

A Roadmap to Integrated Digital Public Health Surveillance: the Vision and the Challenges

Patty Kostkova
Department of Computer Science
University College London
Gower Street
London, UK
P.Kostkova@ucl.ac.uk

ABSTRACT

The exponentially increasing stream of real time big data produced by Web 2.0 Internet and mobile networks created radically new interdisciplinary challenges for public health and computer science. Traditional public health disease surveillance systems have to utilize the potential created by new situation-aware realtime signals from social media, mobile/sensor networks and citizens' participatory surveillance systems providing invaluable free realtime event-based signals for epidemic intelligence. However, rather than improving existing isolated systems, an integrated solution bringing together existing epidemic intelligence systems scanning news media (e.g., GPHIN, MedISys) with real-time social media intelligence (e.g., Twitter, participatory systems) is required to substantially improve and automate early warning, outbreak detection and preparedness operations. However, automatic monitoring and novel verification methods for these multichannel event-based real time signals has to be integrated with traditional case-based surveillance systems from microbiological laboratories and clinical reporting. Finally, the system needs effectively support coordination of epidemiological teams, risk communication with citizens and implementation of prevention measures.

However, from computational perspective, signal detection, analysis and verification of very high noise realtime big data provide a number of interdisciplinary challenges for computer science. Novel approaches integrating current systems into a *digital public health dashboard* can enhance signal verification methods and automate the processes assisting public health experts in providing better informed and more timely response. In this paper, we describe the roadmap to such a system, components of an integrated public health surveillance services and computing challenges to be resolved to create an integrated real world solution.

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.
WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

Categories and Subject Descriptors

H3.4. Social Networking **Experimentation**

General Terms

Public Health Informatics, Data Mining, Web Science, Context-awareness, Social media

Keywords

Epidemic Intelligence, Outbreak Detection, Risk Communication

1. INTRODUCTION

Communicable diseases remain a major health threat for citizens worldwide. New diseases such as SARS are emerging and increasingly frequent travel contributes to the need to detect and respond to public health threats in more timely and coordinated fashion. Complementing traditional case-based microbiological laboratory reports and syndromic surveillance, event-based surveillance is monitoring unstructured events, such as news and Internet media, and has been a significant component of public health early warning and response over the last decade (e.g., GPHIN, MedISys).

However, with the increase of Web 2.0 platforms and social media there is a new real-time source of readily available intelligence provided by citizens directly into public domain. Enhanced by geographic and spatiotemporal tags, population movements can be monitored in virtually real time. In particular Twitter, allowing posted information to be shared in public domain and available in a machine readable format over APIs, is the leading platform in this revolution providing an invaluable source of real time data about more than 500 millions citizens worldwide.

Further, digital epidemiology [1] is opening new horizons through mining information from sensor networks, store loyalty cards (in particular, recording information about medicines purchase), travel cards and mobile phones. This provides yet another realtime stream of situation-aware geo-located information aiding early warning and population monitoring.

Finally, the popularity of participatory systems is rising sharply, further complementing these data streams with self-reported information about citizens' symptoms and diseases (systems such as 'Quantify Me', Influenzanet, etc).

There have been a number of new projects utilizing the Internet, online media and user-generated Web 2.0 content and demonstrating their potential for public health. However, there has not been an attempt to conceptualize the process, develop robust validation and verification methods (utilizing standard disease epidemiological models and new set of user online behaviour models) and *integrating* these systems into a full public health service reliably supporting professionals by semi-automating verification and communication processes.

In this paper, we discuss the background of the new generation public health informatics, surveillance and epidemic intelligence systems (section 2). In section three, we present the data sources for digital case-based and event-based surveillance focusing on novel streams from mobile, sensor and participatory systems. In section 4, we present the vision for an integrated framework for a digital public health dashboard. Section 5 brings the further challenges for outbreak management: outbreak investigation and control, and risk communication. While most of this paper focuses on the computer science challenges, section 6 presents a discussion about the wider implications of the framework - the legal context and international dimensions, while section 7 concludes.

2. BACKGROUND

Public health protects individuals, communities and populations from infectious disease at local, national and international levels. With increased global travel and emergence of new diseases (such as SARS), a more robust early warning systems and coordination support became an imperative.

The case-based and syndromic surveillance traditionally relies on national surveillance systems with established bottom-up reporting processes from local through regional to national and international levels. Agreed case definitions and legally enforceable set of notifiable diseases (EC decision 2000/57/EC and 2008/351/EC and WHO International Health Regulations, legal frameworks are discussed at the end of this paper) make this pillar of surveillance, at least in developed world, very reliable and provide an invaluable source of confirmed longitudinal epidemiological data of diseases prevalence held by national public health agencies and at the international level, by WHO and at EU by ECDC (such EU database is called the European Surveillance System TESSy¹). However, the process of reporting, collating and analyzing surveillance data takes normally several weeks.

The second pillar of surveillance, complementary to traditional syndromic and case-based data signals is the event-based surveillance with the *epidemic intelligence* (EI) at the cornerstone of the activity. EI is defined by ECDC as “the process to detect, validate, analyze, monitor and communicate on public health events that may represent a threat to public health” [3]. The following Figure 1 illustrates the process required to detect, verify and control an outbreak.

Epidemic intelligence systems, customized to the needs and priorities of each country, are typically using one or more early warning systems scanning media news and email warning systems (BioCaster, Argus, GPIHN, HealthMap, MediSys,

ProMED-mail, Puls) as identified by a comparative study on a detection of A/H5N1 Influenza Events [2] highlighting the need for “*more efficient synergies and cross-fertilization of knowledge and information*”. The systems, as required for the purpose of this roadmap article, will be discussed greater detail and referenced in the following section.

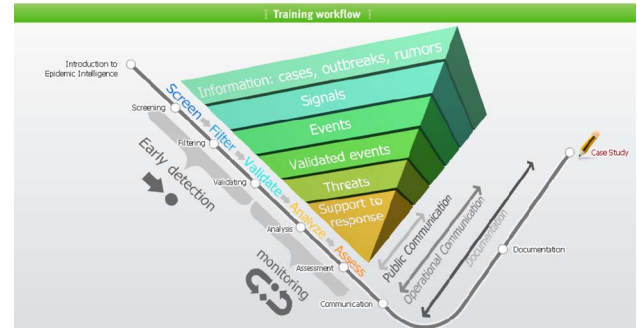


Figure 1. Process of Epidemic Intelligence (taken from [3])

The recent unprecedented explosion of user-generated content on the Web 2.0 Internet and increasingly more on realtime social media provide an enormous amount of data in machine-processable formats. These data streams geo-located with situation and context aware information about the users provide realtime data source about citizens health for public health benefits. However, as these big data signals are characteristics by high heterogeneity and diverse coverage, they each require different computational methods for processing and analysis for public health purposes: these methods include resolving and mapping spatiotemporal tagging; utilizing and verifying personal profiles; metadata-annotating multimedia files; and combining multiple methods for processing test-based tweets. Further, the real time speed comes at cost in a form of a high noise ratio and unreliability. More robust verification methods are required to effectively and reliably analyze these realtime streams for EI purposes, reduce number of false positives and enable deploying them in epidemic intelligence systems in a real public health operation.

Finally, mobile devices have become an indispensable part of everyday life. Integration of mobile information, environment and personal sensor-based digital traces provides a great potential for multi-faceted continuous data streams to be integrated in public health systems to improve quality, timeliness and response. While a number of prototype systems demonstrated a potential for public health (as discussed in section 3), the full integration with verification and response remain a major challenge.

3. DATA SOURCES FOR CASE-BASED AND EVENT-BASED SURVEILLANCE

In this section, we will describe the data sources providing the essential data streams for event-based and traditional case-based/symptom-based surveillance and discuss how these data

¹<http://ecdc.europa.eu/en/activities/surveillance/tessy/pages/tessy.aspx>

streams are monitored, how a signal is detected and verified. This sets the scene for the vision for the *public health dashboard* for verification and a roadmap towards this integrated solution as discussed in section 4.

The aim of this section is to bring description and illustrate on examples the types of traditional and novel data streams providing input for epidemic intelligence components: data mining, modeling and verification algorithms as required for the proposed dashboard, rather than presenting a fully comprehensive review of all systems (for a comparison study we refer to [2]), however, our classification is treating each system separately according to the *type* of data source and the *creator* of the content (e.g., user, news agency, etc) as these are key characteristics for effective application of computing algorithms and methods.

In Table 1, we categorize the data streams in this section by the following characteristics:

- T = timeliness (realtime, ‘days’ after event, ‘weeks’ after event)
- R = reliability (high reliability, medium reliability, low reliability)
- I = intention of sharing (personal, news, professional, seamlessly collected)
- A = collection/mining by EI systems (automated, human moderated)
- G = geo-located (Y, N, some)
- S = situation-aware (user/case context known, not known)
- ST = structured (yes, no)

Data Stream	T	R	I	A	G	S	ST
News/Online Media	Days	Med	New	A	Y	N	N
Digital Traces	RT	High	Per	A	Y	Y	Y
Social Media	RT	Low	Per	A	Varied	Y	N
Pro-Med	Varied	High	Prof	H	Y	Y	N
Labs/Clinical Reports	Weeks	High	Prof	H	Y	Y	Y
Participatory Systems	RT-days	High	Per	A	Varied	Y	Y

Table 1. Classification of Data Streams for EI

3.1 News and Online Media

Online news and Internet media resources are constantly covering public health events. Official media outlets, online newspapers, professional and lay blogs as well as personal home pages all became essential sources of information for early warning systems. Since EI systems have been scanning local

newspapers, it was demonstrated that media coverage of a small outbreak appeared days sooner than health authorities were informed through traditional surveillance processes.

Key EI systems screening media for epidemic events include: Health Canada developed system called GPHIN² (restricted access, multilingual), Joint Research Centre (JRC) funded monitoring system MedISys utilizing keyword extraction based on experts defined weighted taxonomies (open access, multilingual)³ [4], information retrieval is used in the PULS⁴ system [5] developed by the University of Helsinki. Further, the US-based BioSense initiative is a CDC (Centre for Disease Control and Prevention) project enhancing early detection and localisation of possible biologic terrorism attacks is discussed in [6]. HealthMap, developed at Harvard Medical School, is a Web-based monitoring system running since 2006 including data sources from other news feeds, such as Google News, the Pro-MED mailing list, World Health Organisation announcements and Eurosurveillance publications, among others [7-8].

3.2 Sensors, Digital Traces and Mobile Devices

This data stream enables monitoring seamless movement of populations and citizens’ digital traces via GPS-enabled phones, sensor networks, and credit/store cards continually and seamlessly collecting and recording information about our moves, physical locations, purchases, online preferences and payments. Without users actively nor explicitly providing personal data traces, location-aware applications take advantage of collected realtime personal information – these revolutionized the way we travel, drive a car, navigate on a street and find local information in our everyday life. However, there are new applications taking advantage of user location aware real time data sharing for health benefits, such as: ‘Flu Near You’⁵ and the EpiCollect [9-10].

3.3 Social Media Streams

Unlike numerous realtime digital traces collected seamlessly, the increase of Web 2.0 user-generated content actively shared via social networking tools such as Facebook and Twitter has been an unprecedented explosion. These signals also provide EI systems with a highly accessible source of real-time online activity. Facebook’s privacy setting allow users to restrict their profile content and activity, however, content posted on Twitter [11] is available in public domain and therefore freely searchable and analyzable using a provided API [12].

Further, Twitter provides an excellent way to sample large populations. In terms of epidemic intelligence, Twitter can be used to both track [13,14, 15] and even predict [16] the spread of infectious diseases several weeks before the official public health authorities in the UK.

² <http://www.phac-aspc.gc.ca/gphin/>

³ <http://medusa.jrc.it/medisys/homeedition/en/home.html>

⁴ <http://puls.cs.helsinki.fi/medical/>

⁵ <https://flunearyou.org/>

3.4 Pro-MED mailing list

In addition to automatically monitoring Internet news, human moderated email systems, such as the global Pro-MED⁶ - the mailing list for infectious disease professionals around the world, provide the forth data source. Unlike automated system, Pro-MED is human moderated thus can manually analyze redundancies and false positives but require time delay and inevitably has lower coverage due to human bias and the need to discuss unusual case. Pro-MED is run by the International Society for Infectious Disease.

3.5 Microbiological Labs and Clinical Reporting

Traditional surveillance systems provide hierarchical reporting channels for a set of defined diseases reported from primary and secondary care, and microbiological labs at local, regional and national levels. Microbiological laboratories contribute to surveillance by providing the data source of the highest reliability by lab confirmation of unusual disease patterns, specimen and/or dangerous pathogens. Clearly defined, enforceable and very reliable, the main disadvantage of these systems is the slow nature of reporting and the fact that they are prone to human error. Further, in low income countries without a robust surveillance systems public health reporting is often subject to low quality due to lack of resources, capacity or attempts to deliberately underreport outbreaks in order to avoid undesirable economic sanctions. Public health reporting is subject to inevitable bias and discrepancies during medical emergencies, humanitarian disasters or in non-democratic regimes [17].

3.6 Participatory Systems

The popularity of applications like ‘Quantify me’ recently provided another data stream for real-time public health monitoring – citizens’ self-reported symptoms and health-related information [18]. Citizens participatory surveillance is on rise in the US⁷ as well as in Europe, for example the Europe-wide Influenzanet platform⁸ running the Europe-unified Flusurvey citizens reporting project reached over 100 000 volunteers providing their symptoms on a weekly basis; another popular application is the Sickweather platform⁹. Rather than seamlessly sharing unstructured information on social media that could be mined for public health needs, these are dedicated platforms enabling users to pro-actively report structured data about their health, symptoms and conditions providing a magnitude more reliable epidemiological data source.

⁶ <http://www.promedmail.org/>

⁷ <http://www.cdc.gov/biosense/correlate>

⁸ <http://www.influenzanet.eu/>

⁹ <http://www.sickweather.com/>

4. INTEGRATED DIGITAL PUBLIC HEALTH SURVEILLANCE

In this section, we will present elements of an integrated digital public health dashboard and the process of mining data streams and subsequent analysis. This section defines the challenges and research roadmap highlighting future research directions.

4.1 The Dashboard Platform

The framework illustrated in Figure 2 depicts processes and components required for automated monitoring across multiple realtime data channels. We propose six computation phases for threat verification preceding manual actions taken by PH and EI experts.

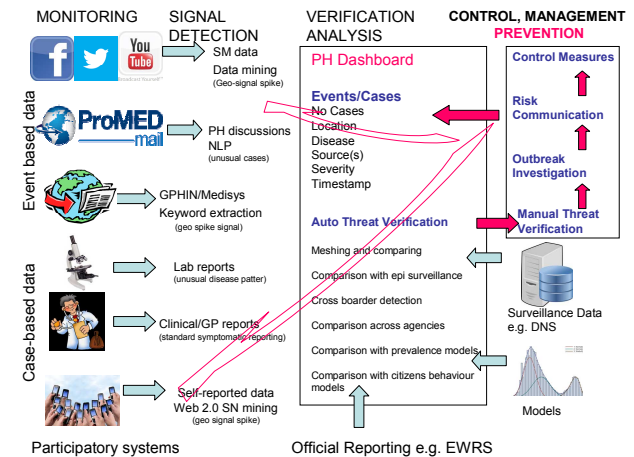


Figure 2. Integrated Digital Public Health

4.2 Signal Detection

According to the six data streams discussed in the previous section, signals monitoring and detection across the multiple channels requires different computational methods according to their structure and reliability (data mining, NLP, ML, data science, complex systems, social networks, knowledge extraction, context awareness, etc) in an interdisciplinary combination with epidemiology, disease modeling and public health. Identified signals through each data stream are validated according to the noise reduction required by the *reliability* of the data stream using appropriate methods (e.g., a Twitter generated threat will have different validation criteria than a lab-confirmed case).

4.3 Analysis and Threat Verification

The following steps are required to verify detected threats (semi) automatically validating the event and minimizing the number of false positives. Each steps indicates a set of research questions and computing challenges to address.

4.3.1 Meshing and Comparing Data Streams

Firstly, data streams (e.g., media rumors, social networks, laboratory reports, etc.) are compared using statistical methods,

GIS methods, data mining according to *reliability* and *situation-awareness* criteria.

This includes a cross-boarder detection as GPS-enabled streams could identify clusters of threats that would not traditionally trigger alerts in either region/country.

4.3.2 Comparison with Surveillance Data

Secondly, a comparison with appropriate surveillance data for the detected disease, season and country/region is performed in order to assess whether, for example, a detected spike in signal is indeed a seasonal increase rather than a threat. For example, at European level, the signal data signal will be verified against TESSy data. A comparison with laboratory results can be included at a later stage when these become available.

4.3.3 Comparison across Public Health Agencies

Further, to aid verification, the integrated dashboard system will check whether the threat was identified by other agencies, such as WHO EURO, ECDC, the PH institute in the country where the incident was identified (using knowledge extraction on reports or a database query on identified *structured* threats).

4.3.4 Comparison against Models

Further, the analysis tool also provides a correlation with models of disease spread and estimated likelihood of increase and/or effectiveness of, for example, targeted vaccination. This step strongly relies on availability of accurate models. Similarly, user-generated content in participatory systems and social media streams need validation against models of online and self-reporting behaviour, as these might follow significantly different patterns than the disease prevalence. This step is one of the most challenging as these models have not yet been developed.

4.3.5 Visualization

Finally, the verified events will be visualized on a map or other appropriate spatiotemporal visualization platform to further illustrate the time and location of the verified events in the geographical and global context, such as outlined in [19].

5. CONTROL AND MANAGEMENT

After the automated identification and verification process, experts-lead outbreak investigation can start with significantly lower manual effort. The platform can further support outbreak team compositions for field investigations by accessing profiles, expertise rather than relying on an ad hoc personal network of contacts.

5.1 Communication with Citizens: Risk and Prevention

In the past, governments' public health agencies and news media organizations had fully control over what was published by with regards to public health information and risk communication, especially during emergencies. However, Web and increasingly social media dramatically changed this landscape, a recent example was reporting of the swine flu in 2009 [20]. On one hand, these platforms provide a real-time channel for risk communication virtually at real time effectively targeting those in affected areas and/or exposed to a risk. On the other hand, these platforms could be a source of misinformation

and quickly spread rumors to large populations without editorial comment or experts' moderation. A recent study evaluating Twitter-based propagation of WHO declaration of swine flu as the pandemics on 11th June 2009 demonstrated that more reliable media (such as BBC) were shared more widely on Twitter than outlets of lower quality [21]. However, designing an effective risk communication strategy over new media remains a major challenge.

Public health prevention is an essential activity aiming to raise the awareness of evidence, knowledge of diseases, improve understanding of the risk and ultimately change the behaviour. The pink back arrows in Figure 2 illustrate a method of feeding the results of completed investigation back to citizens. This could be achieved two ways – either directly to the users (through Internet and social media) and populations (through participatory systems). This is a way of closing the loop while further encourage user sharing.

5.2 Citizens Engagement

Active citizens' engagement in participatory platforms is ultimately a voluntary activity. Effective methods for recruiting users, encouraging activity and retaining citizens' interests in sharing personal information thus contributing to self-reported longitudinal epidemiological databases require new user engagement research methods. Gamification, prizes and Web 2.0 collaborative competitions complemented by promotions through traditional mass media channels such as national television are currently being explored (for example, a positive outcome was measured by the Flusurvey project in January 2013 when over 150 new Italian users signed up to providing a weekly self-observed ILI (influenza like illness) symptoms having seen about 30 seconds shot on national TV in Italy¹⁰). Design of incentivizing methods and validation through a suitably designed metrics across promotional channels is essential [22].

6. DISCUSSION AND FUTURE CHALLENGES

The aim of this roadmap paper is to highlight the analytical steps required to be undertaken by computational research while making several important assumptions. In this section we briefly discuss the real challenge of public health surveillance and outline the existing legal context and dynamics between national and international data sharing

Firstly, not all data streams are available in public domain - data privacy policies requiring explicit user consent play an important role when designing data mining algorithms (e.g. Facebook). Secondly, querying surveillance databases as discussed in the previous section require access permissions from countries sharing the data with ECDC and WHO (e.g. TESSy dataset). Finally, official disease reporting is not a voluntarily activity but is enabled by contractual agreements at EU and WHO levels. At this section we give an overview of the current legal context.

6.1 Legal Framework

Further details are required to be incorporated to fully reflect the agreed reporting requirements signed by all WHO member

¹⁰ TV shot appeared on Italian TV channel 1 in January 2013

states in 2005. This agreed legal framework is called the International Health Regulations (IHR) defining the case definitions and conditions for sharing surveillance data with WHO and other international partners. The equivalent EU legislation [23], as well as their national level equivalents [24] constitutes of the Early Warning and Response System (EWRS), defined by EC decision 2000/57/EC and amended by decision 2008/351/EC, describing a series of steps required to be taken by member states in the EC in case of specific threats to the public of international importance. Exchange of information among member states has a clearly accredited structure and the EWRS network is private. While important for validation, decision about control measures and cooperation between member states and ECDC this information source does not provide a data stream for automated monitoring due to the private nature of the network.

6.2 Future Goals and Challenges

In order to achieve this vision, we have identified a number of major challenges including:

- The proposed integrated digital public health dashboard is not a single stand alone system but rather an interoperable set of federated and dynamically communicating dashboards. The platform is envisaged to run at local level (configured for local coverage for social media filtering etc), at regional levels, national and international level with customization and interoperability support in accordance with each agency needs.
- Signal detection across heterogeneous data using multidisciplinary computation and epidemiological methods including models for Web 2.0 user-generated data and digital traces providing a verification baseline
- Address practical issues surrounding mobile user data protection and anonymization within a heterogeneous international legal context
- Further research into methods for user engagements and retention for participatory systems

7. CONCLUSION

Digital technologies are rapidly changing traditional public health by bringing novel computational challenges and realtime big data sources in aid of epidemiology. Constant stream of user generated data on Web 2.0 platforms, mobile digital traces, social media and participatory systems are continually creating a stream of big location-aware data to be mined by early warning systems, analyzed for threats detection, automatically verified by correlating data streams and performing multiple verifications, before handing over the task of investigation and control to public health experts. Enhanced by risk communication features and feeding information back to user and populations for better engagement and increased prevention, the proposed integrated digital public health dashboard provides a roadmap illustrating the steps ahead of computer scientists to substantially enhance public health services in the 21st century.

8. REFERENCES

- [1] M Salathe, L. Bengtsson, TJ Bodnar, DD Brewer, JS Brownstein, C Buckee, EM Campbell, C Cattuto, S Khandelwal, PL Mabry, A Vespignani. Digital Epidemiology. *PLoS Computational Biology* 8(7): e1002616. doi:10.1371/journal.pcbi.1002616
- [2] P Barboza, L Vaillant, A Mawudeku, NP Nelson, DM Hartley, LC Madoff, JP Linge, N Collier, JS Brownstein, R Yangarber, P Astagneu. Evaluation of Epidemic intelligence systems Integrated in the early Alertin and Reporting Project of the Detection of A/H5N1 Influenza Events. *PLOS One*, March 2013, Volume 8, Issue 3, e57272
- [3] J Mantero, T Mollet, F Santos, P Kreidl, P Pattinen. Epidemic Intelligence, Guiding principles, ECDC, September 2010
- [4] Jens P. Linge, Ralf Steinberger, Flavio Fuart, Stefano Bucci, Jenya Belyaeva, Monica Gemo, Delilah Al-Khudhairi, Roman Yangarber, Erik van der Goot. MedISys — Medical Information System In *Advanced ICTs for Disaster Management and Threat Detection: Collaborative and Distributed Frameworks*. Eleana Asimakopoulou, Nik Bessis (eds.), (2010) IGI Global Press
- [5] Martin Atkinson, Jakub Piskorski, Erik Van der Goot, Roman Yangarber Multilingual real-time event extraction for border security intelligence gathering Counterterrorism and Open Source Intelligence. *Springer Lecture Notes in Social Networks, Vol. 2*. (Uffe Kock Wiil, editor). (2011) pp. 355-390
- [6] C. A. Bradley, H. Rolka, D. Walker, and J. Loonsk. BioSense: implementation of a national early event detection and situational awareness system. *MMWR. Morbidity and Mortality Weekly Report*, 54:11, 2005
- [7] HealthMap. Global health, local information. <http://healthmap.org/en>, 2013
- [8] Brownstein JS, Freifeld CC, Reis BY, Mandl KD (2008) Surveillance sans frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS Med* 5: e151 doi:10.1371/journal.pmed.0050151
- [9] <http://www.spatalepidemiology.net/epicollect/>
- [10] DA Aamesen, DM Huntley, EJ Feil, F al-Own, BG Spratt. EpiCollect: Linking Smartphones to Web Applications for Epidemiology, Ecology and Community Data Collection. *PloS ONE*, September 2009, Vol 4, Issue 9, e6968
- [11] Twitter, <http://www.twitter.com>
- [12] Williams, D.: API Overview <http://apiwiki.twitter.com/API-Overview>
- [13] V Lamos, T de Bie, N Cristianini Flu Detector – Tracking Epidemics on Twitter. In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2010), pp. 599-602.
- [14] V Lamos, N Cristianini Tracking the flu pandemic by monitoring the Social Web. In Proceedings of the 2nd

- IAPR Workshop on Cognitive information Processing (CIP2010), pp 411—416, 2010
- [15] V Lampos, N Cristianini. Nowcasting Events from the Social Web with Statistical Learning. *AM Transactions on Intelligent Systems and Technology*, Vol 3, No 4, Article 72, September 2012
- [16] M Szomszor, P Kostkova, E de Quincey. *#swineflu: Twitter Predicts Swine Flu Outbreak in 2009*. M Szomszor, P Kostkova (Eds.): *ehealth 2010*, Springer Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering LNICST 69, pages 18-26, 2011
- [17] Coulombier, D., Pinto, A., Valenciano, M.: Epidemiological surveillance during humanitarian emergencies. *Médecine tropicale: revue du Corps de santé colonial* 62(4):391-5. (2002)
- [18] <http://radar.oreilly.com/2012/05/quantified-me-personal-health-data.html> (accessed 3rd April 2013)
- [19] SI Hay, DB George, CL Moyer, JS Brownstein. Big Data Opportunities for Global Infectious Disease surveillance. *PLOS Medicine*. April 2013, Volume 10, Issue 4, e1001413.
- [20] B Duncan. How the Media reported the first day of the pandemic H1N1) 2009: Results of EU-wide Media Analysis. *Eurosurveillance*, Vol 14, Issue 30, July 2009.
- [21] M Szomszor, P Kostkova, C St Louis: *Twitter Informatics: Tracking and Understanding Public Reaction during the 2009 Swine Flu Pandemics*, IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology 2011, WI-IAT, Vol. 1, pp.320-323, 2011
- [22] E de Quincey, P Kostkova, G Jawaheer. *Promoting e-Health Resources: Lessons Learned*, M Szomszor, P Kostkova (Eds.): *ehealth 2010*, Springer Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering LNICST 69, pages 103-110, 2011
- [23] Commission Decision 2008/426/EC)
- [24] Roush S, Birkhead G, Koo D, et al. Mandatory reporting of diseases and conditions by health care professionals and laboratories. *JAMA* 1999;282:164–70.