The *CommonGround* Visual Paradigm for Biosurveillance

Yarden Livnat, Elizabeth Jurrus
Scientific Computing and Imaging Institute
University of Utah
Salt Lake City, Utah

Adi V. Gundlapalli, Per Gestland
School of Medicine
University of Utah
Salt Lake City, Utah

Abstract— **Biosurveillance is a critical area in the intelligence community for real-time detection of disease outbreaks. Identifying epidemics enables analysts to detect and monitor disease outbreak that might be spread from natural causes or from possible biological warfare attacks. Containing these events and disseminating alerts requires the ability to rapidly find, classify and track harmful biological signatures. In this paper, we describe a novel visual paradigm to conduct biosurveillance using an Infectious Disease Weather Map. Our system provides a visual common ground in which users can view, explore and discover emerging concepts and correlations such as symptoms, syndromes, pathogens and geographic locations.**

Keywords— ***biosurveillance; visualization; interactive exploration; situational awareness;***

I. INTRODUCTION

The mission critical need for the early detection of disease outbreaks coupled with an exponential growth in our ability to collect and analyze vast amounts of data has led to the development of a multitude of modern disease surveillance systems. In the past ten years, there have been many efforts by government agencies to better detect emerging biodisasters. Much of these efforts have gone into the creation of biocontaminant laboratories and public health infrastructure[2-4]. The main emphasis of current biosurveillance systems has been to facilitate the aggregation and analysis of large amounts of data from numerous and disparate sources. Their goal has been to assist public health officials by shifting the load from the resource-limited, error-prone, slow human-based analysis to a more robust, sophisticated, and fully automated computer-based approach.

In order to achieve these goals for biosurveillance, tools must be designed around the user's needs and accessibility of the data. This requires the engineering of solutions that can handle multidimensional data acquired from different sources, the ability to detect threats in real time, and perform pattern recognition to identify alarms[5]. Surveillance systems, such as ESSENCE[6,7], RODS[8], GermWatch[1], the pandemic quick look tool[9] and BioSense[10], have shown great promise in their abilities to address these goals. In particular, such systems provide the user with unparalleled access to a wide range of analytic algorithms and numerous views of the data. Nevertheless, access to the data through these tools is constrained by a minimal user interface. Often, views of only one or two variables are allowed at a time, making it difficult to infer information from disparate sources.

This work was funded in part by a grant from US Army Medical Research and Materiel Command, Telemedicine and Advanced Technology Research Center (TATRC, project # W81XWH0710699)

Actual bioterrorism events can be easily misdiagnosed as other, unrelated, infections and mistakenly go undetected for too long [3, 11, 12]. Visualization can elevate the comprehension of this information by fostering rapid correlation and perceived associations. To that end, the design of the display must support the decision making process: identifying problems, characterizing them, and determining appropriate responses. It is imperative that information be presented in a manner that facilitates the user's ability to process the information and minimize any mental transformations that must be applied to the data.

In this paper we describe CommonGround (Fig. 1), a visual paradigm [13] that aims to improve situational awareness in biosurveillance. Our work focuses on visualization of signature data that other systems collect, aggregate and analyze. We emphasize the discourse between the user and the surveillance system and leverage human perception and cognitive processes in order to facilitate and enhance comprehension. Our approach provides a *common ground* system for monitoring, exploration and discovery of diverse heterogeneous biosurveillance data. The system can also be used to identify known signatures and discover new ones.

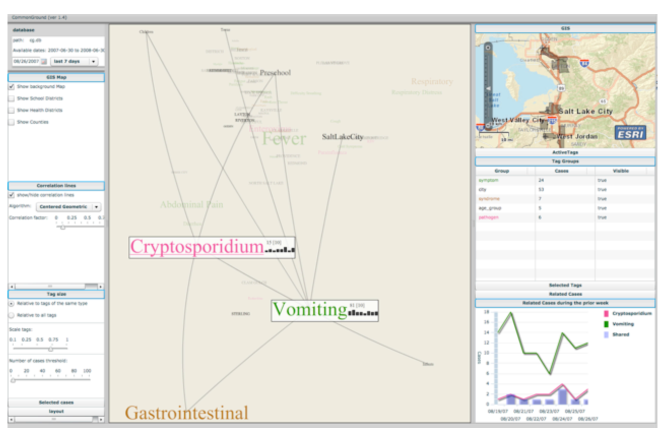

Fig. 1 Screenshot of the CommonGround interface showing various components of the system. Left panel: controls for selecting variables. Middle panel: the interactive tag cloud depicting active concepts and temporal correlations. Right panel: additional information such as a GIS map, a list of selected tags and charts.

II. BACKGROUND

Current surveillance systems concentrate on the aggregation and automated analysis of the data while neglecting the issue of information visualization. Most often, such systems rely on three forms of data visualization; Geographic Information Systems (GIS), graphs and tabular based presentations. These approaches emphasize qualitative or quantitative presentation, each with its own merits and pitfalls.

GIS based approaches are well suited for presenting the geographic distribution of a single variable but are limited in their ability to display relationships between multiple variables. GermWatch[1] is a website that successfully demonstrates this visualization by mapping virus outbreaks to zip codes. Such a map makes it easy to see the spread of viruses and identify geographically distinct outbreaks. However, the GermWatch map cannot show related symptoms, age distribution, and other syndromes at the same time. In essence a map is restricted to show only over variable at a time.

Graphs on the other hand provide an effective means to visualize relationships between several variables and for monitoring changes over time. The challenge is choosing the appropriate number and type of variables for presentation. In addition, some variables may relate to a localized geographic area, such as a city, while others may relate to information over large regions, such as across an entire state. Yet other variables may relate to discrete rather than contiguous locations.

Finally, tabular information is a familiar tool for analysts to use comparing high dimensional data, but its numerical nature can make it hard for the user to compare qualitatively. This is easily demonstrated if one considers the ease of comparing two images versus the equivalence of two tables of numbers. Differences or patterns in images are revealed quickly when we can view a visual representation of the data.

Recently, several systems have attempted to address these shortcomings by focusing on the interaction with the user via the use of advanced information visualization methods. One system in particular, the situational awareness tool[14] that is part of the Argus project, is similar to our earlier work on visual paradigms for network intrusion detection[15] and situational awareness[16]. This project based at Georgetown University is noteworthy for its handling of vast amounts of data and allowing user interaction.

III. SITUATIONAL AWARENESS

In this work we focus on developing a visualization paradigm that takes advantage of human perceptive and cognitive facilities in order to enhance the users' situational awareness for signature detection and support decision-making. Situational awareness is the ability to identify, process, and comprehend the critical elements of information about what is happening. The formal definition of situational awareness deals with the concepts at three separate levels[17]:

- Level 1 – *perception* of the elements in the environment.
- Level 2 – *comprehension* of the current situation.
- Level 3 – *projection* of future status.

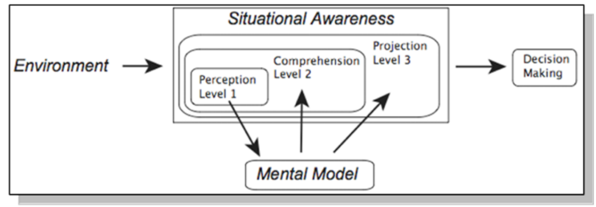

Fig. 2 Decision making model for situational awareness

Fig. 2 depicts the relationship between the three cognitive levels, the mental model, and the decision-making. In level 1, perception of attributes and dynamics of the environment, may be gained by a combination of visual, tactile, and auditory senses. Level 2, involves the comprehension of what the received data mean in relation to the relevant goals and objectives of the cognitive task being supported. This may include integration of the data to generate information, prioritizing and associating specific goal-related meanings and significance. Level 3 is achieved when one can predict how the environment will be affected in the future, based on the perceived data and its meaning. One must have a good understanding of the situation and the dynamics of the system in order to achieve level 3 situational awareness. The perception of time and the temporal dynamics of the environment are also important factors. The trends developed over time can play a critical part as well. Time is a strong component of Level 2 and Level 3[18].

It is important to note that comprehension relies on a mental model, which we create, or most often augment, based on our perception of the environment. In order to facilitate comprehension, we need to streamline and simplify the perception process and ensure that the data presentation matches, as much as possible, the user's mental model.

The cognitive process as depicted in Fig. 2 also highlights some of its limitations. In particular, comprehension depends on the quality and detail of the user's mental model. Information overload leads to an incomplete model or erroneously encoded information. Mental errors also stem from distraction, such as that resulting from too much irrelevant information, or the need to shift our focus of attention to another task to seek additional information.

IV. EXPLICIT AND IMPLICIT KNOWLEDGE

Knowledge encoding and representation of signatures is another important requirement. We distinguish between *a priori* explicit knowledge, that can be encoded and incorporated into a system, and implicit knowledge about the domain, that cannot. Examples of *a priori* explicit knowledge include the zip code of a hospital, or that a particular pathogen that causes specific symptoms. Other examples relate a medication to risks, or that a county comprises a collection of municipalities. A school can be an entity by itself, part of a group (based on its type; elementary, middle, and high) or be part of a single group comprising of all schools in a particular geographic area (i.e., a school district). User expertise represents implicit knowledge that is not part of a system and is not typically, or is technically challenging to encode. For example, an experienced public health practitioner in Utah will be aware that school absenteeism may be high on the first and last day of the hunting season. User expertise is also a crucial component in appreciating correlations between seemingly unrelated bits of information. It is thus important to facilitate and encourage the user to formulate and explore signature correlations that the system did not deduce based on its a priori knowledge.

Incorporating user expertise into a system is a challenging task. By its very nature, the system cannot be aware of what the user knows. The only way to take advantage of the user expertise is by creating a synergy between the system and the user via a dynamic discourse. Catering to the user's cognitive task by reducing information overload and reducing interface clutter that leads to distraction can help maintain focus. In particular, it is important not to force information on the user. For example, a carefully designed interface will only display relationships that won't distract the focus of the user. Rather, it will help in forming and testing hypotheses by supporting rapid confirmation or refutation of association.

Another subtle issue with user expertise is that users do not like to see information, such as correlations or dependencies between known signatures that they are already familiar with. If the system repetitively presents information the user considers trivial, or irrelevant to the cognitive task, then developers risk reducing the willingness of the user to participate in a dynamic discourse with the system.

This analysis of the implications of integrating implicit knowledge that is part of the user's expertise, leads us to the conclusion that *less is more*. That is, biosurveillance system developers should reduce the amount of information initially presented to the user but provide the user with easy access to all of the available information. Note that we do not advocate randomly removing or hiding information from the user. It is important to provide enough information scent[19] to help guide the user during the discovery and exploration phases.

V. THE COMMONGROUND PARADIGM

This paper describes a visual paradigm that provides a common ground (a unified framework) for representing, organizing, visualizing and interacting with a wide range of concepts and signatures. The approach focuses on a qualitative visual presentation of multi-dimensional concepts and their temporal correlations rather than the traditional quantitative display of raw data (e.g. charts or maps). The CommonGround conceptual model is based on an abstract representation of the raw data using meta data tags, similar to the notion of a tag cloud [20]. The size and color of a tag conveys information, such as the temporally relevant importance, the number of reported cases, or the semantics of the tag. In contrast to traditional tag clouds, which are rigid and static in nature, our approach employs a dynamic, free form layout that supports an interactive interrogation of the data. The layout of the tags is based on projection of high-dimensional data into a two-dimensional display using a multidimensional scaling (MDS) algorithm. Using this approach, concepts (tags) that relate to similar cases will tend to cluster together, even if these concepts don't have explicit relations between them. For example, a collection of tags -- *Toddlers*, *SkinIrritation*, *RespiratoryDistress*, *City1*, and *City2*.

To assist in identifying known signatures and discovering new signatures, we introduce or suggest correlations (i.e., edges) between tags that exhibit strong temporal correlation. The entire collection of active tags and the edges between them form a complex graph that can easily overwhelm the user and create an unintelligible display. There has been much work in recent years on graph reduction and simplification in particular with respect to social networks**.** The general approach is to rearrange the graph, combine nodes and bundle edges in the hope of reducing clutter. These approaches however are not appropriate in our visual paradigm, as the locations and sizes of the tags convey important information that we do not want to alter.

Our approach is based on the notion of delayed disclosure of information where only the tags, ignoring all the edges, are initially displayed. The aim of the initial display is to promote situational awareness and provide an overall view of all the active tags and their relative importance and temporal clustering. The relationships between tags are only disclosed though a dynamic discourse between the user and the system. As the user engages in an interactive exploration and discovery session by hovering or selecting tags, the system discloses relations (edges) between the selected tags and the rest of the tags. Our interface design attempts to reduce the clutter by showing only correlations that are higher than a user-controlled threshold. Using this approach, only the relations most appropriate for user at that moment are presented, i.e. only those that are relevant to the tags that the user is focusing on. This approach addresses the issue of implicit vs. explicit knowledge, as discussed in section IV. For example, consider the two tags, *SchoolAbsentee* and *HuntingSeason*. Seeing these two tags together, an expert physician in Utah will likely deduce, based on implicit knowledge, that school absentee is temporarily high due to the opening or closing of the hunting season. Other users, who may not be aware of such known relationships, will wonder about the implication of the *HuntingSeason* tag and will select it. The system will then disclose the potential dependency between the hunting season and school absentees by displaying an edge between the two tags, essentially incorporating additional explicit knowledge into the display.

Fig. 3 shows a screenshot of a typical display before and during and exploration phase, as it applies to Influenza Like Illnesses (ILI) data. In the following sections, additional implementation details of our software prototype are provided.

A. Tag Cloud:

A *tag* is a keyword assigned to a piece of information. Using tags as a form of associating meta-information to assist in searching and classification is most prominent in Internet and social bookmarking[20], and in presenting keywords that are shared between documents. Often several freeform tags are associated with a web page by one or more people. Each tag provides an additional piece of meta-information and thus implicitly associate knowledge about the tagged entity.

Fig. 3 Initial display (left) and during investigation (right)

A typical tag cloud comprises of a collection of tags that are organized in an alphabetic order. The font size of each tag reflects the number of time the tag has been used to annotate an entity. In general, the color of a tag does not convey any additional meaning expcept to potentially distinguish between adjacent tags.

B. *Knowledge Base*

A key difference between our visualization paradigm and traditional tag clouds lies in our definition of a tag. In our implementation, we use tags as first-class entities, that is, objects with semantics that have independent meaning and context. Our system's knowledge base includes a collection of tags and their semantic relationships. The semantic can include presentation information such as the tag's label, color and font. We also associate a group: such as 'symptoms', 'age group' and 'pathogen', with each tag. In general, the semantic can include relationships such as 'is-a', 'part-of' and 'suggests'. The knowledge base is independent of the rest of the system and different knowledge bases can be used for different domains. The knowledge base can be augmented with new tags and their semantic without interfering with the rest of the system.

C. *Tagging*

The CommonGround visualization is independent of the underlying raw data. Instead, we rely on independent, offline tagging processes that assign tags from the knowledge base to the raw data. The tagging results will then consist of a table of many-to-many relations, i.e. which tags are associated with which data item. Note that we do not need access to the raw data from a remote data source and it is sufficient to use a unique id for each data item.

For example, assume the raw data consist of a list of Influenza-Like-Illness (ILI) reported cases. A tagging process may associate a *Sex* tag (Male, Female) with the data to create a list of pairs ((id_1, Male), (id_2, Female), …). A separate tagging process may examine the age field in the data and assign an *AgeGroup* tag (Infant, Toddlers, Child, Teen, Adult, Elderly) to each reported case. Finally, a third tagging process may associate a *HighRisk* tag only to those reported cases that are either very young, elderly, or pregnant women. The only requirement is that all these tagging processes use the same ids for the reported cases. The list of assignments, i.e., (id, tag) along with the knowledge base are the only input the visualization needs.

The separation of the tagging process from the visualization system enables us to work with disparate sources without the need to actually gain access to the raw data. This has the added benefits of allowing the data curator to determine what portion of the data can be disclosed via the tagging processes.

In general one may view all the available data at once, however in biosurveillance one is typically concern with temporally relevant data. For real-time analysis, it is important to view only recent data but in forensic analysis users may choose data that occurred within certain time periods. In our work, we associate the incident day with each data item. When the user specify a period we extract only the tags associated with items from this time period and rearrange the display based on the new data.

D. *Layout*

In contrast to the traditional static alphabetic order approach used in tag clouds, we employ a dynamic and free form arrangement of the tags on the display. Our aim is to cluster tags based on temporal similarities between them. For example, if many Toddlers exhibit signs of Respiratory Syncytial Virus (RSV) than the two tags will be positioned close to each other on the display. Furthermore, if there are many Toddlers with RSV in both City1 and City2 than these two cities will be placed closed to the Toddlers and RSV tags. Note that this will lead to City1 and City2 to be placed near each other even if they are geographically very far apart. A quick look at the display can thus suggest to the user that an outbreak is occurring in multiple places in the state. It is important to note that this approach can only *suggest* that there might be correlation between RSV, Toddlers, City1 and City2. The alerted users can further investigate the situation in

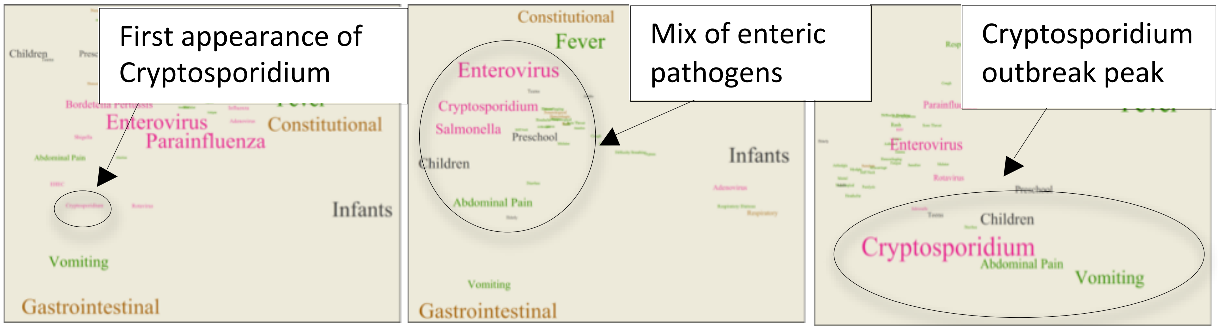

Fig. 4 Three screenshots depicting the course of a 2007 cryptosporidium outbreak in Utah

CommonGround as we describe bellow or use additional tools at their disposal to confirm to disprove this hypothesis.

In general, one can use a variety of layout algorithms to determine the best location of each tag. We have opted not to use algorithms, such as spring layout, that iteratively converges to a solution, as these methods tend to be slow and users do not like to see the tags moving around on the screen in undermined paths. Instead, we opt to use one-step algorithms such as Principle Component Analysis (PCA) and Multi Dimensional scaling (MDS)[21]. Each time the user selects a different period or changes the layout algorithm, or a parameter that affects the layout, the system recomputes a new layout on the fly, and animates the movement of each tag from its current location to the new one. New tags fade-in, while visible tags that do not participate in the new layout, fade-out.

To compute the layout we consider a tag as an n-dimensional vector, v, where n is the number of reported cases and v_i represent the probability that the tag is associated with data item i . The distance between any two tags, v_i and v_j, can thus be computed as the dot product

$$D_{ij} = \langle v_i, v_j \rangle \quad \text{for } i,j=1..m,$$

where m is number of a tags. To layout the tags we need to find a projection of the tags from the n-dimentional space into the two-dimensional screenspace that attempts to preserve these distances. More formally, the goal of MDS is to find a set of points, p, in R^2 such that

$$\|p_i - p_j\| \approx D_{ij} \text{ for all } i,j=1..m$$

The MDS is usually formulated as an optimization problem of finding a set of points, p, that minimize some cost function such as

$$\min \sum (\|p_i - p_j\| = D_{ij})^2 \quad \text{for } i<j.$$

In out work, we employed a classical MDS algorithm, which operates on a matrix of dissimilarities between pairs of tags. We compute the dissimilarities using a variety of correlation functions including Pearson, geometric and centered geometric functions. The user can choose which correlation function to use and the same function is used in the MDS algorithm to layout the tag and to compute temporal correlations, which we show as edges between the tags. The complexity of a naïve MDS is $O(n^3)$ and can be prohibitively high for large data sets. However, modern algorithims exist with complexity of only $O(n^2)$. In general, the number of tags is much smaller than the number of data items ($m<<n$) and thus the size of the similarity matrix D is relatively small.

E. Exploration

When the user engages in exploration and selects a tag, the system filters the underlying data to consider only raw data items that are associated with the selected tag. As a result, the relative importance of all the other tags will change based the new subset of data items. More formally, assume

T_i associated with data items $\{D^i_j\}$

and assume the user has selected tag T_s, then the filtered association of tag T^s_i will be,

$\{D^i_j\}$ iff D^i_j is associated with T_i and T_s

For example, assume that there are 100 cases of RSV and 30 of them are teens. Initially the *RSV* tag will be associated with the 100 cases but when the user selects the tag *Teen* then the *RSV* tag will be associated only with the 30 cases that are teens with RSV. As a result, when the user selects a tag, all the other tags will rescale to show their new relative importance. Some tags will shrink in size, others may grow, while others may disappear all together. While we do rescale the size of the tags, we do not recompute the layout of the tags, as this will introduce too many changes on the screen (tags moving to new locations), making it harder for the user to understand the effect and meaning of the selection.

The discovery and exploratory processes depend on a continuous, iterative discourse between the user and the surveillance system. To compensate for the limited amount of information one can process and store efficiently, our approach suppresses details that can easily be retrieved later. By assigning importance and priority through visual cues, our approach allows us to discard what is perceived to be less valuable information. This process depends on the ability to access additional details on demand, when our mental model is ready to incorporate them and when they become important enough to warrant storing them.

The common ground approach facilitates the discovery process by removing as much information as possible from the display but disclosing additional information when the user focuses on specific tags. To assist the user, the interface augments each tag the user selects with both quantitative information (the number of actually data items that a tag is associated with) and qualitative information (a sparkline of the tag's trend, i.e. the number of data items the tag was associated with). Fig. 4 (right) shows two selected tags, *Cryptosporidium* and *Gastrointestinal*, and the additional information they contain. The number of data items for each tag is shown in the format of $n[m]$, where n is the number of data items the tag is associated with and m is the number of data items that are associated with all the selected tags. In the example in Fig. 3, *Cryptosporidium* is associated with 13 cases, *Gastrointestinal* is associated with 175 but there are only 7 cases that are associated with both tags.

F. Signatures

The CommonGround visual paradigm conveys information at a glance through the layout: correlated tags are arrayed based on the strength of the correlation and individual tags are dynamically sized based on their relationships relative to all tags or tags of the same type. Fig. 4 depicts the course of the 2007 outbreak of cryptosporidium in Utah and illustrates this information conveyance. From this interaction, the user can see the initial case of Cryptosporidium and watch the outbreak progress.

VI. Future Work

In this work, we focused on facilitating user exploration of the data and as such, we do not currently have a backend system that can automatically identify known signatures. We imagine this knowledge would be derived from the incorporation of historical health data, such as CATCH[22]. In addition to being able to compare current data with historical data, data mining on known outbreaks from databases will help users identify similar patterns in illnesses. When combined with emerging relevant data, additional challenges lie in visually presenting known signatures to users so as not to confuse or bias their conclusions.

Another possible extension of this work is to incorporate local infrastructure and health resources into the CommonGround biosurveillance system. This might include information that directs where individuals suffering from a contaminant should go to seek the closest medical assistance [12] or help medical personnel isolate areas for possible quarantine. Finally, incorporating models that can run simulations and predict emerging outbreaks would help prevent and predict containment for threatening scenarios. We believe the CommonGround infrastructure, with the right user interaction, would be capable of being extended into these areas for broader use.

References

1. Healthcare, I. Germ Watch: Respiratory Virus Surveillance. Salt Lake City: Intermountain Healthcare; 2009. 2010; Available from: https://intermountainphysician.org/gw/Pages/default.aspx.
2. Gibbs, W.W., An uncertain defense. How do you test that a human Ebola vaccine works? You don't. Scientific American, 2004. 291(4): p. 20, 24.
3. Kaiser, J., Taking stock of the biodefense boom. Science, 2011. **333**(6047): p. 1214-1214.
4. Shea, D.A. and S.A. Lister. The biowatch program: detection of bioterrorism. 2003. Congressional Research Service [Library of Congress].
5. Berndt, D., A. Hevner, and J. Studnicki, Bioterrorism surveillance with real-time data warehousing. Intelligence and Security Informatics, Lecture Notes in Computer Science, 2003. 2665: p. 958-958.
6. Electronic Surveillance System for the Early Notification of Community-based Epidemics (ESSENCE) Available from: http://health.mo.gov/data/essence/index.php.
7. Lombardo, J.S. and D.L. Buckeridge, Disease surveillance: a public health informatics approach. 2007: Wiley-Interscience.
8. Tsui, F.-C., et al. Data, network, and application: technical description of the Utah RODS Winter Olympic Biosurveillance System. in Proceedings of the AMIA Symposium. 2002. American Medical Informatics Association.
9. Brigantic, R.T., et al. Development of a Quick Look Pandemic Influenza Modeling and Visualization Tool. in Proceedings of the 7th International ISCRAM Conference–Seattle. 2010.
10. Loonsk, J.W., BioSense—a national initiative for early detection and quantification of public health emergencies. MMWR Morb Mortal Wkly Rep, 2004. 53(53): p. 5.
11. Cole, L.A., The Anthrax Letters: A Bioterrorism Expert Investigates the Attacks that Shocked America. 2009: Skyhorse Publishing Company Incorporated.
12. Prakash, N., P. Sharada, and G. Pradeep, Bioterrorism: Challenges and considerations. Journal of forensic dental sciences, 2010. 2(2): p. 59.
13. Gesteland, P.H., et al., The EpiCanvas infectious disease weather map: an interactive visual exploration of temporal and spatial correlations. Journal of the American Medical Informatics Association, 2012.
14. Choi, J., et al. Situation Awareness Tool for Global Argus. in Visual Analytics Science and Technology, 2007. VAST 2007. IEEE Symposium on. 2007. IEEE.
15. Livnat, Y., et al. A visualization paradigm for network intrusion detection. in Information Assurance Workshop, 2005. IAW'05. Proceedings from the Sixth Annual IEEE SMC. 2005. IEEE.
16. Livnat, Y., et al. Visual correlation for situational awareness. in Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on. 2005. IEEE.
17. Endsley, M.R., Designing for situation awareness: An approach to user-centered design. 2nd ed. 2009: CRC Press.
18. Endsley, M. Design and evaluation for situation awareness enhancement. in Human Factors Society, Annual Meeting, 32nd, Anaheim, CA. 1988.
19. Pirolli, P.L., Information foraging theory: Adaptive interaction with information. Vol. 2. 2007: Oxford University Press, USA.
20. Begelman, G., P. Keller, and F. Smadja. Automated tag clustering: Improving search and exploration in the tag space. in Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland. 2006.
21. Duda, R.O., P.E. Hart, and D.G. Stork, Pattern Classification and Scene Analysis. Vol. 2nd. 2001: Wiley-Interscience.
22. Berndt, D.J., A.R. Hevner, and J. Studnicki. CATCH/IT: a data warehouse to support comprehensive assessment for tracking community health. in Proceedings of the AMIA Symposium. 1998. American Medical Informatics Association.