

200942029A

厚生労働科学研究費補助金

健康安全・危機管理対策総合研究事業

積極的健康危機情報の収集と分析および

健康危機管理行政への情報提供

のための情報探索機構に関する研究

平成21年度 総括研究報告書

平成22年3月

研究代表者 重松 美加  
(国立感染症研究所)

厚生労働科学研究費補助金

健康安全・危機管理対策総合研究事業

積極的健康危機情報の収集と分析および  
健康危機管理行政への情報提供  
のための情報探索機構に関する研究

平成 21 年度 総括研究報告書

平成 22 年 3 月

研究代表者 重松 美加

(国立感染症研究所)

## 目 次

### 平成21年度総括研究報告書

研究の目的	-----	1
研究方法	-----	2
研究結果	-----	4
考察	-----	14
結論	-----	16
参考資料	-----	17

# 総括研究報告書

## 積極的健康危機情報の収集と分析および健康危機管理行政への 情報提供のための情報探索機構に関する研究

研究代表者 重松 美加 国立感染症研究所感染症情報センター 主任研究官

研究要旨 新しいサーベイランスによる健康危機の早期検知能力の獲得は、改正国際保健規則で加盟各国に課せられた2012年までの達成課題である。日本とアジア諸国には系統的なWeb情報監視のシステムが無く、必要情報入手の遅れが懸念される。リアルタイム検知には週7日24時間対応可能なシステムデザインに加え、最終的な「人＝アナリスト」による「警戒信号」の複合的判断が行われることが不可欠であるが、安定した判断と迅速な初動への連携に必要な理想プロファイルは定まっていない。健康危機時の初動までの時間の短縮、被害拡大の抑制、最短での現状復帰に必要な、質の高い情報の検知と共有の為の情報評価の要件と担い手に必要な素養と訓練に関して研究し、人材要件と養成機構について提言を最終年度の目標とし、初年度は、各国で稼働中のシステムから分析担当に関する情報収集を実施し、作業環境の改善のための検討を画面構成及び文字等のコンテンツデザインから開始した。

### 研究分担者

谷口 清州 国立感染症研究所 室長  
藤井 聡 京都大学大学院 教授  
重松 美加 国立感染症研究所 主任研究官

### A. 研究の目的

改正国際保健規則（IHR2005）は、各加盟国の2012年までの課題のひとつとして、新しいサーベイランスによる健康危機情報の早期検知能力の獲得をあげている。従来の患者発生の報告を集めるサーベイランスに加え、診断前の患者発生情報を集める症候群サーベイランスや、何らかの健康危機

（イベント）が発生したことを検知しようとするイベントサーベイランス（クラスターサーベイランスや緊急搬送情報収集）、健康危機が噂に過ぎないのか判断ができる以前の、予兆のような情報の検知や、近隣国や地域の情報を報告となる前に検知しようという、インターネット等の最先端の情報メディアを駆使した健康危機情報収集の方法を開発し活用することなどで、IHR2005の要請を満たすことが期待されている。

既存の世界的に利用されているシステムで、国あるいは国家レベルの機構が保有し、実用になっているものはわずかに3例である。

カナダの GPHIN、米国の Argus、欧州共同体の MediSys がそれである。システムによっては、公開部分を設けたり、営利目的でなければ契約（高額ではあるが）により情報が共有される仕組みはあるが、取り扱っているものが国益に関与する情報であり、それぞれの国の多省庁が利用していることから、情報セキュリティを担保するためにシステムアクセスは厳しく制限されている。

これまでのところ、日本とアジア諸国には上述のような系統的な Web 情報監視のシステムが国に無く、必要情報の入手の遅れが懸念されている。個別の機構や個人で、公開クローラーを用いて情報を検索したり、民間機関を活用して情報収集をしている。

一方で欧州各国では、国として早期検知のシステムを保有していない場合には、欧州共同体の機構である ECDC（欧州疾病対策予防センター）の情報網や、DC SANCO（欧州共同体 保健・消費者保護総局）の主導による健康危機管理情報の共有 WEB サイトによるリアルタイムの情報共有が図られている。同じ大陸に多くの国境を接した国々が多いことから、口蹄疫、重症呼吸器症候群（SARS）、パンデミック 2009 を通して、一国の影響が周辺各国へ拡大することを実体験し、「相互利益のための共有」を対策として推進している。

新興感染症やバイオテロ、自然発生的または人為的な化学物質汚染などの健康被害の情報は、いつ起こるか予測がつかないため、公式な発表や紙媒体の新聞などの情報ソースからの入手からでは、情報を確認して公開に至るまでに時間的ロスがあり、行政的対応の遅れや被害拡大へとつながる。また、国内の患者サーベイランス（発生動向調査）に報告

される例は、患者の発症、受診、診断を経て初めて疾患名が決まり報告されることになる。こう言った時間的遅延は、初動までの時間を短縮し、被害拡大を防ぎ、最短で現状復帰を実現するという健康危機管理の目標達成にはマイナスである。パンデミックや SARS の様な新興感染症、拡大の速い口蹄疫やウエストナイル熱などに加え、テロリズムなどの際には、1 時間でも、2 時間でも対策導入が早まることで、人の健康被害や国家の社会的あるいは経済的被害が縮小できる。数十分から数時間、場合によると数日の時間を初期に得ることで、より視野の広い、戦略的な、統制のとれた対応が可能となり、各方面に配慮された対応を取ることができ、社会活動の正常復帰までの期間が短縮される。本邦においても、パンデミック 2009 および SARS の際に、国内の体制決定に周辺国、人や物の交通量の多い国々の動向入手に奔走し、情報がバラバラの情報源から、異なる経路で散文的に入ったことから、方針や政策決定の担当者への到達が遅れ、結果的に現実の対策に反映されるのが遅れたり、猫の目の様に変化する対応指示が出たことは記憶に新しい。

Web サイト上の膨大な情報は、日々拡大しており、ラジオなどの音声ソースに加え、有望な情報ソースのひとつであることに疑問は無い。リアルタイム検知には週 7 日 24 時間対応可能な収集システムのデザインに加え、意思決定に資する質の高い情報が必要であり、情報処理の最終的な段階で、「人＝アナリスト」による「警戒信号」の複合的判断が行われることが不可欠である。しかし、国際的にも、安定した判断と迅速な初動への連携に必要な理想的なアナリストのプロファイルは定まっていない。

平成 20 年度特別研究「健康危機情報の積極的収集と分析および健康危機管理行政への情報提供のためのシステム開発と運用に関する研究(研究代表者ナイジェル・コリアー)」のシステムデザインの提案を踏まえ、「人」とコンピュータが調和した労働環境下で、「人＝アナリスト」の安定した判断力と速度の提供を可能とし、健康危機に際して初動時間の短縮し、被害拡大を防ぎ、最短で現状復帰を実現するためのインターネット等の最先端の情報メディアを駆使した質の高い健康危機情報の早期検知と共有を可能とする条件や環境を検討し、情報評価の要件と担い手に必要な素養と訓練に関して研究し、本邦の現状の改善に必要な人材要件と養成機構について提言をすることを目的としている。特に初年度は、健康危機管理に対して、予算を含む国家の支援体制が確立している米国、カナダと、国際協力により同レベルの情報収集を図る欧州など、先進的取り組みを実施している各国の現状と人材プールの特定や人材養成に関する情報についてまとめる。また、研究代表者の所属機関が保有する既存システムの WEB クローラー機能を用いて、研究協力者の技術支援を得て、コンピュータの改良による効果についても一部検討する。

## B. 研究方法

国立感染症研究所 2 名、京都大学 1 名の研究分担者以外に、国立情報学研究所 1 名、カナダ公衆衛生局 2 名、ジョージタウン大学 1 名、欧州疾病対策予防センター 1 名の研究協力者からなる研究班組織で、研究班の目的である最終年度の提言に向けて、全員が協力する体制で研究を実施する。

初年度は、国際システムの実情調査と、現実の報告の公衆衛生への貢献事例、

アナリストや情報分析官に求められているグローバルスタンダードに関する情報などの調査および、情報と事例の収集を行った。各国アナリストとの連携により収集した判断基準、現在の運用方法などを含む情報から、次年度以降に検討すべき課題項目を検討した。また、コンピュータ機能と作業環境の疲労軽減に必要な改良として、作業量とのバランス、画面表現などについて、より作業効率が高くなる様なインターフェースの検討を開始した。

### (倫理面への配慮)

本研究においてはその対象が人では無いため、倫理面の問題は生じ無い。取り扱い情報が国あるいは地域の利害にかかわる場合、個人の特定につながる場合、人心の無用な混乱を招くと考えられた場合には、匿名化するか、公開しないなど、その取り扱いに配慮するとともに、本報告書への記載についても、公開情報を基本とし、協力関係にある機関やシステムの利害に反する場合には、事前にその了承が得られたもののみを掲載する。国の行政上の仕組みに関する情報の取り扱いは慎重に行い、アナリストからの情報提供の依頼は、文書で説明し、同意を確認した上で、匿名にて収集し、総合的意見として整理を行う。個人情報収集しない。情報については、著作権が発生し、その使用に際しての制限が生じるものもあり、その場合は、使用許可を得るか、著作権を侵害しない形で本報告書には概要のみ記載する。以上に照らし、公開が許可されなかった詳細な情報は、我が国の体制を議論する参考資料としてのみ使用することとし、厚生労働省関連部局に別途提供する。

## C. 研究結果

### 1) 本邦の現状と問題点

これまでのところ、日本とアジア諸国には系統的な Web 情報監視のシステムが国に無く、必要情報入手の遅れが懸念されている。個別の機構や個人で、公開クローラーを用いて情報を検索したり、民間機関を活用して情報収集をしている。

SARS、A/H5N1 亜型の鳥インフルエンザウイルスの家禽における流行とヒトへの感染事例、A/H1N1 亜型によるインフルエンザパンデミック、口蹄疫、ウエストナイル熱、リフトバレー熱と、環境破壊などによる新種のウイルスの出現や交通と流通の Global 化による感染症の国際化とか、ボーダーレスな感染症といわれて久しいが、本邦の情報サーベイランスは他の先進国に比べ立ち遅れている。警察や国防関係の情報収集機能が向上したことに比べ、健康危機管理部門での進捗は遅い。

日本には系統的な国際サーベイランスシステムも組織も存在していない。1997年に国立予防衛生研究所から国立感染症研究所に組織再編が行われ、感染症情報センター設置されてより、数人の担当者や研修生である FETP (実地疫学専門家養成コース) が、公式情報や個人的なネットワークを使い各国からの情報を収集し、また WHO の Outbreak Verification List や ProMED などのネットワーク情報を集めたり、独自に世界各国のニュースメディアの英語サイトなどを手動で検索して情報を収集してきた。現在も、こうした活動をする人数が増えたり、他の機関でも同様のことを開始したり、情報収集のためのクローラーが進化した以外は、国立感染症研究所や地方衛生研究所、海外勤務健康管理センター、厚生労働省検疫所が

基本的に同じことを繰り返しているのが実情である。非効率的な重複した努力が払われ、人件費抑制による効率化に逆行している。また、WHO や ProMED などの情報から日本国内向けに和訳して配信するといった方法が、地方衛生研究所ネットワークや検疫所ネットワークで取られてきたが、この方法では、ある程度問題になりつつあるものであれば、探知できるが、まだよくわからない段階での早期探知は難しい。系統的ではなく、言語も日本語と英語に限られるため、網羅的な情報収集ができていない。

日本語が 2 バイト文字であり、アジア言語のほとんどが 2 バイト文字であり、本邦がアジアに位置することもあって、国際的にもアジアを中心とした情報の収集や評価、発信と共有が期待されているが、そのための基本的な仕組みが存在していない。

IHR2005 の改正で加盟国へ要請されているのは、加盟各国が 2012 年夏までにコア・キャパシティとして、原因が何によるものかに関わらず、化学物質や放射線を含むすべての国際的に重要な公衆衛生的危機の早期検知と情報共有のためのサーベイランスと対応の整備をすることである。この取り組みの一環として本研究班が取り上げている情報のサーベイランスがある。研究分担者らは平成 20 年度の特別研究の提言でも指摘したが、日本は先進国として認知されていることでもあり、国家的事業として IHR2005 に即した情報収集能力を向上することが、国際社会の一員として求められている。

### 2) 海外のシステムとその活動状況調査

研究代表者らは、平成 20 年度特別研究「健康危機情報の積極的収集と分析および



健康危機管理行政への情報提供のためのシステム開発と運用に関する研究（研究代表者ナイジェル・コリアー）」において、「WEBベースの国際的健康危機情報システムに関する検討—将来的な日本のシステム設計における方法論的提言—」として、各国システムのコンピュータシステム部分の情報は報告している。従って本報告では、アナリストの候補（公衆衛生対応者、情報処理技術専門家、言語・表現学専門家、人間行動学専門家、医師、国際政治専門家、微生物学者、バイオセキュリティ・対テロ対策専門家等）に必要な資質、知識背景、現時点での職種等について、また、人材雇用や養成方法などに関する情報収集を行った。

#### Global Public Health Intelligence Network (GPHIN)

カナダ公衆衛生局（PHAC）が保有しているシステムであり、傘下の Centre for Emergency Preparedness and Response（CEPR）によって運営されている。12年間におよび運用歴があり、テロリズムの情報を含む公衆衛生学上の健康危機情報の早期検知と警戒アラートの発信のために、1日24時間、週7日間運用されている（<http://www.phac.aspc.gc.ca/gphin/index-eng.php>）。



図1. GPHIN の紹介ページ  
（<http://www.phac.aspc.gc.ca/gphin/index-eng.php>）

本システムは、設立早期から世界保健機関（WHO）とのパートナーシップを確立しており、1998～2001年のWHOによる評価では、WHOが確認したアウトブレイクの56%が本システムで検出されたことを受け、ウェブソースの情報が公衆保健上の脅威の早期警戒として有効であることが確認された。

Webコンテンツ管理およびテキストマイニング技術など、システムの技術的構築には専門家としてカナダのモントリオールのNStein technologiesとの協力がなされたが、アナリストについては、すべてCEPRの職員である。2004年にリリースされた現在のGPHIN IIは、6つの国連公用語（アラビア語、簡体字中国語、英語、フランス語、ロシア語、スペイン語）および繁体字中国語、ポルトガル語、ペルシャ語の情報を分析している（[http://www.phac.aspc.gc.ca/media/nr/rp/2004/2004\\_gphin\\_rmisp-eng.php](http://www.phac.aspc.gc.ca/media/nr/rp/2004/2004_gphin_rmisp-eng.php)）。多言語に対応した公衆衛生オントロジーを開発し、ヒト疾病をはじめとした健康危機管理の用語の意味とその重要度をエンコードして、システムが先ず、報道内容の警戒レベルを検出および評価するようにし、アナリストの負担を軽減している。人間のアナリストの役割は、コンピュータが事前処理した情報を、その処理の妥当性や内容の確認を行い、詳細な調査が必要かどうか、また、警戒を呼び掛けるアラートを公衆衛生対応者や政策意思決定の担当者へ発信するかどうかを判断することにある。

GPHIN アナリストと呼ばれる人々は、平成20年で14名、パンデミック対応を経て平成21年には15名となり、一部局を構成する。これ以外にシステムメンテナンスのためのIT専門家がスタッフとしている。情報分析は、生物学、化学、ジャーナリズム、サーベイランス、経済学、および環境科学

などの様々な専門知識を有する者が担当している。特に、公用語を二言語持つカナダの特徴として、ほとんどの分析者は二言語以上（例えば、WHO 公用語+英語）の複数言語を操ることができ、その技能を考慮に入れた人員のシフト配置を組むことで、同じ情報を二重に、複数の専門的視点から確認することができるように構成している。また、アナリストには、言語の使用地域に合わせて、担当地域の理解や現在の状況に関する経験を背景とした直感も駆使したニュース情報のレベル判定が期待されている。これに加え、副次的なタスクとして、ニュースソースのマニュアルでの追加調査、翻訳システムの出力内容の確認や編集などがある。

1日24時間、週7日間の運用を担保するために、アナリストは基本的にシフト制で就業している。分析者の生活リズムを「太陽のリズム」、つまり通常的生活リズムを維持できる勤務を基本とし、最初のニュースが報道された時に、その対象言語を解するアナリストが待機しているように工夫されている。つまり、中東、中国、米国と異なるタイムゾーンをカバーするため、カナダ時間の夜間および早朝のシフトが含まれるために、複数言語を操る人の間で調整して、生活リズムを破綻させない様なシフトを組むことに努力がはらわれている。情報の判定は、改訂国際保健規則（IHR 2005）の Annex 2 ([http://www.who.int/gb/ghs/pdf/IHR\\_IGWG2\\_ID4-en.pdf](http://www.who.int/gb/ghs/pdf/IHR_IGWG2_ID4-en.pdf)) のアルゴリズムに従い、経対象地域の種々の背景情報や言語やジャーナリズムの知識に加え、経験（蓄積データ）から決定されている。

### Project ARGUS

米国ジョージタウン大学の Argus

Research Operations Center で運用され、2004年以降に公開されている WEB 情報から疾病アウトブレイクの迅速な検出を行っているシステムである。テロリズムの情報を含む公衆衛生学上の健康危機情報の早期検知と警戒アラートの発信のために1日24時間、週7日間運用されており、コンピュータシステムの自動分析と、アナリストによる分析が行われるところは前出の GPHIN と同様である。特徴的なのは、米国内のサーベイランスは、米国疾病対策予防センター（CDC）やその他の米国連邦政府機関が担当しているため、ARGUS は米国以外の情報だけを取り扱っている点である。

本システムも約10年の歴史があり、MITRE Corporation の MiTAP (MITRE Text and Audio Processing) システムを採用し ([http://www.mitre.org/work/tech\\_papers/tech\\_papers\\_04/04\\_0804/index.html](http://www.mitre.org/work/tech_papers/tech_papers_04/04_0804/index.html))、2001年から構築が始まり、2004年に早期警戒のためのアウトブレイク検知システムとして公開された。米国連邦政府の複数の部署で共有し、運用されていることから、人、動物、植物の疾病やテロリズムに対応しており、40言語に対応するなど、収集情報の守備範囲は最も広い。

アナリストの人数も最も多く、スタッフが45人いる。言語および文化的知識は広範囲におよび、コンピュータシステムの保守および開発を担当する10名弱の技術専門家の支援を受け、公衆衛生に関する専門家が管理チームとして協力し、情報分析が進められる。アナリストの活動を支援するシステムとして、公衆衛生用の Argus オントロジーから様々な語彙と重要度のウエートを付けた関連情報が組み込まれており、これはインターネットが普及する前にまで遡って、過去のアウトブレイクの報告を詳細かつ



継続的に検証した調査の結果に基づいている。また、自然災害や気象学で使用される警戒の段階的モデルを採用した生物学的イベントの記述や、直接的（臨床診断、地理、罹患／死亡例数、人口などの疫学指標）あるいは間接的な（自然災害、社会的混乱、公式発表、業務習慣の変更）指標を統合して、情報を選別する機構が初期の分析機能としてコンピュータシステムに組み込まれている。

アナリストの人数が十分いることを背景に、夜間シフトを含む体制はとっている。基本構成としては、カバーしている言語を中心に担当地域チームに分かれて、グループ構成している。まずこのグループ単位で通常とは異なるイベントに関する情報がまとめられ、それを上位のアナリストへ報告として提出し、人による二段階目のスクリーニングがなされ、さらに首席アナリストの確認を経て、国別の警戒アラート情報として公開される。どこにいても対応できる様に、リモートアクセスできる形で、システム内での作業としてこの確認、承認のメカニズムは実施される。機械処理がリアルタイムであることから、タイムラグは非常に小さい。

本システムのアナリストの雇用時には、対象地域に関する専門知識と言語能力に加え、公衆衛生などの特定専門領域に熟達していることが重視される。ある地域で特定の出来事が普通か、普通でないかを理解するには、社会経済的ストレスの早期に見られる間接的なシグナルを検知できるかどうかと密接に関連している。この技量は、直接的な疫学指標とともに、訓練や活発な意見・情報交換を介して磨いて行かれるため、グループの中で切磋琢磨されるために必要な潜在的な能力と基本知識が採用時に求められる。

2008年の同プロジェクトの求人案内の例には (<http://ja-jp.facebook.com/topic.php?uid=7336878017&topic=3683>)、ARGUSの構成メンバーは、医学、言語学、遠隔検出、歴史、情報システム、政治科学、生物学、社会学、人類学、地域学などを専門としており、ラテンアメリカ地域のアナリストは、情報を分析できる程度にスペイン語とポルトガル語が堪能であり、学士（特に社会学あるいは文化研究の分野）の学位を保有し、他種類の情報源からの情報を総合して分析する能力を有し、ラテンアメリカでボランティア、研究活動、雇用などの経験が十分にあり、どう地域の文化に造詣が深く、柔軟性に富む者となっていた。

#### **Medical Information System and Pattern-based Understanding and Learning System (MedISys/PULS)**

イタリアのイスプラにある Institute for the Protection and Security of the Citizen (IPSC) の Joint Research Centre (JRC) によって運用されている MedISys と、ヘルシンキ大学にある言語学的分析を担う PULS の合同システムである。PULS 部分と情報の分析には、ECDC も協力している。予算が欧州委員会から提供されている関係上、欧州各国の共同所有となる。本システムは当初、無休で稼働する全自動公衆衛生サーベイランスシステムとして DG SANCO の Health Threats Unit との協力のもとで作成された。システムの目的は感染症の早期警戒と、CBRN テロと自然災害の検出を公開ニュースメディアで報道される公開情報から行うことにある。MedISys は 2004 年 8 月に起動され、JRC が開発した Europe Media Monitor によって収集されたニュースを自動で処理している。

対応言語は随時拡大しており、欧州各国で

使用している言語をすべて網羅する予定である。現在ウェブポータルから使用できるのは 26 言語（アラビア語、ブルガリア語、中国語、クロアチア語、チェコ語、デンマーク語、オランダ語、英語、エストニア語、フィンランド語、フランス語、ドイツ語、ギリシア語、ハンガリー語、イタリア語、ラトビア語、ポルトガル語、ルーマニア語、ロシア語、スロバキア語、スロベニア語、スペイン語、スウェーデン語、リトアニア語、マルタ語、ポーランド語）であるが、40 言語すべての情報に対応している。

収集メカニズムの Europe Media Monitor と重複して、20 人のスタッフがいますが、MedISys では分析プロセスにアナリストの介在は無い。2007 年以降に、ヘルシンキ大学の PULS が接続され、テキストマイニングにより症例数や地理的位置などから事実を判断する文書内容の言語学的分析が付与された。両システムにより処理された情報は、IPSC によって運用される共通のポータルサイトにおいて報告されている。

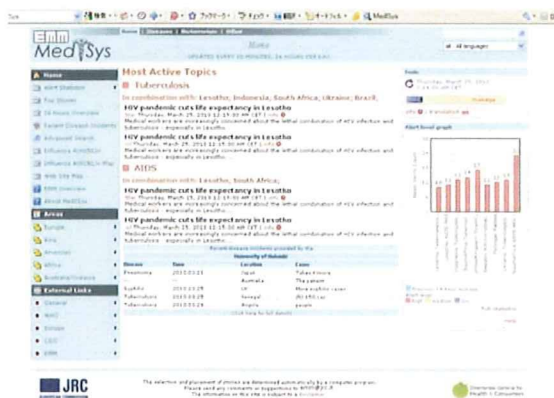


図 2. MedISys のホームページ  
(<http://medusa.jrc.it/medisys/homeedition/all/home.html>)

MedISys のコアチームは、研究開発職員を含む 5 名ほどで構成され、コンピュータ言語

学、知識工学、ソフトウェア工学、公衆衛生の専門知識を有しているに過ぎないが、このシステムは JRC のより大きなメディアモニタリングのグループのメンバーでもあることから、実際に利用可能な人的リソースはさらに広範囲に及んでいる。PULS は、Roman Yangarber 博士が学術的プロジェクトの一環として運用しているもので、コンピュータ言語学、ソフトウェア工学、専門用語処理の専門家からなる約 5 名で構成されている。前二者のシステムと大きく異なるのは、アナリストが含まれていないことである。自然言語処理技術である高度な知識獲得アルゴリズムの公衆衛生および他領域における応用の可能性を検証することが PULS プロジェクトの目的であるため、MedISys と併せて、特定のトピックに関する情報の突発的な増加の自動的検出など、技術革新と改良の面の進展は大きい。最終的な分析を行う人のエレメントであるアナリストについては、情報の使用者に任せている。現在は、アナリスト視点でのフィードバックは Yangarber 博士の学術グループの協力者と、ECDC の担当者 1 名に依存している。

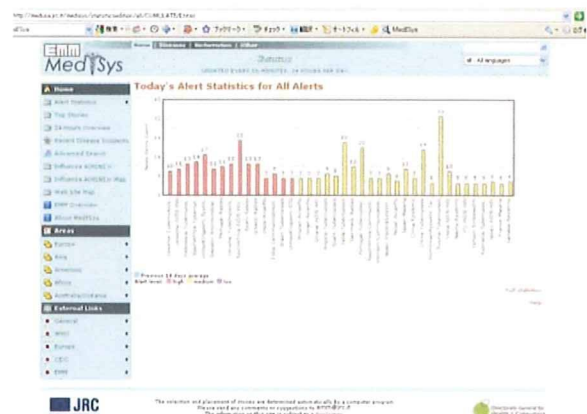


図 3. 自動機能の一例（アラート統計）



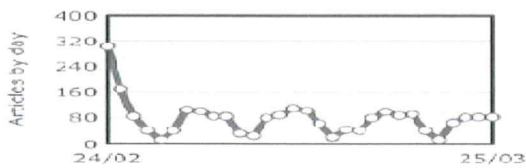


図4. 自動機能の一例 (情報量計測)

### その他のシステム

平成 20 年度特別研究「健康危機情報の積極的収集と分析および健康危機管理行政への情報提供のためのシステム開発と運用に関する研究」の研究代表者であった国立情報学研究所の Nigel Collier 博士が運用する BioCaster も PULS 同様に、日本学術振興会 (JSPS) や新領域融合研究センターの育成融合プロジェクトからの資金による学術的プロジェクトの一環として運用されている。年中無休で、ほぼリアルタイムの収集を可能としているが、スタッフ数は 5 名で、コンピュータ言語学とソフトウェア工学を専門としている。本システムも機械による情報の選別がなされており、アナリストの関与は無い。2006 年から、アナリストの機能を補うため学術的パートナーのネットワークを作り、公衆衛生や遺伝学などの知識を導入している。Collier 博士\*は本研究班でも研究協力者として、技術協力をしている。同システムは現在 3 言語 (英語、日本語、ベトナム語) に対応しているが、8 言語 (中国語、英語、フランス語、スペイン語、日本語、韓国語、タイ語、ベトナム語) による多言語オントロジーが既に用意されており、分析情報の対象拡大が予定されている。

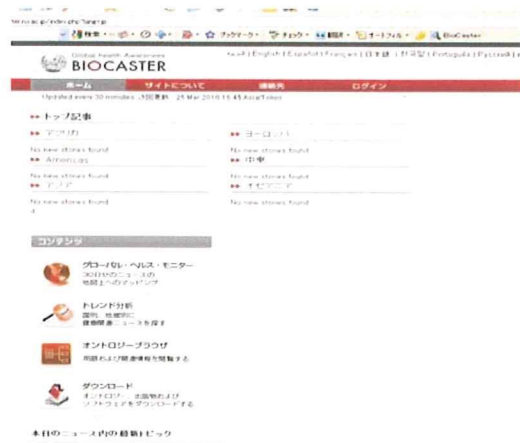


図5. BioCaserer 日本語ホームページ (<http://biocaster.nii.ac.jp/index.php?lang=jp>)

ハーバード大学と民間企業の協力による HealthMap プロジェクトは、無料公開の自動サーベイランスツールである。影響の大きなニュースに焦点を絞り、既存の公衆衛生サーベイランスを補完するという考えで、国際感染症学会が支援する、専門家ボランティアによる ProMED-mail のネットワークと協力し、オープンプラットホームのアプリケーションを活用して、ユーザーが利用しやすい視覚的情報として提供している。Google 社が設立した慈善団体である Google.org からの資金援助と協力があり、Google マップの利用、Google ニュースからの情報の掲載も行っている。英語、中国語、フランス語、ロシア語、スペイン語の主要な 5 言語による提供をしており、ARGUS の国際情報を補完する米国内の情報も掲載している。情報源そのものが既に人のスクリーニングを受けた情報が多いこともあり、アナリストの関与はわずかで、マップ上に公示後、手動で 1 名のアナリストと、自動電子メールの外部購読者の協力を得た人による情報の分析や確認がなされている。



図 6. HealthMap ホームページ  
(<http://www.healthmap.org/en>)

システム以外で、メーリングリストとして最も古くから活躍しているのが ProMed mail である。ボランティアの研究者、医師、国際保健の専門家、公衆衛生関係者などからなる 30 名ほどの編集担当者が情報の分析をして、アナリストの役割を果たしている。しかし、業務外の活動である性質から、情報の発信の遅れは防げず、対応言語も 4 言語にとどまる。アナリストは全員、それぞれの専門領域や個人のネットワーク、さらに支援学術団体である国際感染症学会のネットワークを活用し、経験に基づいた分析を行っている。近年、それぞれの地域ごとのネットワークが娘ネットワークとして立ち上がり、より地域に密着した情報の交換が始まっている。



図 7. ProMed mail のホームページ  
(<http://www.promedmail.org/pls/apex/f?p=2400:1000>)

### 3) システムによる検知の事例

これらの健康危機情報収集システムの効果について疑問視されることが多い。システムの運用者も、十分な人材が投入され、経験の蓄積や訓練が行われ、活発な情報共有と協力体制が確立できるような公衆衛生のネットワークが現実に機能したなら、これらシステムより早期に検知し、対応に繋がることは否定していない。しかし、予算不足、人材不足、人材登用と教育訓練の仕組みの欠如、政治・社会的理由での情報公開の遅れ、経済的理由による圧力など、各国それぞれの理由で正確な情報の公開が遅れたり、共有できない事情が存在する。もちろん、情報の価値が理解されずに報告が遅れることも多々ある。典型的な事例が SARS であった。図 8 には、GPHIN が公開したシステム検知と公的機関の対応の時系列の流れである。中国語情報 (2002 年 11 月 27 日) での検知がいかにか速かったか、また、その情報の共有が十分でなく、分析が遅れた結果、世界的な流行への初動対応が遅れたことが分かる。WHO の情報公開は 2003 年 2 月 25 日である。このとき、国内の新聞は特派員報告として、2002 年 11 月にすでに原因不明の肺炎発生の小さな記事を報道している。また、分析がいかにか重要かを示す例として、この中国語情報は、抗ウイルス剤の売り上げによる、広東省の薬剤会社の経常利益の急増を報道する経済ニュースであり、感染症の集団発生を報じたものでは無かったことが指摘できる。これから、新興感染症を疑うには、経験と学術的な背景が必要である。



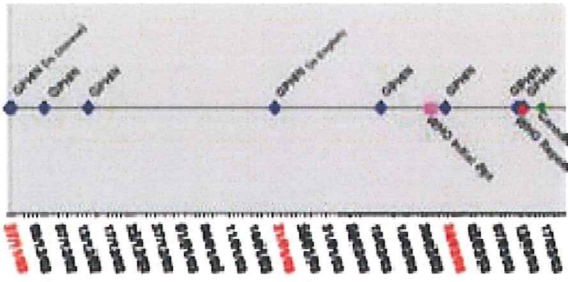


図8. 時系列(左から右へ)順に示したSARS情報の公開の流れ (<http://www.mt-archive.info/MTS-2005-Mawudeku.pdf> Figure1 を再掲)

A/H5N1 亜型の鳥インフルエンザウイルスのヒトへの感染に際しては、多くの情報が現地にいる研究者を介して入ってきた。しかし、一部の国に関しては、国の経済政策に大きな影響を与えるため、情報が公開されないままのところも多く、噂を取り上げたニュース情報などから、情報の検証をせざる負えないことも多かった。

成功事例もあるが、失敗の事例もある。特に、北朝鮮からの情報は検証も難しく、同じ年に毎年の様に同じ話がニュースとして検知される。かなり伝染性の高い疾患が疑われる内容であるが、周辺諸国への拡大も無く、他の国からの報告も無い。このような事例は、ニュースやブログの様な発信が自由な媒体としてはコントロールすることはでき無い。最近の事例としては、赤十字の職員から通報のあった4例のウガンダの天然痘疑いの水痘の事例がある。これも、電話などの個人的な連絡網による情報のため、最初の情報をシステムで検知することはできなかった。また、翻訳の間違いによる情報の交錯もある。ニュースなどの情報は、翻訳を介して複数言語でインターネット上に掲載される。翻訳の技量により、全く違ったニュースとして伝わることも往々にある。テロリズム関連の情報など、用語が特異な

時に見られる。アナリストのスクリーニングは、こういった間違いを取り除くことができる。

#### 4) アナリスト支援の改良の検討

初年度は、作業効率の面からの改良と、情報量の整理を行った。

具体的には、アナリストの作業画面の改良を検討し、既存のWEBクローラー機能を保有するシステムに搭載して、入力や評価作業の効率を検討した。改良にあたり、既存のシステムが担当者の眼性疲労、肩こり、頭痛などを誘発している原因を検討した。操作画面の視認性と操作性の悪さが最も大きく寄与していると推測された。この点を解決するため、色彩、配置に変更を加えただけでなく、最新の民間企業のWEBデザインなどを参考に、図9から11に代表される変更を加えてみた。運用中のシステムを材料に検討しているため、定量的検討は完了していない。主観的には、目の疲れが軽減されている、マウス操作が楽になった等があり、再行継続可能時間が約25%伸びている。



図9. 大きなボタンと感覚的に選択できる表示  
表示を読むストレスを低減し、小さなポインター移動を楽に行えるようにポイント先を大きくした。当り前のデザインではあるが、これまでの公的システムでは配慮されてこなかった。

図 1 0. 行間や情報の区別をし易いフレームをつけ視認性を向上させた情報の評価用一覧表



図 1 1. 評価用のオリジナル画面の表示例

情報を評価するときには、関連情報の呼び出しなど、複数の画面を同時に開くことが多い。モニターを複数台設置する余裕がある場合には、視野に入る最大数（横へ3台、上下2列程度）まで使用することができるが、現在の本邦の施設の備品整備状況では、1台で最大の効果をあげる工夫が必要である。

インターネット上の情報が増加の一途をたどるため、情報量は今後も減少する可能性は無い。従って、目的に合致した情報のみを、人が評価、検討するようにできることが目標である。このために、テキストマイニングと機械学習の機能を付加したフィルターを利用した改良を試みた。このフィルターは研究協力者 Collier 博士が BioCaster のために製作したものであるが（図 1 2）、連結プログラムを作成し、一部の機能を改修する

ことで他のシステムでも機能することを確認したのち、上記システムへ搭載した。これにより、システムが入手した新規情報 100~200 件/日が、感染症のそれぞれのカテゴリへ振り分けられる前の段階で、毎日数件から数十件へと減少した。1 ヶ月検証で、特異度は 95%以上あるが、感度は今後入手情報を増やして確認の予定である。陽性予測力は 98%以上あり、疑陰性として間違っ



図 1 2. フィルター機能の情報の流れの中で  
の位置を示す BioCaster の構造図  
(<http://biocaster.nii.ac.jp/index.php?page=about&lang=jp>)

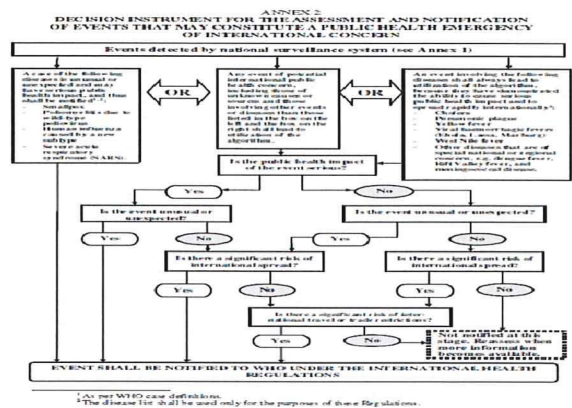


図 1 3. WHO International Health Regulations (2005) second edition 43p より



今回搭載したフィルターは、原因は問わず、感染症の集団発生をターゲットに調整されている。その機能は、機械学習用の教材を何にするかで、基本のアルゴリズムをベースに訓練し、情報の蓄積を基に単語の出現頻度のパターン認識に加え、言語学技術で単語と単語の関係性などを用いて文脈・文意の認識を向上させ、情報の選別能力を改善させるものである。基本理論は単純ではあるが、現実の人間の判断を突き詰めていくつかのステップとする Yes、No の質問を作ることから始まる。例えば、IHR2005 を念頭に置くと、「主題は国際的な公衆衛生学上の健康危機につながる情報ですか？」という質問がある。これを、「主題は感染症の患者ですか？」Yes/No、「人から人へ感染する」Yes/No、「致死率が高い」Yes/No、といった様に、細かく分けて文章内にある単語とその出現パターンや、単語の関係性で判断を積み重ねて行き、採用する情報か、テーマと関係無しとして却下するかという機械の判断がなされる。フィルターは、この様な仕組みを段階を追って実施するため、システムの情報を使う者が、「明らかに必要とする情報」と「明らかに必要でない情報」、さらに、「判断に苦しむ情報が必要・不要」と三種類に大別した教育教材とする情報をたくさん用意して訓練を重ねることで、データが蓄積されて、判断の質が高くなることが期待できる。過去の情報のマニュアルサンプリング以外に、今後の情報を収集して行くことが必要である。

##### 5) 国際的なサーベイランスの取組み

G7 + メキシコの保健大臣級会議において、Global Health Security Initiative の推進が合意されたのは 9.11 以降のことである。この下部会議としての閣僚級会合で、2001

年に CBRN のテロ対策の相互協力と支援をすることを目的に、いくつかのワーキンググループ (WG) を作り実行に移すことが決まった (GHSAG: Global Health Security Action Group)。これを受けて、リスク管理とリスクコミュニケーションの WG (Risk Management and Communication Working Group) が立ちあげられ、2007 年 9 月に京都で第一回の会議が開かれた。会議の席上、具体的なプロジェクトを実施し、協力体制を顕在化させることが決まり、早期検知とアラート発信のための情報共有のプラットフォームの作成を行うことで合意した。

この活動は欧州が強く推進して進められているが、国際的にも科学的、経済的、政治的理由から WHO への通報も情報提供も遅れがちになり、公開あるいは情報共有がさらに遅れている実態を踏まえて、先進 7 カ国とメキシコでの情報の早期共有による対テロ対策への連携対応の効果を狙っている。プロジェクトの実態は、本報告書に記載されているシステムによる早期検知の長所を生かして、ひとつのプラットフォームで同時に共有し、各国が対テロ対応において意思決定し、対応を開始するまでの初動時間を 1 時間でも 2 時間でも短縮し、被害を小さくすることにある。従って、情報の収集、統合、情報が内包するリスクの評価、評価結果の共有と必要に応じたアラートの発信、情報の確認結果の迅速な共有の場の提供などが関連しており、積極的健康危機情報の収集と分析および健康危機管理行政への情報提供そのものが主題である。

続く、イタリア イスプラ、ルクセンブルグ、イタリア ローマ、米国アトランタの会議を経て、実際に患者が発生し、診断

される前の情報を集める症候群サーベイランスの進化型として、ニュースや SMS、ブログ書き込みなどのウェブ情報源をデータとする情報サーベイランス（別名ルーモア・サーベイランス：RuS）の情報共有を行うプロジェクトを遂行することで合意した。RuS には三つのステージがあり、機械的な情報の収集と整理、ノイズや重複の除去、次が情報の分析により「疑わしい」、「将来的に問題へ発展しそうな」情報を抽出する段階、最後がリスクマネジメントの担当者や政策対応者が関与する段階で、整理され、評価された情報がどのように使われたか、また対策や対応へのインパクトがあることの確認である。2009 年末のアトランタでは、第二段階のリスク評価の方法がほぼ決まり、現在システム評価とプロジェクト評価方法についての検討が開始されている。我が国も、この一員として参画しており、独自のシステムを保有していないことで、一方的なユーザーとなることから、情報評価などにおいて積極的関与が期待されている。しかし、情報の検証には系統的な情報収集の機構が必要であり、現在、その機構が国内に存在していないことが情報の検証に問題となるという悪循環にある。参考資料として、2008 年 2 月の会議の要旨で、EpiSouth という南欧州のサーベイランスサイトで公開共有されているものを添付する（[http://www.episouth.org/doc/r\\_documents/GHSAG\\_short\\_report\\_for\\_PHEA\\_web.pdf](http://www.episouth.org/doc/r_documents/GHSAG_short_report_for_PHEA_web.pdf)）。

この WG の活動の成果の一部は、システム保有メンバーの連名で、“The Landscape of International Event-based Biosurveillance” のタイトルで、Emerging Health Threats Journal に、2010 年に掲載された（<http://www.eht-forum.org/ehjt/journal/v3/pdf/ehjt10003a.pdf>）。

## 6) 人材について

既存のシステムから必要とされる人材に関する情報を集約すると、目指すシステムが完全自動化されたコンピュータ主体のシステムで、最終的な情報の分析と判断はユーザ・クライアントに任せるか、提供する情報に人によるリスク評価と加工という付加価値をつけるのかで異なっている。前者は基本的にコンピュータの専門家、コンピュータ言語学やシステム工学といった専門領域の者を必要とし、後者は医学、言語学、歴史、情報システム、政治科学、生物学、社会学、人類学、地域学、ジャーナリズム、環境学など広範な背景の者を必要としている。さらに、後者は発信された情報の文化的背景への理解や、複数言語を解する能力を必要としていた。

いずれの場合も、on-the-job training の傾向があり、実際に指導を受けつつ実務を行う中での技能の習得が主体である。したがって、しっかりとした階層型の管理の仕組みを必要としている。

## D. 考察

諸外国の主なシステムにおけるアナリストの位置づけに関する調査と検討は概ね終了した。国内の現状を考慮すると、現場の人材は健康危機の発生時に限らず、常に現実への対応に追われており、改めて部署を作り人員を配置することが可能としない限り、ユーザ・クライアントに情報の分析と判断を任せるのは現実的では無い。また、これらの情報の中には社会、政治、経済の視点から、非常に取り扱いが難しいものも含まれる。本邦の文化風土の元で、完全に公開された、透明性 100%の仕組みで、個人が情報のリスク評価を行うことができる

であろうか。メディアの収集した意見や、世論といわれるものでは、情報の公開は求めるものの、専門的知識を自ら学ぶのではなく、専門家がコメントをすることを求めている。その意見を並べて、選択をするという構図が見える。公衆衛生の現場でも、その守備範囲の広さから、少ない担当者が様々な知識をすべて保有することが難しく、一旦感染症が発生すると特定の WEB サイトのアクセスが急増する、専門機関への電話が殺到するということが起こる。SARS やパンデミックで繰り返し実証されている。従って、なんら評価のされていない情報を如何に早期検知して提供しても、利用できない可能性の方が大きい。

では、透明性 100%の公開された仕組みの下で、リスク評価を実施できるのかという疑問が生じる。国あるいは地方政府が公衆衛生対応を取るための情報であるから、分析はその視点に立つて行う必要がある。お菓子会社や IT 産業が海外工場を増やすかどうかの際してのリスクと、健康危機に至るか否かのリスクは全く評価に必要な知識が異なる。情報の信頼性の面から、国家が利用する情報は、国が専門家と認める機関あるいは人々により評価される必要がある。また、国家危機の視点で評価されたリスク評価結果は、現実の問題として情報の与えるインパクトへの配慮すること無しに公開することは、無用な社会不安や混乱を招く。情報の垂れ流しは無く、健康危機を最小限に留め、乗り切るためにより良いと考えるコミュニケーションの方法の選択が必要となる。GPHIN も ARGUS も、情報そのもののリスク評価をしたものは、政府関係機関に限りリアルタイムに発信するが、学術関係者その他へはタイムラグを設定して、遅れて提供している

(<http://www.phac-aspc.gc.ca/gphin/index-eng.php>)。また、これらの大規模なシステムのほとんどが、政府の公衆衛生と国防のために巨額を投じて制作したという背景もある。

我が国はこう言ったシステムを持たない。自然災害の対応などで示された様に、能動的に情報を収集し、発信して提供し、リアルタイムに他国からの情報も共有することによる本邦の利益は計り知れない。すでに研究分担者らは平成 20 年度の特別研究でも同様のシステムを国家プロジェクトとして、安定した運用を担保して保有することの重要性を提言した。各国のサーベイランスは国家事業であり、公的資金とイニシアチブが無いと、継続した安定運用ができないからである。研究班がシステムを作成したとしても、研究班終了後に運用する資金は残らない。各機関は、近年の予算削減のために余剰体力は無く、新規事業を引き受けることはまず不可能である。友好国からの情報提供を期待しても、提供国がその国内の提供のためのステップを踏むことによる提供の遅れは取り除くことができないため、危機対応の初動に必要な迅速性が犠牲になる。迅速な公衆衛生対応に重きを置く場合には、非公式情報を積極的に検出することが重要なのであるから、積極的に国際的な感染症情報を収集するメカニズムを自国で保有できないなら、欧州の様に、共同で開発するあるいは、他国の開発したシステムからリアルタイムで情報を入手する約束（契約）をすることが選択肢となる。当然ながら、対価が必要となる。いずれにしても国家主導の事業として実施されない限り、継続性は担保されない。

こうしたことを踏まえて、国家プロジェクトとして提言したシステムでの

アナリストについて考察すると、海外のシステムからいくつかの特徴が明らかになる。まず、最大のものは言語能力である。情報は翻訳により変わるため、元の言語での理解と分析が必要である。次に、複数の専門領域からの分析であり、ひとりの場合も複数の協力による場合も考えられる。アナリストや情報分析官は、一定の現場経験や、国際経験を有していることも必要と考えられる。

仮に、本邦がこのまま独自のシステムを保有しないなら、GHSAGの枠組みなどで提供される情報に頼ることになるが、これは事前に特定の前提でリスク評価した情報をそのまま受け入れることになる。「日本にとって」という視点無しに情報を受け入れると、他国と同じリスクを必ずしも日本は負っているとは限らないので、その情報に基づく対策は必ずしも我が国に適していることにはならない。同じグローバルスタンダードの手法でのリスク評価を、日本のためにする必要は残ってしまう。従って、アナリストのグループを目的に合わせて養成する必要も結果的に残る。

次年度は、国内で実際に情報の分析やリスク評価を行っている各分野の現状の調査を行い、今年度の成果と併せて検討して行く。また、国内外の研修プログラムや、on-the-job trainingの詳細を明らかにし、本邦でどのようなことができるかを考察する。また、システム側の改良によるアナリストの負担軽減について、GHSAGの場での情報収集、国際協力、情報サーベイランスにおけるリスク評価手法の設計も引き続き行い、最終年度にアナリストの人材要件と養成機構など健康危機情報の収集に係る枠組みのアナリストフェーズの構成について提言する。

## E. 結論

20世紀末から今世紀にかけて、新興再興感染症やパンデミックが出現した他、テロリズムの脅威も去っていない。これに対応するために、健康危機情報の早期検知が国際的にも取り組まれている。危機対応においては、正確なリスク評価に基づいて、最終的に「人」による「警戒信号」の複合的判断が必要な事は明白であることを踏まえ、本年度は、危機情報のリスク評価の要であるアナリストに焦点をあて、既存システムの情報収集と検討、効果的な活動環境についてのシステム改良の検討、最新の国際現状の情報収集を行った。次年度以降は、国内環境の確認と、リスク評価手順の検討に加え、最終年度の提言に向けてのまとめを行う。

## F. 健康危険情報

なし。

## G. 研究発表

1. Collier, N., Doan, S., Matsuda Goodwin, R., McCray, J., Conway, J., Shigematsu, M., Kawazoe, A. Navigating the Information Storm: Web-based Global Health Surveillance in BioCaster. In: *BioSurveillance: A Health Protection Priority*. Kass-Hout, T. and Zhang, X. (eds), 2010 (in press).
2. Hartley, D., Nelson, N., Walters, R., Arthur, R., Yangarber, R., Madoff, L., Linge, J., Abla, M., Collier, N\*, Brownstein, J., Thinus, G., Lightfoot, N. The Landscape of International Event-based Biosurveillance. *Emerging Health Threats Journal* 3:e3, 2010 (doi: 10.3134/ehjt.10.003)

## H. 知的財産権の出願・登録状況

(予定を含み)

なし