は日本国内の隅々に普及し、もはや利用できない地域はない。そ

して医療関係の ML に参加する医師も、呼びかけに呼応する医師も、人口に比して存在するはずである。ならば地域ごとに医療機関を定め依頼する手順を省き、自発的に報告する医師からの検出報告を集計しても地域の流行を反映するであろう。本研究はこうした想定に基づいて、ML 参加者を中心に自主的に RSV の検出状況を登録する医師を呼びかける事から調査が始まった。

調査・研究の呼びかけに、全国から128名の臨床医が、自主的に報告して下さった。全国から協力者が現れた理由は、次のような事柄が考えられる。

- ・臨床現場で求められている情報を作り出す具体的な手法を示した。
- ・集められた情報を、極めて迅速に還元した。
- ・日本全体の報告数だけでなく、地域ごとの報告数を表示するなどの臨床に即した情報もリアルタイムに還元された。
- ・通信コストがほとんどかからず、報告者に負担が少ない。
- ・事前登録も不要で参加しやすい調査であった。
- ・報告者個別の集計結果もリアルタイムに還元し、自らの診療を解析するのに即役だった。

特に情報の還元が極めて早く行われた事には、大きな意義があるように思われる。報告されたすべての情報がすべてリアルタイムに集計され、情報提供者は速やかに臨床に役立てることができたであろう。そしてなおかつ「MyData」として、報告者個別に報告の集計を表示し、臨床現場に還元したことも、参加者のモチベーションを高めたであろう。

こうした、ML で自主的に検出状況を報告する調査研究は、「ML インフルエンザ流行前線情報データベース(ML-dlu-DB)【j】」でも行われて

いる。ML-flu-DB は、報告医が 250-300 名程度と本調査よりも多く、IDWR との相関が 0.8125 から 0.9960 と極めて高い。同じ手法を用いても、IDWR の報告数との相関は、必ずしも高いとは限らないことが分かった。

RSV-OS が IDWR との報告数の相関が低い理由として、迅速診断キットの保険適応の問題があげられる。インフルエンザ (flu) は、迅速診断キットは保険請求が可能であるが、RSV は今シーズンよりようやく外来診療でも保険請求が可能となった。これまで検査に費やしたコストを回収できず迅速診断キットの使用は、最小限に留められていた。今後は RSV の迅速診断がさらに普及し真の流行が把握できると期待する。

また、都道府県毎に見ると報告者数がない都道府県があった。RSV は成人領域では臨床的に大きな問題となっておらず、報告医は乳幼児を診療する小児科医に限られ、flu に比べると報告医は少なくなる。実際に流行がなかったのか、報告医がいなかったのか、本法では判然としない。

本法にて流行情報を収集する際には、報告者数の地域差を無くすことが今後の課題である。

3. 運用コスト

RSV-OS は、GPL II 【k】で配布されている無償のソフトウェア群で構築されており、高機能であるが開発コストが低く抑えられた。また症例報告には INET を使い、集計はサーバーが自動的に行うために、人材を必要とせず運用コストも安い。そのために RSV-OS は、流行期のみならず通年運用が可能である。

E. 結論

INET に RSV の検出情報を入力・出力するシステムを準備し、RSV の検出情報の提出を ML で呼びかけたところ、運用期間中に 186 名の情報提供者が現れた。RSV-OS は、検出状況がリアルタイムで解析表示され、迅速な情報還元が実現した。しかし、情報提供者が少なく、それをい

かに増やしていくかが本調査の課題である。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

なし

2. 学会発表

1. 西藤成雄：P-89 有志医師による「RS ウイルス・オンライン・サーベイ」の運用について．第 114 回日本小児科学会（2011 年 08 月 12 日・グランドプリンスホテル新高輪）

H. 知的財産権の出願・登録状況

1. 特許出願

なし

I. 参考 Web サイト

- 【a】 国立感染症研究所情報センター <http://idsc.nih.go.jp/index-j.html>
- 【b】 小児科フリートークメーリングリスト <http://www.ebisu.net/pedft.htm>
- 【c】 日本小児科医電子メールカンファレンス <http://jpmlc.med.tohoku.ac.jp>
- 【d】 FreeBSD <http://www.freebsd.org/>
- 【e】 Apache Software Foundation <http://www.apache.org>
- 【f】 MySQL <http://www.php.net>
- 【g】 PHP <http://www.php.net/>
- 【h】 Boutell.Com GD <http://www.boutell.com>
- 【i】 JpGraph <http://www.aditus.nu/jpgraph/>
- 【j】 ML インフルエンザ流行前線情報データベース <http://ml-flu.children.jp>
- 【k】 GPL GENERAL PUBLIC LICENSE <http://www.gnu.org/licenses/gpl.ja.html>

J. 参考文献

- 1] 西藤成雄:メーリングリスト有志によるインフルエンザ流行情報のオンライン集積と公開.病原微生物検出情報,27:p16-17,2006.
- 2] 西藤成雄:ML インフルエンザ流行前線情

報データベースの紹介. 日本医師
会,136:2439-2443,2008.

- ・ 都道府県・市町村
- ・ 報告日（自動的に記録される）
- ・ 発症日、診断日
- ・ 性別、年齢・月齢
- ・ 診断キット、判定結果
- ・ 既往歴（低出生体重、シナジス®接種歴）
- ・ 既往症（喘鳴、湿疹の有無、RAST）
- ・ 家族歴（気管支喘息、アトピー性皮膚炎）
- ・ 臨床経過（通院/入院）

図 1 報告項目

1

図 2 症例登録画面

※このWebページのURLとパスワードは紙にて周知。登録はほとんどが県参加者。

2

図 3 トップページ

3

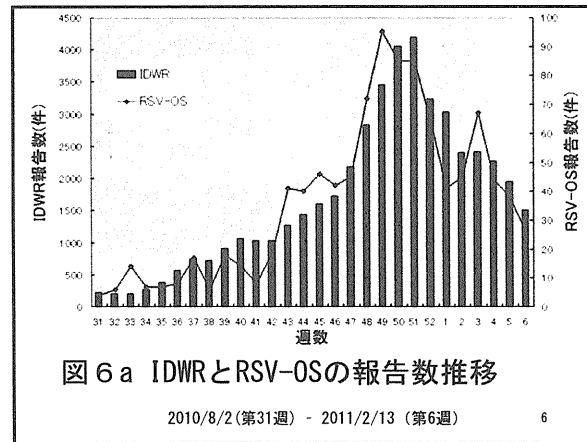
図 4 都道府県の詳細情報ページ

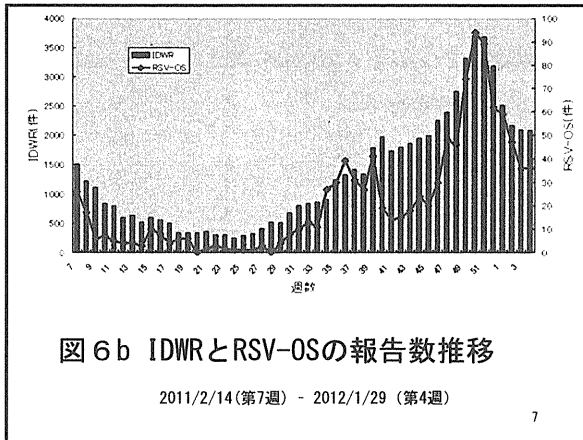
4

都道府県	報告者数	人口	10万人に對する有志の割合
1 北海道	6,151,649	0.11	25
2 青森県	2,148,619	0.14	26
3 岩手県	1,136,404	0.21	26
4 宮城県	2,242,731	0.19	29
5 秋田県	1,116,638	0.27	30
6 山形県	1,299,444	0.11	31
7 福島県	2,246,741	0.17	31
8 茨城県	2,213,770	0.13	32
9 栃木県	4,174,713	0.16	32
10 群馬県	1,916,332	0.16	34
11 千葉県	6,692,915	0.17	34
12 東京都	19,121,931	0.15	35
13 埼玉県	4,894,492	0.16	35
14 東京都	4,284,734	0.17	35
15 東京都	4,116,574	0.20	39
16 東京都	4,116,574	0.24	40
17 東京都	4,116,574	0.24	40
18 東京都	4,116,574	0.24	40
19 東京都	4,116,574	0.24	40
20 東京都	4,116,574	0.24	40
21 東京都	4,116,574	0.24	40
22 東京都	4,116,574	0.24	40
23 東京都	4,116,574	0.24	40
24 東京都	4,116,574	0.24	40

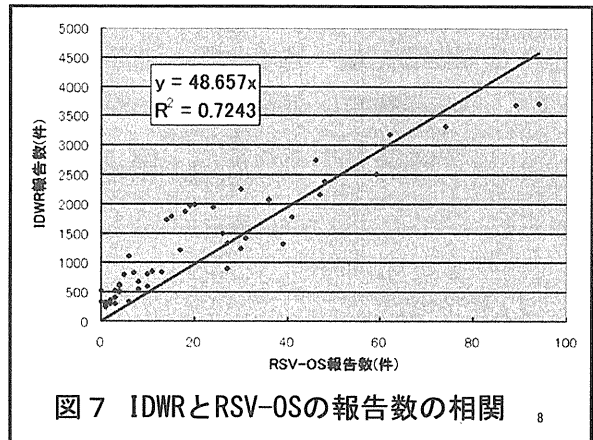
図 5 各都道府県の有志数
(2012/2/17のスクリーンショット)

5





7



8

ツイッターを利用したインフルエンザの流行状況の把握に関する研究

= 総合報告書 =

○分担研究者 西藤成雄 西藤小児科こどもの呼吸器・アレルギークリニック
分担研究者 谷口清洲 国立感染症研究所感染症情報センター
分担研究者 砂川富正 国立感染症研究所感染症情報センター

= 研究要旨 =

インターネットユーザーがいつでも自由に”ツイート(つぶやき)”と呼ばれる書き込みができるインターネット上のサービス”ツイッター”で、インフルエンザに罹患したと推定できるツイートを集計し、その件数の推移が感染症週報との報告数との相関を観た。ツイッターの書き込みから、インフルエンザの流行を探る手法について検討を行う。

A. 研究目的

インターネットユーザー(以下、ユーザー)がいつでも自由に”ツイート(つぶやき)”と呼ばれる書き込みができるインターネット上のサービス”ツイッター”では、しばしばユーザー自身インフルエンザ(flu)に罹患したとのツイートが書き込まれている。

日本語のツイートから、ユーザー自身が flu に罹患したとのツイートを探し、その件数の推移が感染症週報(IDWR)との報告数との相関を探る。また、ツイッターの書き込みから、flu の流行を探る手法について検討を行う。

B. 研究方法

ツイッターが提供している検索条件に合致したツイートやユーザー背景を抽出できるAPI(Application Program Interface:)を組み合わせて、以下のWebサイトを構築した。図1はその先頭ページである。構築したWebサイトを「TweetFlu」と名付け、以下のURLに配置した。

<http://tweetflu.jp>

キーワード「インフル」を含むツイートを抽出し、その中から除外するキーワード(図2)を含んだツイートは、集計から除外した。また「インフル」以外に flu の流行を疑うキーワードとして「学級閉鎖」や flu 治療薬の「タミフル」「リレンザ」「イナ

ビル」も加えた。

構築したWebサイトには、次のような以下の集計ページを準備した。

(1) プロフィールで都道府県が設定されているユーザーのツイートの場合、その都道府県毎に集計を行い、ツイート数に応じて当該都道府県のピンの色を変えて件数を表現(図3)。

(2) 週毎に集計を行い Web サイト上に集計結果を表示(図4)。また flu の流行情報を求めるユーザーに対して、全国版と各都道府県別に、一定時刻に集計結果のツイートを自動的に発信する”ボット”も準備し、Web ユーザーに当 Web サイトの周知を図った。

TweetFlu の集計と IDWR の flu の総報告数の推移を比較検討を行った。また集計に利用されたツイートが、流行状況を把握するために用いられるかどうかを読解し確認した。

5. 倫理的配慮

本調査・研究は、個人を特定する情報は対象とせず、倫理的な問題は発生しない。

C. 結果

1. IDWR と TweetFlu との相関

図4は「インフルエンザ、インフル」を含むツイートの推移と ML インフルエンザ流行前線情報データベース(ML-flu)の報告数推移、そしてその他のキーワードのツイートの推移をグラフに描いた。

a. 2010-2011 年の運用

2010 年第 50 週から 2011 年第 5 週までの、IDWR のインフルエンザの総報告数と TweetFlu の集計の推移の近似式は「 $y = 6.9062x - 61580$ 」となり、その場合の関係係数(R2)は 0.5843 となった(図 5)。

b. 2011-2012 年の運用

2011 年第 49 週から 2012 年第 6 週までの IDWR のインフルエンザの総報告数と TweetFlu の集計の推移の近似式は「 $y = 8.4798x - 159826$ 」となり、その場合の関係係数(R2)は 0.952 となった(図 5)。

2. 抽出精度の検証

a. 2010-2011 年の運用

調査期間を 2011 年 1 月 19 日から 2011 年 1 月 25 日とした。同期間に「インフルエンザ」を含むツイート数は、18089 件であった。そのうち否定語を含むツイート数は 4202 件であった。このツイートを差し引いて残る 13887 件が、ユーザーが flu 罹患した際のツイートの可能性が高い。

2011 年 1 月 22 日深夜 0:00 からの 300 件のツイートを分析した。300 件目のツイートは 2011 年 1 月 22 日 AM2:01 であった。

抽出されたツイートを「ユーザー本人が罹患した」「家族の罹患」「周囲(職場、学校)の罹患」「誰が罹患したのか不明」「flu の罹患ではなかった(否定)」「一般的な flu の話題」「その他」に分類した(表1)。

flu の発生として集計に用いることができるツイートは、本人、家族、周囲、不明を併せて 77 件で、25.7%であった。flu 罹患を明らかに否定しているツイートは 9 件で 3.0%であった。

「その他」のツイートには、ユーザーに対するお見舞いの言葉や、容態を尋ねる、流行に注意するようなツイートが多い。サッカーの選手が flu に罹患すると「インフルエンザ」を含むツイートがふえた。

b. 2011-2012 年の運用

調査期間を IDWR で最も報告が多かった 2012 年 1 月 30 日から 2012 年 2 月 5 日とした。同期間に「インフル」を含むツイート数は、41023 件であった。そのうち否定語を含むツイート数は 4569 件であった。このツイートを差し引いて残る 36454 件が、ユーザーが flu に罹患したか周囲の流行状況に関するツイートの可能性が高い。

2012 年 1 月 30 日深夜 0:00 からの 300 件のツイートを分析した。300 件目のツイートは 2012 年 1 月 30 日 AM0:56 であった。

TweetFlu が収集したツイートを「インフル」「学級閉鎖」「タミフル」「リレンザ」「イナビル」を含むツイートに分け、それぞれの内容を検討した。キーワード毎の割合を図6に示す。「インフル」が最も多く 63%、次いで「学級閉鎖」が 22%であった。

キーワード「インフル」において「ユーザー本人が罹患した」「家族の罹患」「周囲(職場、学校)の罹患」「誰が罹患したのか不明」「flu の罹患ではなかった(否定)」「一般的な flu の話題」「その他」に分類した(表2)。flu の発生として集計に用いることができるツイートは、本人、家族、周囲、不明を併せて 80 件(42.3%)であった。flu 罹患を明らかに否定しているツイートは 10 件(5.3%)であった。

上記外の「インフル」を含むツイートには、特定できない一般的な流行の話題やニュース、ユーザーに対するお見舞いの言葉や、容態を尋ねる、流行に注意するようなツイートであった。

キーワード「学級閉鎖」においては、実際に学級閉鎖が行われた旨のツイートが 22 件(32.8%)、それ以外に学級閉鎖に対する期待感が込められたツイートが 30 件(44.7%)であった。

治療薬のキーワードでは「タミフル」が最も多く 34 件であった。ただ 11 件(32.3%)はニュースに関するツイートで流行を反映する内容ではなかった。

3. ボットのフォロー数

表3に2011-2012年運用時の各ボットのフォロー数を示した。

D. 考察

1. TweetFlu - 発想の経緯、概念

インフルエンザの流行調査には、以下から情報を求める事ができる。

- (1) 医療機関(薬局を含む)
- (2) 学校・保育園
- (3) 罹患した患者から直接

(1)(2)の経路は、すでに研究機関や厚労省が集計し発表している。しかし(1)(2)は、関係者にとっては本来の業務では無く、依頼に応じて行う業務で、報告に常にラグタイムが発生する懸念がある。また集計作業や集計結果の公開に人が介在すると、市民への周知がさらに遅れる。

(3)の経路で、罹患した患者自らが情報発信してもらえば、ラグタイムを小さくする可能性がある。なおかつ人が集計に介在せず、すべて自動化された運用が最も望ましい。

しかし、現状では罹患した患者自らが容易に罹患した旨を情報発信するサービスや運用は存在しない。何らかのサービスの利用状況や発言から、推測していく方法を考える必要がある。

例えば(3)の経路では、大手検索サイト「Google」では、「インフルエンザ」というキーワードを検索する人が多い地区は、インフルエンザが流行していると推論づけ「Google flu trends」としてfluの流行情報を提供するサービスを行っている。

<http://www.google.org/flutrends/jp/#JP>

本調査・研究は、「ツイッター」と呼ばれる、インターネットユーザーが「つぶやき(ツイート)」と呼ばれる短文を投稿し、閲覧できるコミュニケー

ション・サービスに着目した。

Twitter <http://twitter.com/>

2009年6月時点で、日本国内からTwitterにアクセスしているユニークユーザーは約320万人(全世界では約1.1億人)。男女比は、男性が72%、女性が28%。年齢層は、最も多い層が35-44歳の42%、そして45-54歳の18%、25-34歳の17%とされている。

2010年6月の投稿数は国別で日本は世界で2番目にTwitterへの投稿数が多く、米国に次いで全体の18%を占めた。日本のTwitter利用者数は2010年時点で前年の19倍に達しており、実利用者数に限ればすでに米国を追い抜いているとされている。

こうした膨大なユーザーのツイートの中には、fluに罹患した旨の内容が含まれている。その件数の推移は、実際の流行を反映していると考えた。ユーザーの地域情報を取得できれば、流行地域を特定することも可能かもしれない。

Twitterは、様々な利用法や情報提示を可能にするAPIが、多数公開されている。キーワード「インフル」が含まれるツイートのみを抽出しWebサイトで集計表示することも可能である。これらの集計にも人は介在せずに、すべてリアルタイムに処理が可能で、報告からWeb公開までラグタイムが発生しない事も利点である。

以上の発想の元に準備したWebサイト「TweetFlu」の集計結果を検討した。キーワード「インフルエンザ、インフル」が含まれるツイートの件数とIDWRの報告数の推移を比較すると、決定係数が2010-2011年は0.5843、2011-2012年は0.9520であり、後者で高い相関が観られた。その理由として以下の事柄が考えられる。

1. fluの流行が無くても、ユーザーのfluへの関心は高くツイートは“0”にはならないと予想し、

線形近似式の y 切片を“0”とせずで作成した。

2. ツイートは文字数の制限があり、ユーザーは短い単語を使用する傾向がある。そこで抽出するキーワードを「インフルエンザ」の 7 文字から 2011-2012 年運用から「インフル」の 4 文字に変更したことで抽出率が上がった。

しかし、ツイートの件数と IDWR の報告数の推移に、結果として高い相関が観られるが、ツイートの内容を読むと flu 罹患したユーザーは 42.3% 程度であった。一方、キーワード「インフル」以外に、「学級閉鎖」などのキーワードでも flu の流行をよく反映している。また flu の治療薬名も flu 流行を反映する可能性がある。

感染症の流行は様々な情報源より集計を行い、実際の流行を推察するべきである。罹患した患者自身からの情報提供という経路を否定すべきでないとする。最も早く流行を知るという点では、罹患した患者から直接得る情報の収集手段は有効であるとする。

「インフル」を含むツイートから、flu に対する市民の反応が読み取れる。Tweetflu が示すツイート数の推移は、市民や社会が flu で受けた衝撃、影響(インパクト)の強さとも言える。

また抽出するキーワードを「症状」に置き換える事で、まだ臨床症状が明らかでない感染症の流行を早期に察知する症候群サーベイランスとして利用できる可能性もある。

疾病の流行とツイートに関する調査研究は、まだごく僅かしか行われておらず、早稲田情報技術研究所の加藤らの研究が知られている。そして以下の Web サイトを構築し、感染症の早期探知を試みている。

感染症トレンド検知 <http://infection.wit.jp/>

ツイートの研究は未だ確立しておらず、本年

の調査結果は、研究を否定するものではない。今後も様々な調査や研究が行われると考える。

2. 抽出精度を高めるために

次年度の流行では、flu 流行や罹患を示すキーワードをさらに増やし、ツイートの抽出も検討すべきである。そして罹患者のツイートの特異的なキーワードの組み合わせを探る必要がある。

新聞などのメディアのツイートは、ニュースとして自動的に書き出されており、除外ユーザーの設定も必要である。昨シーズンはスポーツ選手が flu に罹患すると、そのツイートが増える現象も確認された。芸能人、政治家、スポーツ選手などが flu に罹患したら、随時、否定語を追加していく必要がある。

3. 罹患時に「インフルエンザ」をつぶやくユーザーを増やすために

ツイッターユーザーが、flu に罹患したときにいつもツイートを必ず書き込むとは限らない。罹患時にツイートを書き込んでもらうためには、インフルエンザの流行を探るプロジェクトが存在していることをユーザーに、広く知ってもらう必要がある。

そのために、能動的な情報発信が一つの方法である。それには TweetFlu の集計結果を一定の時間毎に自動的にツイートするボットを作成した。同ボットをフォローするユーザーはまだ多くないが、有益な情報を発信するボット(ユーザー)には多数のフォローをするユーザーが現れる。本サイトの存在と意義を知ってもらい、ツイートで社会貢献できる事を広く知ってもらえると期待する。

E. 結論

1. Twitter が提供する API を用いて「インフル」という単語をキーワードにツイート数を集計表示する Web サイトを構築した。

2. そのツイート数と IDWR との報告数の推移の比較をおこない、強い相関が観られた。

3. flu 罹患したユーザーに特異的なツイートの抽出方法の検討が必要である。

4. flu に罹患した場合に自主的にツイートを書き込むユーザーを増やす必要がある。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

なし

2. 学会発表

なし

3. メディア

1. 読売新聞 2月8日 夕刊 「インフル情報 つぶやき速報」

2. 日経メディカル・オンライン 「パンデミックに挑む」

H. 知的財産権の出願・登録状況

1. 特許出願

なし

2. 実用新案登録

なし

図1 Webサイト「TweetFlu」

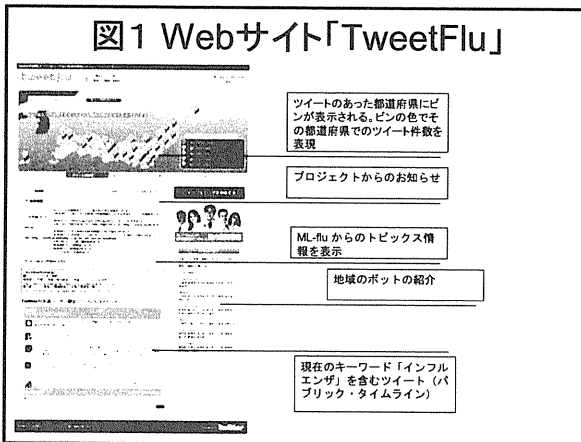


図2 否定語一覧

- 予防
- 注射
- 接種
- 摂取
- 打つ
- 鳥
- 鶏
- ツル
- 対策
- 陰性
- じゃなくて
- だったらどうしよう
- もし

図3 ツィートが書き込まれた都道府県分布

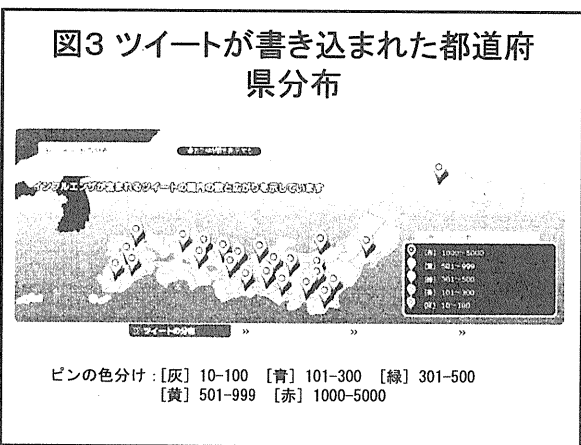


図4 ツィート数の推移(週集計)

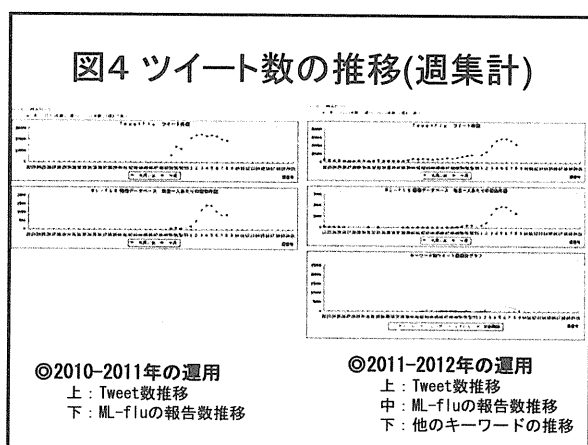


図5 TweetFluとIDWRの相関

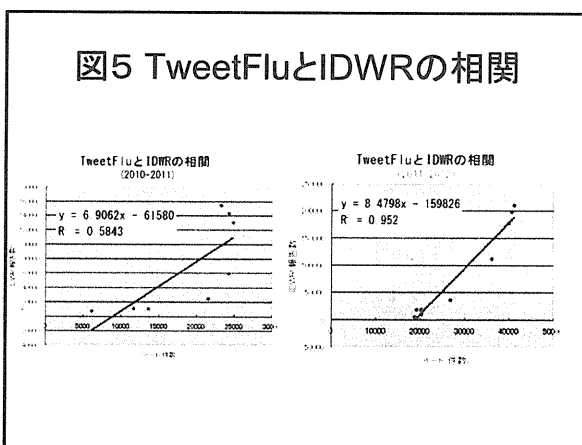


表1 ツィートの分類

分類	件数
ユーザー本人が罹患	40
家族の罹患	9
周囲(職場、学校)の罹患	22
誰が罹患したのか不明	6
=>fluではない(否定)	9
一般的な流行の話題	35
その他	179
*合計	300

図6 抽出されたツイートの割合

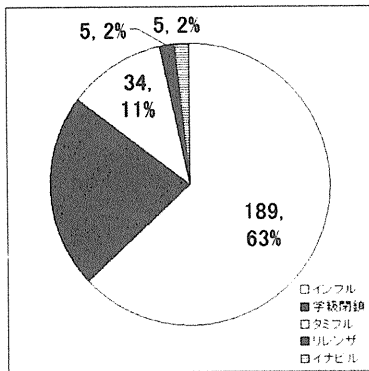


表2 ツイートの分類 (2011-2012)

カテゴリー	総数
インフル	189件
インフル (サブ)	37
インフル (サブ)	8
インフル (サブ)	7
インフル (サブ)	4
インフル (サブ)	23
インフル (サブ)	3
インフル (サブ)	10
インフル (サブ)	7
インフル (サブ)	10
インフル (サブ)	81
学校閉鎖	67件
学校閉鎖 (サブ)	21
学校閉鎖 (サブ)	10
学校閉鎖 (サブ)	3
学校閉鎖 (サブ)	13
タミフル	34件
タミフル (サブ)	2
タミフル (サブ)	2
タミフル (サブ)	11
タミフル (サブ)	19
リンザ	5件
リンザ (サブ)	1
リンザ (サブ)	4
イナビル	5件
イナビル (サブ)	1
イナビル (サブ)	4

表3 各ボットのフォロー数

全国版BOT	10フォロワー
北海道、東北BOT	6フォロワー
関東BOT	7フォロワー
北信越・中部BOT	4フォロワー
関西BOT	4フォロワー
中国・四国BOT	6フォロワー
九州・沖縄BOT	8フォロワー

電子カルテ「ダイナミクス」から ML インフルエンザ流行前線情報
データベースへのインフルエンザ診断症例の自動登録システムの構築
= 総合報告書 =

○分担研究者 西藤成雄 西藤小児科こどもの呼吸器・アレルギークリニック
分担研究者 谷口清洲 国立感染症研究所感染症情報センター
分担研究者 砂川富正 国立感染症研究所感染症情報センター
研究協力者 東金城克治 医療法人西藤こどもクリニック

= 研究要旨 =

目的: 診療所に広く普及している電子カルテ「ダイナミクス」にて、ML インフルエンザ流行前線情報データベース(ML-flu)へのインフルエンザ症例の自動登録システムの構築し、ML-flu へ症例登録している医師の負担軽減。

方法: ダイナミクスの診療記録より言語解析を併用し、インフルエンザ(flu)症例を抽出し、ML-flu へ自動登録システムを構築する。このシステムにより、ML-flu 症例登録に求められる情報を抽出し CSV 形式のデータ・ファイルを作成する。検討対象は西藤小児科こどもの呼吸器・アレルギークリニックと医療法人西藤こどもクリニックにて稼働しているダイナミクス。

結果: ダイナミクスの診療情報から、西藤小児科こどもの呼吸器・アレルギークリニックでは、flu 症例を 2 年続けて 100.0%、医療法人西藤こどもクリニックでは 94.7%と 98.0%の確立で抽出できた。

考案: 診療所で広く普及している電子カルテで、その診療録から flu 症例を高い確率で抽出できた事は意義深い。ML-flu の報告で負担となっている症例登録が自動化されることで、運営の協力する医療機関・医師が増え、より一層の flu 症例の収集が可能となることが期待される。

A. 研究目的

2009年12月現在、2791件以上の診療所で導入されており、診療所向け電子カルテでは第3位のシェアとされている[1]。また、2009年8月の時点で支払基金ページに掲載されたオンライン請求を実施した診療所(病院を除く普及率10.3%)9140件のうち、1082件(11.8%)がダイナミクスであり、診療所で広く普及している[a]。また、そのプログラムソースが公開されており、様々な付加機能をユーザーが開発できる特徴がある。

本研究は、flu患者が多く受診する診療所で広く普及するダイナミクスから、ML インフルエンザ流行前線情報データベース(ML-flu)[2-9]へのインフルエンザ診断状況の自動転送システムの構築し、ML-flu へ症例登録している医師の

負担軽減することである。

B. 研究方法

1. 対象

「西藤小児科こどもの呼吸器・アレルギークリニック」と「医療法人西藤小児科」に導入されているダイナミクスの診療情報を調査する。

2. 調査方法・システム構築

開発したシステムを「ML-flu 用データ出力アプリケーション(ML-flu アプリ)」と称する。図1にその起動時に表示されるウインドを示す。

(1) 基本設定

最初に使用する際に、ML-flu 登録時に求められる情報の内、報告者固有の情報(報告者氏名、メールアドレス、診断キット)、ダイナミクスがインストールされているPCのダイナミクスのデータベースファイルが存在

するディレクトリ・パスを設定する(図2)。一度設定を行えば、次回からはこの操作は不要である。

(2) flu 診断時に記載されるフレーズの設定

flu を診断した際に、ダイナミクスの診療ページに決まって入力されるフレーズを設定する。複数設定することも可能である。そのフレーズは医師によって異なり、医師毎に設定する必要がある(図3)。病名が「インフルエンザ」とされている症例で、設定されたフレーズが所見ページにあれば、ML-flu に登録すべき症例と判断され、CSV 形式ファイルへの抽出対象となる。

(3) flu 診断時に処方される治療薬剤のフレーズ

flu を診断した際に、処方される治療薬剤のフレーズを設定する(図3)。複数の薬剤が設定可能である。

2. 倫理的配慮

本調査・研究は、個人を特定する情報は対象とせず、倫理的な問題は発生しない。

C. 結果

ML-flu アプリにて、両医療機関で診療に使われているダイナミクスで、flu 症例の抽出作業を行った。ML-flu アプリにて、それぞれ ML-flu に登録可能な CSV 形式のファイルとして書き出すことに成功した(図4)。

西藤小児科こどもの呼吸器・アレルギークリニックにおいては、2010-2011 年の運用において 106 例の flu 症例中、106 件を抽出できた(的中率 100.0%)。2011-2012 年の運用では 118 例の flu 症例中、118 件抽出できた(100.0%)。(表1)

医療法人西藤こどもクリニックにおいては、2010-2011 年の運用において 146 例の flu 症例中、149 件と抽出した(的中率 98.0%)。2011-2012 年の運用では 171 例の flu 症例中、162 件抽出できた(94.7%)。(表1) 2010-2011 年の運用同院医師のフレーズを登録行った

が、迅速診断の A/B 判定はほとんどが「判定不能」と抽出されてしまった。また間違えて抽出された症例を個別に検証すると、(1) 周囲の流行、同居者に関する問診の「インフルエンザ」の文字列に反応したケースが 3 件、(2) 迅速診断を行っているのに、未実施と判断ミスが 2 件、(3) 迅速診断が陽性であるにも関わらず、陰性と判断したケースが 1 件、(4) 判定結果の抽出が不完全なケースが 1 例であった。2011-2012 年の運用で抽出から漏れた症例を検討すると、A 型でイナビルの処方されたケースで 7 件、B 型でイナビルが 1 件、A 型で無投薬が 1 件であった。

D. 考察

今回開発した ML-flu アプリは、どちらの医療機関のダイナミクスからも高い確率で flu 症例を抽出が可能であった。

傷病名、迅速診断の有無、flu 治療薬処方の有無、などの複合的な解析に加え、カルテ文中の flu 症例を診察した際の特定のフレーズ解析を実装したことが、抽出率を高めた理由と考えられる。

電子カルテの所見欄は、自由文であるが、その入力には必ず日本語変換ソフト(IM: Input Method)を介して行われている。ほとんどの IM は効率よく文章入力を行うために、利用者の頻度の高い入力フレーズを記憶している。最初の数文字で、利用者が入力しようとするフレーズを予測し、入力したいフレーズを候補として画面上に表示される。その候補から利用者は求めているフレーズを選ぶので、フレーズのバリエーションは、実際にはそれほど多くない。調査対象において、flu 症例で用いられていたフレーズは、調査期間中に西藤小児科こどもの呼吸器アレルギークリニックで 5 種類、医療法人西藤こどもクリニックで 3 種類にとどまった。

解析するフレーズを自由に設定する機能

を実装したことで、定形化されていない自由文から、定形化した情報へ置き換えが可能となった。

しかし、フレーズによる症例の抽出にも問題があった。同居者に関する問診に「インフルエンザ」の文字列が含まれていると、flu 症例と間違っ て抽出された。

抽出された症例を細かく吟味していくと、迅速診断が未実施でも、flu 治療薬が処方されているケースが含まれている。例えば次のケースである。

- (1) 高熱を伴い学校や保育園での流行があり迅速診断するまでもなく flu と診断され flu 治療薬処方
- (2) 翌日が休日であったり通院が困難な場合の flu 治療薬処方

(1) のケースは flu として ML-flu の登録は構わないが、(2) のケースも ML-flu の登録されてしまうのは問題がある。この 2 つケースは、カルテの記載を読めば判断できるが、アプリケーションで区別することは、大変困難である。抽出された症例を登録前に一覧し、登録すべきかどうか主治医が最終的に判断するのが運営上望ましいと考えられた。

膨大な診療情報から flu 症例を高い的中率で一括して抽出できる ML-flu アプリは、ML-flu 報告協力医の負担を大幅に軽減した。ダイナミクスを利用する医療機関に本アプリの存在を紹介し、ML-flu の登録が容易になるよう期待する。

一般にダイナミクスは、セキュリティの配慮により、インターネットに接続されていない院内ネットワークで利用が勧められており、それに従う診療所も多いと考えられる。抽出された CSV 形式のファイルは、一旦、外部の記憶媒体 (USB メモリーなど) に移し替

え、INET に接続された PC から ML-flu ヘファイル転送を行う (図 5)。

効率よく flu 症例を抽出するシステムは実現できたが、ML-flu アプリを普及させるには、ファイルの受け取り作業も煩雑にならないように、高いユーザビリティを備えた運営を心がけなくてはならない。

E. 結論

診療所で広く普及している電子カルテ「ダイナミクス」において、その診療録から flu 症例を高い確率で抽出する ML-flu アプリの開発に成功した。本アプリが普及すれば、ML-flu の報告で負担となっている症例登録が自動化されることで、より多くの協力者を募ることが可能と期待される。

F. 健康危険情報

特になし

G. 研究発表

特になし

H. 知的財産権の出願・登録状況

特になし

I. 参考 Web サイト

[a] 医科診療所向け電子カルテ・レセプトシステム : Dynamics [ダイナミクス]

<http://www.hitachi-softec.jp/dyna/>

J. 参考文献

- [1] 月刊新医療, 2006 年 9 月号.
- [2] 西藤成雄: インフルエンザ情報を知る技 ~ 地域の流行をいち早く入手してインフルエンザ対策を ~ . MEDICAL DIGEST, 55: p53-60, 2006.
- [3] 西藤成雄: メーリングリスト有志によるインフルエンザ流行情報のオンライン集積と公開. 病原微生物検出情報, 27: p16-17, 2006.
- [4] 西藤成雄: メーリングリスト有志によるインフルエンザ流行情報の集積と公開 - ML インフルエンザ流行前線情報データベースの運用について - . 小児

- 科,47:p1928-1934,2006.
- [5] 西藤成雄:リアルタイムでインフルエンザを監視。メーリングリストを使った手作りの新しい試み。The Mainichi Medical Journal(MMJ),2:p164-166,2006.
- [6] 西藤成雄:ML インフルエンザ流行前線情報データベース運用状況とインフルエンザ早期警戒システムの展望。Pharma Medica,24:p104-109,2006.
- [7] 西藤成雄:メーリングリスト有志によるインフルエンザ流行情報の収集と公開。滋賀医学,27:P77-81,2005.
- [8] 西藤成雄:メーリングリスト有志によるインフルエンザ流行情報の集積と公開。東京小児科医会,23:p64-70,2004.
- [9] 西藤成雄:ML インフルエンザ流行前線情報データベースの紹介。日本医師会,136:2439-2443,2008.



図1 ML-flu用データ出力アプリケーション 起動画面

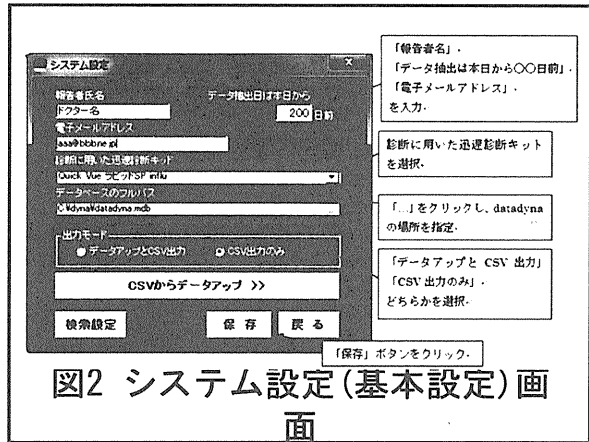


図2 システム設定(基本設定)画面

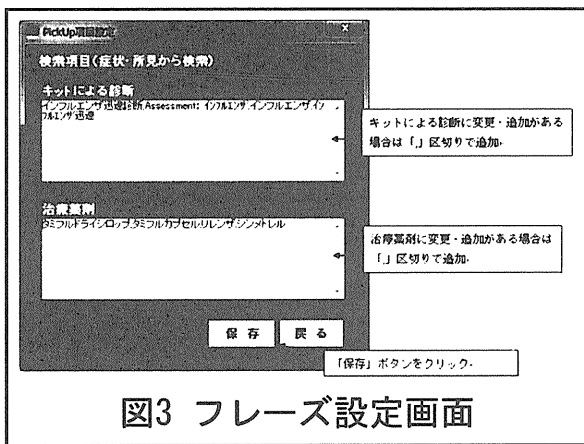


図3 フレーズ設定画面

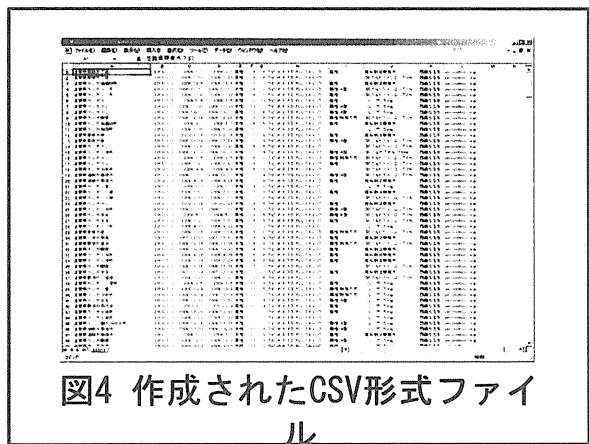


図4 作成されたCSV形式ファイル

表1 的中率

	西条市医師会ともむね検査センター			西条市医師会ともむねクリニック		
	flu症例数	抽出数	正答率	flu症例数	抽出数	正答率
2010-2011	106例	106件	100.0%	146件	149件	98.0%
2011-2012	118例	108件	100.0%	171件	162件	94.7%

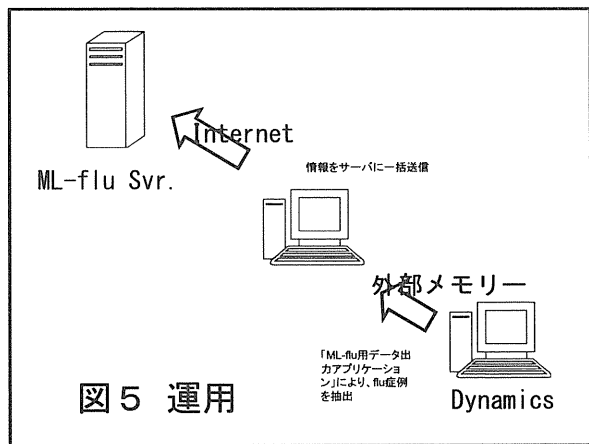


図5 運用

厚生労働科学研究費補助金（新興・再興感染症研究事業）
分担総合研究報告書

緊急時における Fax を用いた感染症情報の効率的な収集・共有

研究分担者 奥村 貴史

（国立保健医療科学院 研究情報支援研究センター 特命上席主任研究官）

研究要旨

緊急時に効率的な感染症対策を行うためには、各種の医療機関や学校、施設等より効率的に情報を収集し、関係諸機関において適切に情報共有を行うための手段の確立が必要である。しかしながら、ウェブシステムは、ユーザーID とパスワードを厳密に管理しなければならず、緊急時に各種の医療機関や施設等から効率的に情報収集を行うような目的で利用することには困難が伴う。そこで、本研究分担では、Fax を活用することにより、緊急時に医療現場側に大きな負担を掛けることなく新たな情報収集を大規模かつ柔軟に行う手法の検討を行っていた。平成 22 年度は、FaxOCR システムのプロトタイプを試験運用時、実利用に向けた課題の抽出を行った。平成 23 年度は、試験運用を通じて得られた問題点を改善し、システムを誰もが簡便かつ低コストに利用可能な形へと整備した。今後、緊急時に実際に役立つ技術としていくうえで、保健医療行政において誰もが日常的に利用できる環境を整え技術への習熟度を高めていく必要がある。

A. 研究目的

効率的な感染症対策を実現するためには、各種の医療機関や施設等から効率的に情報を収集し、関係諸機関において適切に情報共有を行う必要がある。そこで、厚生労働省も、NESID (厚生省感染症サーベイランスシステム)やH-Crisis(健康危機管理支援ライブラリー)を運用するなど、健康危機管理対策を進めて来た。

しかしながら、NESID や H-Crisis などのウェブシステムは、個人情報扱いすることからユーザーID とパスワードを厳密に管理しており、緊急時に各種の組織から効率的に情報収集を行うような利用は想定していない。また、ウェブシステムは、利用者教育が求められるケースが少なく、緊急時に現場に負担を掛けることなく新たに展開することは現実的ではない。

一方で、Fax は、ほとんどのオフィスに

備わっており、誰もが簡便に利用することが出来る。さらに、電話番号通知機能により、発信者の身元を特定することが可能であり、ユーザー管理が容易である。

そこで、本研究分担では、広く普及している Fax と自動文字読み取り装置(OCR)を組み合わせることにより、緊急時に多くの組織から効率的に感染症情報の収集を行い、必要な組織間で効果的な情報共有を行うための技術的な検討を行った。

B. 研究方法

保健医療系の情報システムは、公共調達のルールに縛られるために、事前に仕様書を定め、競争入札により落札業者が決められ、取り交わした契約書に基づいた開発と運用が行われることになる。しかし、契約に基づいて開発したシステムにおいて、実

際に使い初めてから使い勝手の悪さが判明したり、機能が足りないことが明らかとなったりすることは珍しくない。また、新型インフルエンザや東日本大震災においては、事前に想定していなかった事態が数多く生じ、情報システムの多くが機能しない状況が生じた。こうした際も、契約書に縛られる多くのシステムは、刻々と変化していく状況に柔軟に対応することが出来なかった。

そこで、限られた予算で高品質なシステムを開発し、今後も予算の有無に関わらず継続的にシステムの改良を行っていける発展性を確保するために、システムをオープンソース化して公開し、開発者のコミュニティ形成を試みた。その上で、ユーザーインターフェースの改良等を進めた。

また、利用例や利用者マニュアルを充実させると共に、緊急時に役立つシステムとするために、国立保健医療科学院において研修教材として活用し、保健医療関係者への認知向上を図った。

C. 研究結果

FaxOCRシステムのオープンソース化に向けて、開発者コミュニティを構築し、積極的な情報提供を心がけた。そのうえで、他プロジェクトで開発したプログラムの移植を行うなどを経て、システムの品質の要であるユーザーインターフェースを大幅に改良することができた。

現在、さらなる品質向上に向けて、デモサーバを運用継続することにより、稼働実績の蓄積とOCRの精度向上を行っている。また、より分かりやすいマニュアルや教材の開発を進めている。

D. 考察

こうしたシステムを実際の緊急時に役立たせていくためには、平常時から利用する

ことで、操作に習熟しておく必要がある。そのためには、保健医療行政の誰もが無償利用できる環境を整え、日常的な利用を促進することが効果的であろう。

そこで、国立保健医療科学院の情報系研修において利用講習を進めていることに加えて、インストールせずCD-ROMから起動することによりシステムを利用できるFaxOCRシステムの公開準備を進めている。

また、保健医療系の行政機関が自由に利用できる情報システム基盤である「科学院クラウド」上にFaxOCRシステムを構築し、自由に試用できる環境を整えると共に、科学院クラウド上に必要に応じて各自治体専用のFaxOCRシステムを複製構築できる環境の準備を進めている。

集めた情報の還元方法としては、他の感染症サーベイランスシステムと連携することにより地方自治体などとの効率的な情報共有が可能となる。そこで、今後は、NESID等の感染症サーベイランスシステムとの接続手法についても検討を進める必要がある。

E. 結論

本研究分担では、緊急時における感染症情報の効率的な情報収集、情報共有のために、ほとんどのオフィスに備わっており、また、発信者を簡便に確認することができるFaxからの情報収集を効果的に行う手法について検討を行った。

そのために、Faxからの自動文字読み取りを可能とするFaxOCRシステムをオープンソース形式で開発し、ユーザーインターフェースを改良すると共に、マニュアル等の充実を図り、一般公開を行った。

今後、ユーザー教育と、安価な利用手段を提供することにユーザーの習熟度を高め、緊急時における効果的な感染症対策に繋げて行きたい。

F. 研究発表

1. 論文発表

なし

2. 学会発表

なし

3. ソフトウェア

研究成果を、オープンソース形式のソフトウェアとして一般公開している。

G. 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし