Spatiotemporal Zoning: A Framework for Enhancing Spatial and Temporal Granularity in a Web-based Health Surveillance System

Hutchatai Chanlekha\textsuperscript{1*}, Ai Kawazoe\textsuperscript{2}, and Nigel Collier\textsuperscript{1*}

\textsuperscript{1}National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo, Japan
\textsuperscript{2}Center for Women in Research, 2-1-1, Tsuda-machi, Kodaira-shi, Tokyo, Japan

*Corresponding author

Email addresses:

HC: hutchatai@nii.ac.jp
AK: zoeai@tsuda.ac.jp
NC: collier@nii.ac.jp
Abstract

Background
Current public concern over the spread of infectious diseases has underscored the importance of health surveillance systems for the speedy detection of disease outbreaks. Several such systems have been developed, including GPHIN, Argus, HealthMap and BioCaster. A vital feature of these systems is the geo-temporal encoding of textual signals. However, until now, systems have tended to adopt an ad-hoc strategy for processing geo-temporal information, normally involving the detection of locations that match pre-determined criteria, and the use of document publication dates as a proxy for disease event dates. Although these strategies are effective, they usually cannot discover time and places of the outbreaks in a better granularity. In order to improve the capability of current web-based health surveillance systems, we introduce a novel framework called spatiotemporal zoning with the purpose of providing a means for enhancing the granularity of spatiotemporal information.

Method
Our spatiotemporal framework classifies news articles into predefined classes based on their spatial and temporal characteristics, and recognizes the spatial and temporal attributes of each event. In order to study the reliability of the spatiotemporal annotation scheme, we conducted experiments by recruiting a group of annotators to annotate the same set of documents and analyze the agreement between human annotators. To quantitatively measure the agreement, we used 2 statistical measures: kappa, and the percentage agreement.
Results
The inter-annotator results are quite promising, more than 0.9 in average for event types and temporal attributes annotation, with a slight degradation in annotating spatial attribute. However, we have found that the annotators usually agree on the location annotation at country, state or province levels, which are the granularity level normally employed in the health surveillance systems.

Conclusions
In this article, we present and evaluate a novel spatiotemporal zoning annotation scheme. Given the present level of human agreement on annotation, and the current advances in natural language processing techniques, including the availability of language resources and tools, we believe that an automatic spatiotemporal zoning system can be achieved. In the next stage of this work, we plan to develop such a method and evaluate its usability in the context of a health surveillance system.

Background
The International Health Regulations (2005) [1], which entered into force on 15 June 2007, have bound 194 countries around the globe to a new legal framework for coordination of the management of events that may constitute a public health emergency of international concern. The implementation of the framework has underlined the importance of health surveillance technology, including novel methods that look for signal sources on the World Wide Web (i.e., the web). Example systems including MedISys [2] (EU), GPHIN (Canada) [3, 4], Argus [5], EpiSpider [6], HealthMap [7], and BioCaster [8, 9]. These systems generally gather outbreak data from a variety of electronic sources, e.g., online news wires, official reports, ProMED Mail, and so forth. Systems that provide map-based visualization usually geocode alerts to the country scale, with province-, state-, or city-level resolution for select countries [2, 6-10].
Geo-temporal encoding of textual signals is a key requisite in any health surveillance system, but current systems tend to adopt ad-hoc strategies, generally in the form of detecting the first disease and location pair that match predefined criteria or similar heuristics in order to identify the disease-affected location, and use publication date as an approximate occurrence time of the outbreak events. Although these strategies are effective in reducing both computational time and false positive detection, their limitation is that detailed information about the times and places of outbreak events may not always be discovered. This results from a characteristic of news, in which the details of the specific locations, as well as the exact time of an outbreak are usually mentioned later in a story. In order to improve the performance of current web-based health surveillance systems, the capability to reliably identify spatial information about outbreaks with better granularity as well as precisely recognize temporal information of the outbreaks is crucial. This paper addresses this issue, with a focus on public health surveillance, by proposing a novel framework that attempts to analyze news content with regard to the spatial and temporal attributes of events reported in text.

Detecting outbreak-affected locations with better granularity inevitably requires analyzing entire documents. It has been reported, however, that blindly searching for locations in full text, while increasing the detection sensitivity, can lead to excessive false positives [10]. This is because a news story does not always discuss only the current outbreak-affected location but can also refer to locations that are not directly affected but are related to the situation in various ways, e.g., neighboring countries that might be impacted, countries that provide medical assistance, previously affected locations, and so forth. The text capture shown in figure 1 exemplifies this situation. In order to effectively identify outbreak locations with better granularity, while
minimizing false alarms as much as possible, it is necessary to use a more sophisticated approach that enables systems to distinguish “locations” where the “current” outbreak is occurring from other locations. More specifically, the framework must as a minimum provide means to (1) identify outbreak locations at the finest level of granularity offered by the text and (2) distinguish newly reported data from historical and hypothetical data. Recognizing the precise time of the outbreak also faces the similar issue in that can be a lot of temporal expressions mentioned in the documents but some of them may not be directly related to the outbreak situation. One existing linguistic-oriented approach that is capable of overcoming these limitations is information extraction [11-13], which analyzes documents and extracts outbreak-relevant information, such as the disease, location, and time. The inherent problem that any information extraction systems generally face, however, is a trade-off between specificity and sensitivity. Sensitivity of extracted information usually has to be sacrificed in order to have a high-specificity system, significantly increasing the number of affected locations that are not detected. In health surveillance systems, the preciseness of the outbreak detection and alerts is very sensitive, so the information extraction patterns or models employed in such systems tend to be quite specific in order to reduce the false detection. However, this generally leads to a failure in detecting a number of outbreak affected locations. For example, the sensitivity (i.e. recall) of one reported information extraction system for the outbreak reporting domain was less than 50% [11].

To tackle the limitation of the existing systems, while reducing the problem faced in information extraction, we hypothesize that the time and place of an outbreak can be extracted through the capability to associate events reported in each text segment with the most specific geographical and temporal information available in news reports.
With this association, outbreak-affected locations can be identified by using more simple techniques such as text classification to detect text segments that indicate outbreak situations. At the same time, false alarms of past outbreaks can be avoided by taking the temporal information of events into consideration.

As we looked deeper into this kind of data, we found that certain types of news content cannot be anchored along a timeline, and hence, cannot be associated with temporal information. Among these types of content are sentences that provide general knowledge about certain subjects, such as diseases, and sentences that predict or express the possibility of certain situations. Moreover, the content of these kinds of sentences also possesses a different informative level with respect to outbreak situation analysis. Given this observation, we believe that the capability to classify news content is necessary not only for event time identification but also as a means to enable more flexible content processing.

To deal with all the issues mentioned above, in this article, we propose a novel framework called spatiotemporal zoning, which integrates the classification of news content and analyses of the spatial and temporal attributes of events, as another means to mitigate the inherent limitations of current surveillance systems. In designing the framework, we have tailored it to satisfy the basic requirements of analyzing spatial and temporal information about events. At the same time, since the detection of outbreak alerts has a high impact on society and the economy, we have tried to keep the framework simple enough to allow for a reliable automatic system to be implemented, using extant knowledge sources such as part of speech taggers, named entity recognizer, shallow parser. Our framework tries to overcome the limitations of current health surveillance systems in the following ways:
1) Classifying news content into predefined content classes based on its spatial and temporal characteristic

As mentioned earlier, certain classes of news content cannot be located in time. To be able to effectively analyze the temporal attributes of reported events, it is necessary to provide a means to classify news content according to spatial and temporal characteristics. This task is similar to what is proposed in existing works on argumentative zoning [14, 15]. These works, however, focus on scientific articles and classify texts into rhetorical zones in terms of argumentative and intellectual attribution. Since our task focuses on classifying news report content in terms of its spatial and temporal attributes, direct application of these frameworks may not be suitable for our needs.

2) Recognizing the spatial attributes of an event

Our scheme associates each reported event with its location. Although the scheme does not directly extract outbreak locations, its capability of associating reported events with their locations allows for further processing to distinguish outbreak locations from other locations in a news story. Since all event-denoting linguistic expressions in text are associated with the name of location where the event occurred, an outbreak-affected location can be identified by detecting certain sets of words or expressions that are related to the outbreak situation. Automatic recognition of geo-temporal information about events in natural language texts, however, is not a trivial problem. For event location recognition, our preliminary study on 100 randomly selected news documents showed that a location of a event is the one referred by the location name closest to the event expression in the story only 56.1% of the time. In order to correctly identify the locations of events, a more sophisticated technique applying syntactical analysis of sentences is needed.
3) Recognizing the temporal attribute of an event

In addition to spatial information, according to our scheme, each event reported in text is also anchored to the approximate time that it occurred. Recognition of this temporal attribute is crucial in reducing false alarms of past outbreaks, since it provides information that systems require to distinguish newly updated information from previously reported information. For the temporal aspect of the proposed framework, we based our design on one of the existing frameworks for temporal processing, namely, TimeML [16, 17]. TimeML is a rich specification language for event and temporal expressions in natural language text and can be considered a gold standard for temporal information. TimeML captures three phenomena in temporal markup: (1) It systematically anchors event predicates to a broad range of temporally denoting expressions. (2) It orders event expressions in text relative to one another. (3) It allows for a delayed (underspecified) interpretation of partially determined temporal expressions. Although the TimeML framework provides very rich, useful information regarding the temporal aspects of an event, it is quite complex, and it is considered impractical to develop an automatic system that is capable of performing the whole task [18]. Therefore adopting its full expressivity may not be necessary for this task.

In the following, we first define the events considered in our framework and give a concise description of spatiotemporal zoning. Next, we describe the classification of news content according to spatial and temporal information. Then, we introduce the spatiotemporal zoning scheme in detail. Quantitative and qualitative analysis of the resulting annotation is then extensively discussed. Finally, we examine the current limitations of our proposed scheme and the possibility of developing automatic systems based on this scheme. After giving our conclusions, we include a methods...
section to explain the methodology for studying the reproducibility of the annotation scheme, as well as to describe the annotation setting.

Methods

Characteristics of news report on disease outbreaks
In order to effectively design the spatiotemporal zoning framework, we first studied the characteristics of the disease outbreak news report.

News reports on disease outbreaks usually start with a brief summary of a whole story in the first few sentences or the first paragraph. This part generally states what happened, updated for the situation, together with the location and time of the story’s occurrence. While temporal information is usually explicitly mentioned in this summary, information about location is usually given on a broad scale, such as the national or provincial level.

Initial news reports or real-time update reports are generally condensed and include only the gist of information regarding to current situation, such as the disease, location, victims, and so forth.

Follow-up reports or daily reports, however, can be long and composed of various views on a situation. They may consist of updated information, as well as previously reported data about the current situation. They can also include basic knowledge about the disease, control measures for the situation, personal information about victims, or suggestions to residents in handling the spread of the disease.

Another characteristic of news in terms of spatial aspects is worth mentioning. In news reports, when no geographical location or well-known location is reported in the story, the situation typically occurred in the same geographical area as that of the news agency.
**Definition of events**
Since we are dealing with analysis of the times and places of events reported in natural language text, it is necessary to explicitly specify the definition of events.

Here, the definition of an event follows the definition used in the TimeML framework [17], where “event” is a blanket term for situations that happen or occur. Events are also considered as predicates describing states or circumstances in which something changes, obtains, or holds true, and which might need to be located in time.

In our framework, events may be expressed by the following means:
1) Tensed or un-tensed verbs, such as “die”, “occur”, and “spread”.
2) Certain sets of adjectives, such as “underway” and “ill”.
3) Prepositional phrases, such as “on board”, “under construction”, and “in progress”.

Prepositional phrases that indicate locations, such as “in Indonesia”, are also considered as events in our framework.

In the rest of this paper, by “event” or “event expression”, we mean a linguistic constituent consisting of a sentence, finite clause, non-finite clause, or phrase that contains a single event, as in the following example:

*Seventeen people have died and 41 have been admitted to hospitals in Sichuan, China, suffering from an undiagnosed disease.*

**Spatiotemporal zoning**
In this article, we propose a novel framework, namely spatiotemporal zoning, as a step towards future development of surveillance for recognizing geo-temporal information about outbreaks with better granularity. Before discussing the details of the framework, we first give the definition of spatiotemporal zoning.

Spatiotemporal zoning is a task that aims to partition text into segments based on the type of their content. Each segment contains events that occurred at the same geographical location in the same homogeneous time frame. Here, a homogeneous
time frame means that events or actions mentioned in the text segment overlap in
time, occurring either continually or sequentially. Basically, in this work, a zone is
regarded as a coherent text segment whose content is in the same class, occurring in
the same geographical region at the same time. The text capture shown in figure 2
below is an example of spatiotemporal zoning.

Class of news content
There are some text segments whose content cannot be anchored along a timeline.
Since our work deals with analysis of events’ temporal and spatial information, the
capability to distinguish groups of expressions from one another is necessary. In terms
of spatiotemporal characteristics, news content can be classified into three main
groups.
The first group is text segments that contain “happening events”. News content in this
class is composed of events that truly occur in the world and hence can be located
along a timeline. These events’ occurrences could be in the past or present, or will
definitely occur in the future. The second group of news content is “hypothetical
events”. Text segments falling in this group consist of events that may or may not
happen. These events are usually based on an expectation, prediction, belief, or
thought. The last group is “information”. This group contains either generic events
that cannot be specifically located in time, or eternal truths. Text segments classified
into the information group are usually given as basic knowledge, such as general
information about a disease, for news readers.
Regarding utility in a health surveillance system, news content in these classes also
contributes a greater or lesser degree of usefulness to the situation analysis task. Text
segments that contain happening events are usually the main interest of people who
want to learn about newly occurring situations or track the continuation of any
situation. Hypothetical events are usually not considered the main focus of situation analysts, since they are usually based on a prediction or personal opinion. These events might be useful, however, as information for prevention or a control strategy. Text segments that explain basic knowledge related to outbreaks, such as details about a disease or pathogen, are generally not regarded as significant to the event tracking or event detection task.

Since text segments of each class possess different levels of significance from the situation analysis point of view, we believe that it would be useful to include this information in our framework, as well. Therefore, we classified text content into 3 main classes, which are general information, hypothetical event, and Temporally-locatable event. In the following subsections, we describe the details and characteristics of each class.

**General information**
General information is usually non-eventive information, events that can not be positioned in space or time, general knowledge that is always true, or generic events [17]. The following are examples of general information:

1) General knowledge that is always true or events that cannot be located in time.

Example:  *Chikungunya is spread when tiger mosquitos drink blood from an infected person and, if conditions are right, pass the virus on when they bite again.*

2) Imperative and interrogative sentences, as well as recommendations, suggestions, and requests.

Example:  *Residents are recommended to stay away from the facility for one month.*

   *Students with symptoms should stay out of school until they have taken antibiotics for at least five days.*
3) Sentences whose subjects are linked to their predicates (e.g., characteristics, attribute, etc.) via a copula verb.

Example: (4) The victim is a 12-year-old boy.
(5) He is a resident of Boyolali district.

Informative-type events in the second and third groups usually convey information about the current situation, such as the details of victims, control measures, and so forth. In contrast, events in the first group, i.e., general knowledge, only provide basic information to readers.

**Hypothetical event**
Hypothetical events are those that are alternative or occur in other possible worlds.

Events in this group represent only the perspective or anticipation of the speaker.

While hypothetical events may or may not happen, forthcoming events are those that, without any unexpected circumstances, will definitely occur in the future, such as events that are planned. Generally, hypothetical events are the following:

1) Events introduced by verbs such as hope, believe, think, expect, and so forth.

Example: (1) Health agencies in Tripura were expected to launch culling operations later on Monday.

2) Events marked with a modal such as “may” or “might”.

Example: (2) Some of the patients might need to remain on the respirators for a couple of months.

3) Events introduced in a conditional sentence, i.e. in an if-clause.

Example: (3) If the virus mutates it could create a pandemic that would kill millions of people.

**Temporally-locatable event**
Temporally-locatable events are those that have happened, are ongoing, or will definitely happen. Among linguistic expressions that represent Temporally-locatable
events, there is a certain set of verbs that are worth mentioning. These verbs have a communicative function, and we refer to them as reporting verbs. From a grammatical perspective, the timing of reporting verbs has an influence on the temporal interpretation of events in the scope of a quoted speech. This influence is apparent in direct speech construction, where the time of an event inside a quotation is interpreted in terms of both the tense and time of the reporting verb and the event’s own tense. Given this characteristic, we believe that it is advantageous to separate reporting events from other happening events. For our framework, we decided to further classify Temporally-locatable events into two subclasses: reporting events, and normal events.

**Temporally-locatable event: Reporting event**
Reporting events describe the action of a person or organization declaring something, narrating an event, informing about an event, and so forth. [17] This type of event is used to give information about what people or organizations say or think. Reporting events are usually expressed by reporting verbs, such as “say”, “tell”, “announce”, and “report”. Examples of reporting events are shown below:

(1) The ministry **said** the boy might have been infected by sick chickens near his home.

(2) “It's very important to test the vaccine on humans and to produce it,” Van **said**.

(3) At least 15 children had died in the outbreak, Health Department Director General O.P. Singh **confirmed**.

**Temporally-locatable event: Normal event**
Normal events are Temporally-locatable events that are not reporting events.

Statistical investigation showed that more than 50% of the event expressions in our corpus are in the normal class. Examples of normal events are the following:

(4) A total of 14 of the 19 districts in the state, including Murshidabad, **had been affected**.
(5) Five days after returning to her hometown of Khon Kaen, she fell ill with Sars-like symptoms.

Temporal issue in spatiotemporal zoning design
The temporal aspects of spatiotemporal zones are somewhat more complicated than the other attributes. Therefore, we first discuss the details of temporal attribute design, before introducing the zone annotation scheme.

Since we define a zone as a text segment that consists of a group of events that occur in the same span of time, representing the occurrence time of events in a zone with one attribute may not be appropriate. One of the most obvious examples is a news report about continuation of events over a certain period, as in the following sentence captured from the news article:

*From 1 September to 8 November 2006, 16 deaths of meningococcal disease have been reported in Greater Yei County, Central Equatorial State of South Sudan.*

To enable our scheme to handle these cases, we regard the temporal attribute of the zone as a period with a starting time and an ending time. Hence, instead of defining the event time as one temporal attribute, we introduce two temporal attributes to indicate the starting time, STIME, and the ending time, ETIME, of the event occurrence period.

Another issue to consider is the relation between events in the zone and time. As reported previously [19], events and time can exhibit various relations, e.g., before, after, simultaneous, and so forth, as shown in the example below:

*The man was declared brain dead on Aug. 26, three days after suffering a serious head injury.*

Therefore, it is necessary for our scheme to provide a means to reflect the temporal relations between events in a zone and the starting and ending times of events’
occurrence. In our scheme, we introduce two attributes to express such temporal relations: STIME_Dir to indicate the temporal relation between events in the zone and the starting time of the occurrence period; and ETIME_Dir to indicate the temporal relation between events in the zone and the ending time of the occurrence period. Another important element is the reference time. Generally, the presence of reference time is not significant when we know an event’s absolute time. We often find cases, however, in which there not enough information is available to infer the absolute event time. These cases are usually those in which the occurrence time is represented by means of verb tense. With only the tense available as temporal information, only an approximate occurrence time relative to a certain reference point can be determined, as in the following example:

*At least 45 people have died of malaria in Jalpaiguri and Coochbehar Districts of North Bengal, senior health department officials said on Thursday.*

In the above sentence, all we know is that the event “died” started to occur at some time before the utterance time and continued to occur until then, at least. In these situations, the reference time plays an important role in temporal interpretation. Hence, we think it is necessary to include the reference time as a temporal attribute of a spatiotemporal zone.

**Granularity issue**

In outbreak news reports, temporal expressions at the level of time-of-day granularity are rarely reported. On the contrary, we usually find that temporal expressions in outbreak reports are at the level of a ‘day’ or a coarser period, such as a week, month, or year,. In terms of requirements, organization of news reports in health surveillance systems with regard to time is done at the level of a day, i.e., news is grouped and presented on a daily basis. Given these considerations, in our framework, we specify temporal attributes with ‘day’ granularity.
Application of temporal attributes
While spatial information is an attribute that is inherent to any type of zone, temporal information cannot be considered in that way. Temporal attributes can be applied to the events that can be located in time only, i.e. either normal or reporting events.

Spatiotemporal zoning annotation scheme
Earlier in the article, we defined the specification of the task, analyzed the types of zones, and discussed the temporal issue of the zone annotation task. In the following, we introduce the details of our spatiotemporal zoning scheme. We defined nine attributes for zone annotation, which are summarized in table 1.

ID: Zone identification attribute
This attribute represents a zone’s ID.

TYPE: Zone type attribute
This attribute indicates the type of events in a zone. There are four values for the TYPE attribute. These values are defined according to the classes of events. They are: “Event_Info” for the information type, “Event_Hypothetical” for the hypothetical type, “Event_Report” for the Temporally-locatable reporting type, and “Event_Normal” for the Temporally-locatable normal type.

As mentioned earlier, events with the information or hypothetical type cannot be located along a timeline. As a result, events with the Event_Info or Event_Hypothetical value for the zone type attribute have no temporal attributes (see below) marked in the zone.

LOCATION: A location attribute
The location attribute specifies the geographical location where the events in a zone happened. The location attribute can be a geographical location at any granularity available in a text.
ANCHOR VAL: Reference time attribute
The ANCHOR VAL attribute is introduced with the purpose of giving a reference
time, which is used for interpretation of other temporal attributes. The
ANCHOR VAL attribute consists of a normalized form of an anchoring date.
Generally, the default value of ANCHOR VAL is the document date or news report
date. In the case of direct speech construction, the timing of events in quoted speech is
interpreted with regard to the time of speaking. Hence, if events to be annotated are in
the scope of direct speech for a reporting event, the date of the reporting event is
selected as the value of ANCHOR VAL.

VAL: Relative time attribute
The value of the VAL attribute indicates a relative time with regard to the value in
ANCHOR VAL at which, based on the available textual information, the event in
focus hold true or happened. For example, if the events in the zone occurred in the
past, then the VAL attribute is “past”, but if the events started occurring in the past
and have continued until the present time, the VAL attribute is considered “present”.
Hence, there are three possible values for the VAL attribute: PRESENT_REF for
present events, PAST_REF for past events, and FUTURE_REF for future events.

STIME: Starting time attribute
STIME indicates the (approximate) starting time of the annotated events. The value in
STIME is the normalized form of a temporal expression based on the information
available in the text. If there is no explicit information indicating the starting time of
events in the zone, however, the value in STIME can be PAST, PRESENT, or
FUTURE, in relation to the value of ANCHOR VAL.

ETIME: Ending time attribute
ETIME is the same as STIME, with the difference that ETIME indicates the
approximate ending time of the annotated events.
**STIME_DIR: Relative direction with regards to starting time**

In many cases, we found that an event time is reported by using a preposition to indicate a temporal relation between the time that the event happened and a temporal expression or another event. In this circumstance, neglecting the existing of the preposition would result in the loss of detailed information for locating events along a timeline. In an attempt to lose as little explicit information as possible, we introduce temporal attributes that reflect this type of temporal relation, namely STIME_DIR and ETIME_DIR, which are explained next.

The STIME_DIR attribute represents the relative direction or orientation between the value of STIME and the events in the zone. In the TimeML framework, there are 13 temporal relations between events and temporal expressions or other events. These relations, however, are very detailed and quite complex. To eliminate unnecessary complexity, we decided to group these relations together and classified them into three main classes, which correspond to the possible values of STIME_DIR.

The value of STIME_DIR can be any of the following:

- **AS_OF**
  This class consists of the following types of temporal relations defined in TimeML: “simultaneous”, “including”, “being included”, “during”, “being held during”, “beginning”, “begun by”, “ending”, “end by”. The AS_OF relation is comparable to the OVERLAP relation in the SemEval-2007 TempEval task [18].

- **BEFORE**
  This class consists of the following types of temporal relations defined in TimeML: “before”, and “immediately before”.

- **AFTER**
  This class consists of the following types of temporal relations defined in TimeML: “after”, and “immediately after”.

- 19 -
ETIME_DIR: Relative direction with regard to ending time
ETIME_DIR is the same as STIME_DIR, except that it represents the temporal
relationship between the value of ETIME and the events in the zone.

Temporal attribute annotation example
For the example below, we illustrate the selection of each temporal attribute in zone
annotation, for a publication date of April 14, 2005.

1) Two of the cases were recently detected, between 2 and 8 April, in Hung Yen and
Ha Tay Provinces, respectively.

In the above example, the VAL attribute is selected as PAST, since the event “were
detected” was completed by April 8, before the publication date. The temporal
expression in this sentence indicates that the event occurred at some point in time
from 2 to 8 April. According to this, STIME and ETIME are annotated as “April 2,
2005” and “April 8, 2005” respectively, with both STIME_Dir and ETIME_Dir
selected as “AS_OF”.

Metadata
We introduce a metadata section into our framework with the purpose of providing
extra information related to the zone annotation task, including both the temporal and
the spatial aspect. Currently, the metadata section contains the information described
below.

Temporal-related Metadata
The temporal metadata section of an annotated document provides information related
to temporal attributes. The temporal metadata consists of the following:

1) News publication date

The news publication date is introduced in the metadata section since it is regarded as
the default value for the anchor date in the zone annotation task

2) The ISO normalized form of each temporal entity marked in the text
This second part of temporal metadata consists of the ISO normalized form of each
temporal entity marked in the text. The purpose of this metadata is to canonicalize
relative temporal expressions in news articles, such as “yesterday”, “today,” and
“tomorrow,” to absolute times, as well as to convert each temporal expression to the
same format. Since the smallest unit of time used for zone annotation is a ‘day’,
temporal entities with finer granularity than the ‘day’ level will be associated with the
normalized forms of their dates.
The text captured in figure 3 is an example of temporal metadata. The value marked
up by Anchor_Date is the news publication date. In the TIME_Norm field, the
‘time_id’ is the ID of each temporal expression appearing in the text, and ‘norm’ is
the ISO normalized form of that temporal expression.

Location-related Metadata
The location metadata consists of two parts. The first part provides information about
the site of a news agency. The second part provides relations between each location
appearing in text.

1) Location of news agency
Generally, a news agency’s location is specified at the country level. If the news
agency is local to a lower level of administration, however, a geographical location at
the city or province level is used.
This information can be used as the default value of an event’s location, since
observation indicates that many news agencies usually omit the locations of reported
events when they are local to the agencies’ locations.

2) Geographical relation
This part of location metadata provides information about the relations holding
between location expressions that are often found in text. Currently, two relations are
considered in this scheme: “IS_A”, and “PART_OF”.

- 21 -
“IS_A(A,B)” indicates that locations A and B are geographically the same. For example, IS_A(“USA”, ”United States of America”) indicates that the two location names both refer to the same geographical area.

“PART_OF(A,B)” indicates that location A is located in location B. For example, PART_OF(Tokyo, Japan) indicates that Tokyo is located in Japan.

In the future, we plan to add more information to the metadata section, consisting of the geographical grounding of locations in text.

**Unit of annotation**

As we mentioned earlier, events can be expressed by various kinds of linguistic component. It is likely that a single sentence may consist of multiple constituents expressing spatiotemporal-separated events. The data observation confirmed our hypothesis. From data observation, we found many cases where a single sentence consists of constituents (e.g. clauses or phrases) pertaining to different event type, or mentioning about events that occurred in different place or time. For example:

*The father became ill on 2 July with fever, mild cold, then coughed and was taken to the district hospital on 7 July where he died 10 days after onset.*

In such a case, giving a single annotation to the whole sentence would cause inappropriate aggregation of the spatiotemporal attributes of independent or irrelevant events into a single zone. As a consequence, detailed information, which is possibly important, may be made obscured. The clause-level annotation is considered to be more complicated to the annotator than the sentence-level annotation. However, as a trade-off for the purpose of retaining the detailed information as much as possible, we believe that the clause-level annotation should be more appropriate. Grammatically, a clause is an expression consisting of a subject and a predicate. It is a linguistically
stable and easily recognizable unit of annotation for the general readers. Therefore we decide on the clause-level annotation.

As we investigated a number of news reports, however, we found that events conveyed in different types of clauses contribute different levels of importance to the main situation reported in news. For example, noun-modifying clauses (e.g., relative clauses, noun-modifying non-finite clauses, etc.) usually give supplemental information, including possibly unimportant past events. Moreover, these clauses can be considered as the events that are not in focus from the reporter’s perspective. Therefore, we consider it necessary to specify in syntactic terms the constituent types that qualify for an independent zone. With this restriction, we can avoid over-generation of small, scattered zones, which would lead to results that are too complex and unnecessarily detailed. Table 2 lists the allowed constituent types.

For our framework, in general, an annotator processes text sentence by sentence. If all events conveyed in a sentence are temporally coherent events that occurred at the same geographical location, the whole sentence is annotated with a single zone, with attribute values derived from the events. If the events expressed in a sentence appear to occur at different times and/or locations, however, a new zone is instantiated. If an adjacent sentence (or clause) contains spatiotemporally similar events, then it is annotated together in the same zone as that of the first sentence or clause.

**Nested annotation**

Generally, zones are annotated sequentially. We observed, however, that some sentence structures express syntactical dependency between events for which it may be beneficial to allow them to be annotated in a nested manner. The typical cases of nested annotation in a sentence are a subordinate clause or non-finite clause (see Table 2), as in the following example:
He said that [the Indonesia case count has climbed to 226].

From our data investigation, we found that this nested construction mainly occurs when subordinate or non-finite clauses are the complements of verbs. The most frequent cases are a subordinate clause that is an internal complement of a verb such as “hope” or “expect”, or of a reporting verb such as “said” or “report”. Nested annotation not only is necessary because of complex sentence structures but also provides the possibility of greater flexibility in event analysis. For example, a health surveillance system might pay more attention to the main events in a main zone than to the complement events in a nested zone, or it might considered the importance or recentness of the main events to proportionally transfer to nested events.

**New zone instantiation**

In the zone annotation task, the boundary of a spatiotemporal zone is sequentially extended as long as it consists of only a set of clauses or sentences expressing events in the same class, occurring at the same location in the same period of time. Whenever a consecutive event is considered to be in a different class, or occurs at different locations or in a new, non-homogeneous time frame, a new zone is initialized with attributes according to its member events. Generally, a set of events is considered to occur in the same homogeneous time frame when each event in the set connects to every other event, either directly or indirectly, through the chain consisting of the 13 relations defined in TimeML, without having a narrative macro-move [20] in the chain. Applying the results of a study [21] on the relation between the usage of temporal markers and topic continuity/discontinuity, we hypothesize that without explicitly mentioning the dates and times of events, differences in event timing are not significant from a reporter’s perspective. Without temporal expressions or temporal adverbials, reporters seem to emphasize the continuity of sequentially
reported events rather than the timing of each event. Thus, when there is no strong evidence indicating discontinuity of events, we consider these events to occur in the same homogeneous time frame. A separate zone is also created when two textually consecutive events are mentioned to occur with different granularity of time or geographical location. In the case when a group of textually consecutive events occur in the same time frame but at many locations, the events can be annotated within one zone, where all events’ locations are identified by the LOCATION attribute.

In the case when two textually consecutive events are reported to occur on consecutive days, in order to keep all detailed information available in a news report, we decided to annotate the two events in different zones.

**Spatiotemporal zone annotation scheme evaluation**

In order to evaluate the proposed scheme, we conducted an annotation experiment on 100 pre-selected news articles. In this experiment, we were interested in one property, namely, the reproducibility, of our annotation scheme. The reproducibility property is the extent to which different annotators produce the same annotations. It measures the consistency of shared understanding held by more than one annotator [22]. Although the stability property, i.e., the extent to which one annotator will produce the same annotations at different times, is also important, we did not perform an experiment to evaluate the stability of our proposed scheme because of time limitations.

Nevertheless, it is commonly assumed that proof of the reproducibility of such a scheme implies its stability [22].

Here, we measured the reproducibility by recruiting a group of annotators to annotate the same set of documents according to the annotation guidelines developed for the spatiotemporal scheme. Three annotators, denoted as A, B, and C, participated in our experiment. Annotators A and B were given the same set of 50 documents, called
corpus Set1, while annotators A and C were given another set of 50 documents, called corpus Set2. We asked the annotators to perform full-coverage annotation, meaning that they judged all zone attributes for each span of text that contained any single verb, and then send us back the annotation results.

**Agreement measurement**

For agreement analysis of zone annotation, we considered agreements between annotators in annotating events for each type of zone attribute: the zone type, the location attribute, and the temporal attributes (i.e., the starting time, STIME; the ending time, ETIME; and the relative event time, VAL).

For quantitative agreement analysis, we used 2 statistical measures: kappa, and the percentage agreement.

**Kappa**

There have been different ways to evaluate agreement between humans for a task characterized as mutually exclusive category assignment. The evaluation matrices use such methods as the percentage agreement and Cohen’s kappa coefficient [23]. The kappa coefficient, $\kappa$, is a statistical measure of inter-annotator agreement for categorical items. It is generally thought to be a better measure of agreement than a simple percentage agreement calculation, since $\kappa$ takes into account agreement occurring by chance. The equation for $\kappa$ is the following:

$$\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)},$$

where $Pr(a)$ is the observed agreement among annotators, and $Pr(e)$ is the hypothetical probability of chance agreement. Regardless to the number of annotators, the number of items to be classified, or the distribution of the categories, $\kappa \leq 0$ means that there is no agreement other than what would be expected by chance, whereas $\kappa = 1$ means that the annotators are in complete agreement.
Since our zone-type annotation could be regarded as mutually exclusive category assignment, we used kappa as an agreement measure for the annotation task.

**Percentage agreement**

In annotating the location and temporal attributes of marked-up events, the annotators could freely select an event’s location as any location name appearing in the news report. Since the nature of the task was not exactly a mutually exclusive classification, the kappa coefficient may not have been totally suitable as an agreement measure. Hence, we used the simple agreement percentage as a measure to show the agreement characteristics between annotators in assigning the location and temporal attributes.

The percentage agreement was calculated by the equation below.

\[
PA = \frac{\text{Number of events in classA, with the same attribute value marked-up}}{\text{Number of events in classA}}
\]

**Data for zone annotation experiment**

The corpus for the zone annotation experiment task consisted of 100 news reports about disease outbreak events, randomly selected from the BioCaster gold standard corpus [24].

The first 50 files, denoted as Set1, were annotated by annotators A and B. There were 1086 events to annotate in Set1. The other 50 files, denoted as Set2, were annotated by annotators A and C. There were 886 events to annotate in Set2.

The number of events to be annotated and the number of sentences in each document set are shown in table 3. In figures 4 and 5, we show the distributions of documents that contained various numbers of sentences and events, respectively. Here, events were regarded as linguistic expressions that conform to the definition of an event given earlier.
The data represented in figure 4 indicate that, overall, documents with a length of 6-10 sentences had the highest proportion. According to figure 5, the largest proportion of documents contained 16-20 events to be annotated.

Figures 6 and 7 represent the details of outbreak news reports in our corpus in terms of the publication date and affected country respectively. Our corpus covered news articles published from 1996 to 2007. They reported outbreak situations of 44 diseases, on 45 countries world wide. In some articles, one disease outbreak were reported on multiple countries. On the other hand, some articles reported the spreading of multiple diseases within one country.

**Annotation guideline design process**
The annotation task and the design and revision of the annotation scheme were done as a loop process. That is, first, a draft of the annotation scheme was developed in terms of the usefulness of the event-based processing system from our perspective. Then, news articles were annotated document by document according to this scheme, further elaborations of the annotation specification were made, and so forth in a continuously looping manner.

**Annotators**
After finalizing the annotation scheme, we conducted an experiment in which three annotators participated. The first annotator, annotator A, was the first author of this paper. The second annotator, annotator B, was a linguist. The last annotator was an undergraduate student in the faculty of Liberal Arts at our institution. Since annotator A was involved in the design and development of the annotation scheme, she was very familiar with the task.
**Annotation tool**
For the annotation task, we developed an annotation tool for spatiotemporal zoning. Since the documents to be annotated could be marked up with various kinds of linguistic information (e.g., named entities), which might cause a visualization problem for an annotator in the zone annotation process, this tool hides these abundant tags from the annotator. It also provides a friendly interface for selecting the boundaries and attributes of each zone. The interface of the annotation tool is shown in figure 8. This tool is available online at spatiotemporal zoning’s code project web page [25].

**Annotation setting**
In the training phase, the zone annotation guideline was given to the annotators, together with a certain number of examples. After reading the guideline, the annotators were asked to annotate 10 articles selected as training materials, one article at a time. Each file's annotation was followed by a discussion session, with the purposes of clearing up misunderstanding of the guideline, settling disagreements between the annotators, and refining the guideline. After the training, the 50 articles of Set1 were given to annotators A and B, and the 50 articles of Set2 were given to annotators A and C, for annotation.

In the annotation task, the annotators performed full-coverage annotation, meaning that they judged all zone attributes for each span of text containing any single verb. The procedure for annotation was as follows. We asked the annotators to annotate the documents in the same order and send back the marked-up results after every 10 articles. The 10 resulting pairs of articles were thoroughly checked. When we found apparent mistakes as a result of human error, incorrectly annotated articles were sent back to the annotator for re-annotation, without telling the annotator where the mistake was or what the correct annotation should have been. When there was a sign
of misunderstanding, or when disagreements occurred repeatedly in the same context, discussion sessions were held in order to settle the confusion. These discussions aimed to reduce the number of undesirable mistakes in future annotation. The numbers of agreements of previously submitted articles, however, were counted before the discussion sessions.

In the following, the annotation results are analyzed both qualitatively and quantitatively. The quantitative analysis was done for every zone attribute annotation, with the purpose of revealing the reproducibility property of our annotation scheme. The qualitative analysis was done in order to provide insight into the data in terms of the detailed characteristics of each zone type. This information will be useful for future development of an automatic annotation system.

**Results and Discussion**

**Qualitative analysis of each zone class: Linguistic feature**

Before reporting our qualitative analysis of the experimental annotation results, we first mention the linguistic clues signaling each news content class, i.e., zone class. The following list represents the main linguistic feature types that we found helpful in judging the class of a spatiotemporal zone.

- Lexical item: words and phrases
- verb, verb phrase: including infinitive and gerund
- modal auxiliary (e.g., would, might, can)
- subject of the verb

In the subsequent sections, we present a detailed discussion and analysis of the characteristics of each zone class with respect to these features.
Qualitative analysis of Information zone
As we conducted the qualitative analysis of the information class, we found that the clauses in this class express three different characteristics. In order to effectively discuss the qualitative analysis results, we further separate the information class into three groups.

Clauses that represent attribute or state of entities
Clauses in this group represent the attributes or states of entities, such as a person, thing, or organization. These entities appear non-uniformly in news content. The most obvious feature signaling information in this group is the usage of a copula verb to assign a verb’s subject to its complement, through verbs such as “be”, “seem”, and “appear”. The following are examples of clauses that fit into this group of the information class:

(1) The fifth victim was a 21-year-old woman.
(2) The mother appeared to be healthy.

The existence of the auxiliary verb “can” in the sense of indicating capability also signals this group.

(3) It can cause serious problems, such as pneumonia and blood poisoning, and can even prove fatal.

Clauses that explain about general knowledge or things that are always true
This group of the information class is the most difficult category to detect, since there are almost no prominent linguistic features to distinguish it from the normal category, which is statistically a major class in news reports. One of the clues that we observed is that the subjects of clauses falling under this group usually refer to classes or conceptual-level entities, instead of to individuals, such as certain types of diseases, viruses, or patients. Examples of clauses regarded as members of this group are the following:
(1) **Encephalitis** causes an inflammation of the brain, **resulting** in brain damage or death.

(2) **HTLV-1** is transmitted through unprotected sex.

Regarding the location within text, clauses in this group usually appear at or almost at the end of news content. They can occur as part of a person’s quoted speech or as a sequence of stand-alone sentences. Moreover, they are generally separated from other types of events. When these clauses occur in a quotation, normally there are only clauses of this type. When these clauses are stand-alone sentences, they are often positioned in a separate paragraph.

**Clauses that represent events that can not be spatiotemporally located**
The most apparent clue to signal the existence of this group is the structure of a clause or sentence, which is in the form of an interrogative or an imperative. Clauses that express a suggestive meaning, recommendation, or request are also examples of this group. This type of clause is signaled by the presence of a modal, such as “should” or “must”. Examples of sentences in this group include the following:

(1) **Was it a migratory bird that got way off track?**

(1) **Residents** should remember their basic common-sense health practices.

(2) **Provinces with the disease** must immediately set up steering boards to control the disease.

**Qualitative analysis of hypothetical zone**
The strong signal of the hypothetical class is the presence of certain words that express the sense of possibility or an expectation. These words are modal auxiliaries, e.g., “would” and “should”, or words such as “probably” and “possible”, as in the following examples:

(1) A vaccination campaign launched last week should be able to bring the outbreak under control.
(2) The virus it had detected at a chicken farm was probably not the H5N1 strain dangerous to humans.

The occurrence of certain group of verbs, such as “hope”, “expect”, and “predict”, also indicates the hypothetical class, where the subordinate clauses of these verbs are considered members of the hypothetical class.

(3) In fact we expect that there will be more cases.

The modal “can” used in the sense of possibility is another indicator of the hypothetical class.

(4) I am always afraid that any illness can spread.

Qualitative analysis of report zone
The strongest clue to signal the reporting class is the verb itself. There is a set of verbs that has a communicative sense, such as “say”, “report”, and “tell”. The existence of these verbs suggests the reporting class.

(1) He said that hospitals in this southern Pakistan port city had been put on alert.

(2) She added that the severity of the imminent crisis would be worse than previous cases.

Another obvious indicator is when the verb is used as a main verb in direct speech construction, which is marked by the presence of quotation marks, as in the following example:

(3) “There have been no yellow fever reports since 1942 and it has been eradicated since then,” he said.

Verbs used in this manner are normally regarded as indicating the reporting class.

Qualitative analysis of normal zone
From our date investigation, which is later illustrated in detail in section xxx, we found that the normal class is the predominant class in news reports. It could be
considered the default class for clauses when they do not conform to the characteristics of any of the other classes.

There are no apparent features signaling the normal class. A wide range of verbs are used to express events in this class, as in the following examples:

(1) A 16-year-old Indonesian girl has died of bird flu.

(2) The government transferred drugs to those regions in order to control the epidemic.

The tense used with verbs in normal class clauses can be present, past, or future. It can have the simple, progressive or perfective aspect. The tense that we most typically found in news reports, however, is past tense, followed by present tense, with future tense occurring least frequently. This is a normal characteristic of news reports about recent or current situations. Present tense is usually used to stress the continuation or recentness of a situation. We often found the usage of present tense with the perfective or progressive aspect. Future tense is usually used in a planning context, or for a scheduled event, such as the expected time of laboratory confirmation.

**Quantitative analysis of zone attributes**

To evaluate the reproducibility of the proposed scheme, we analyzed the annotation results for each zone attribute in terms of an inter-annotator agreement measure. Since the annotation of each zone attribute can be different in nature, quantitative analysis of annotation agreement on different attributes by only one metric may not be appropriate. In our zone annotation task, the zone type annotation can be viewed as mutually exclusive category assignment, in which a certain value and number of categories are predefined. On the other hand, in annotation of the location and temporal attributes, the annotators can freely select attribute values according to the information available in news text. Therefore, in our analysis, we used two different
agreement measures suiting the different annotation characteristics: the kappa coefficient [23] for zone type annotation, and the percentage agreement for location and temporal attribute annotation.

In the following, we report our analysis of the annotation agreement for each of the spatiotemporal zone attributes, as well as an intensive analysis of cases of disagreement.

**Zone type annotation: annotation results**
Table 4 lists the proportions of events that were classified by each annotator. The trend in classification was the same for each of the three annotators, for both corpus sets. As might be expected, the number of normal events was the highest, since news reports generally talk about current situations, which are considered normal events according to our scheme. The second most frequent event was the reporting event, which we usually found in the context of reported speech, followed by events in the information and hypothetical classes.

**Zone type annotation: agreement analysis**
For the zone type annotation task, we used the kappa coefficient as a measure for agreement between annotators.

The results showed that our annotation scheme for zone types is reproducible, with $\kappa=0.87$ for annotators A and B, and $\kappa=0.9$ for annotators A and C. These numbers show that, given the annotation guideline, trained annotators could make distinctions between the different types of events in the same way.

In order to see which category distinctions are hard to make, we applied Krippendorff’s diagnostic for category distinctions. In this setting, all other categories but the one of interest are collapsed. The results are listed in table 5. Krippendorff’s diagnostic indicated that the most difficult distinction was one that results in omission
of the best K values. In our case, the most difficult classification is that among the normal, hypothetical, and information classes. The statistics support our data observation, in that we have never found (possible) reporting verbs used in the context of the information or hypothetical class, while there is confusion in categorization among the normal, information, and hypothetical classes.

Another tool for analysis of annotation is the confusion matrix. Table 6 shows the confusion matrices between each of the two pairs of annotators: A and B, and A and C. The cells along the diagonal show the decisions on which they agreed, while all other cells show decisions on which they disagreed. From the confusion matrices, we can see that there was no disagreement in judging between the reporting class and the hypothetical or information classes, between both annotators A and B, and annotators A and C. The disagreements between annotators A and B and between annotators A and C were found mostly in classification between the normal and information classes (40 times for annotators A and B, and 26 times for annotators A and C).

To gain insight into the disagreements in event classification, we conducted a detail analysis of these errors, which is described next.

**Zone type annotation: error analysis**

**Disagreements between the normal and reporting classes**

There are certain verbs that usually cause disagreements between annotators. While there is a certain set of verbs that are always considered to indicate reporting events, such as “say”, “inform”, and “report”, there are also many verbs that can be considered to indicate either reporting or normal events, depending on the context. These verbs include “show”, “concede”, “order”, “urge”, “recommend”, “ask”, and so forth.
Disagreements between the normal and information classes

Disagreements between the normal and information classes are the most common among all disagreements. The causes of these disagreements come mainly from the following two situations:

1) Difference in perception between an event situation and a non-event situation

In our guideline, verbs that indicate non-event situations, as in a sentence that describes the attributes or complement of its subject, are considered to indicate the information class, such as “The victim is a 12-year-old boy” and “His condition is very severe”. We often found, however, that there were many disagreements when the attributes of the subject were represented by adjectives. An examples of such a sentence is given below.

(1) A red rash is also visible on the bodies of affected persons.

Annotators who consider the above example as an event regard this sentence to express a perception of state by the author. This reading can be paraphrased as “I see a red rash ...”. We think that this type of sentence is naturally ambiguous as to whether it represents a state or an event.

2) Difference in perception of general knowledge or between a general event and a specific event

The events marked by boldface in the examples below are events for which there was disagreement between the two annotators.

According to our annotation guideline, a typical case of the normal class is temporally located and described by clauses with specific subjects, while a typical case of the information class is not temporally located, and the subject of the sentence can refer to either specific or non-specific entities.

Generally, when the verb’s subject is considered to refer to non-specific entities, this event should be classified as the information type. Different annotators, however,
might have different views of the verb’s subject, with one annotator considering the subject to refer to a general entity and another considering it to refer to a specific entity. The classification becomes more complicated when a temporal expression appears in the sentence, as in the following examples. When there is such a temporal expression, one might feel that the event can be located in time and thus classify it as the normal type. This situation, as shown in the examples, causes disagreement between annotators.

(1) People working in the wool industry used to be prone 50 years ago.

(2) The mushroom has been eaten in Japan for centuries.

These two situations happened quite often, and their judgment depends on the experience and perspective of each annotator. A decision tree that provides an explicit guideline to annotators in annotating the zone type might help reduce this type of disagreement.

Disagreements between the normal and hypothetical classes
Disagreements in this group come mainly from different interpretations between normal future events and hypothetical events. This situation occurred when one annotator felt that the linguistic element of focus (e.g., a sentence) represented something that will definitely occur in the future (i.e., a normal event), while another annotator thought that the element referred to a prediction or a conditionally possible situation (i.e., a hypothetical event). The example below illustrates this disagreement:

(1) “Our hospitals are not clean and we will find when this infection has passed that we will have another that takes its place. Until we get our hospitals clean we’re going to have infections all the time.”

In the above example, while one annotator annotated the events marked in boldface as normal events, the other annotator considered them as hypothetical events, giving a prediction of a situation.
From our data analysis, we found that there were disagreements in deciding whether “would” was used to signal the future aspect or the hypothetical sense (i.e., indicating possibility or willingness), as in the following example:

(2) The Red Cross said it *would spend* nearly one million Swiss francs (602,000 euros / 867,000 dollars) in a four-month awareness drive.

This situation occurred mostly in indirect speech construction. In conversion of direct speech to indirect speech, the tense of verbal elements within the reported speech is changed according to tense of the reporting verb. When the reporting verb is in past tense, such as “said”, the modal “will” is usually changed to “would”, while “would” is left as is. This construction sometimes causes difficulty in judging an original form of “would” used in indirect speech. Since the auxiliary “would” usually signals the conditional or optative mood (e.g., the hypothetical type), while “will” signals the future aspect of an event occurring, different judgments may result in different classification of such events.

**Disagreements between the hypothetical and information classes**

Disagreement in terms of the hypothetical and information classes occurred very often when there was hypothetical mention of general concepts or general knowledge, as in the following example:

(1) Because West Nile virus antibodies can stay within a person’s bloodstream for up to 500 days, it can be difficult to determine the date of infection.

While one annotator viewed “can be difficult” as indicating general information about West Nile virus, the other annotator considered it to indicate a hypothetical situation relating to a certain West Nile virus infection.
Zone location attribute annotation: annotation results
Since annotating the location attributes of marked-up events is not exactly a classification task, we used the percentage agreement as a measure to obtain inter-annotator agreement statistics.

Zone location attribute annotation: agreement analysis
The agreement statistics on the location attributes of marked-up events are listed in table 7. The numbers in the table were calculated only from pairs of events that were classified into the same class by both annotators.

With strict analysis, only location attributes that were annotated with exactly the same location(s) would be considered to indicate agreement in annotation. From the results, we found that the annotators seemed to disagree on location selection more often for events in the hypothetical and information classes than for events in the normal and reporting classes. For the information class, disagreements occurred most often when the event to be annotated consisted of general knowledge. We hypothesize that it is more natural for a human to consistently locate events that actually occur than to specify the locations of non-occurring events, such as information, as in the following example:

(1) Human T-cell Lymphotropic Virus, Type 1 (HTLV-1) occurs mostly in Japan, Caribbean countries and Africa. Doctors say most people who contract it will show no symptoms, but in about five percent of cases, it can lead to cancers of the blood and diseases affecting the nervous system.

The events marked in boldface were classified as the information type by both annotators. While annotator A considered these events as world knowledge, selecting the location as “World” for all three, annotator C considered them as information about specific locations and selected Japan, Caribbean countries, and Africa as the locations of these information events.
As we thoroughly examined the data, we found that even when the annotators selected different locations, these locations mostly appeared to be related. Specifically, either the locations selected by one annotator are located within the location(s) selected by the other annotator, or the locations selected by both annotators are partially the same. Especially in the case of locations that are partially the same, we observed that many of these selections occurred when one annotator selected only locations at a lower level of administration (such as selecting only villages), while the other annotator selected locations at both lower and higher levels of administration, which included the lower-level locations. Although we cannot say that these annotations represent 100% agreement, they are not totally different. They simply convey different levels of information. With loose agreement analysis, in which partial agreement or inclusion of a location is acceptable, the percentage agreement was very high, at almost 100% for most event classes for annotators A and B. The situation was the same for annotators A and C, except for the hypothetical class, in which the agreement was a little bit lower.

The percentage agreement can be regarded as the number of examples that can be annotated without ambiguity. If we assume that this approximate number of unambiguous cases can be 100% correctly annotated by an ideal system, we can treat this percentage as the approximate maximum score that an automatic system could achieve.

In surveillance systems, alerts currently are geocoded at the country scale, with province-, state-, or city-level resolution for select countries. Viewing the percentage agreement as an approximate maximum score for automatic annotation, the results show that reliable, automatic event location recognition can be obtained at the country level. As we further analyzed the data, we found that spatial annotations of normal
events usually had agreement or partial agreement at the state or province level, especially in well-characterized countries. This result indicates a promising possibility for identifying outbreak locations with a finer geographic resolution, which is a critical area in future development of effective outbreak detection.

Note that the annotations were done without providing the annotators access to a gazetteer or geographical ontology. We believe that with full access to geographical data, the agreement between annotators should be improved.

**Zone location attribute annotation: error analysis**
As we examined the raw data to find the characteristics of disagreements between annotators, we observed that the disagreements mostly occurred for the following reasons.

1) Event related to the movement of an object

There are certain sets of event-representing verbs that convey a sense of spatial movement, i.e., verbs that can be used with a prepositional phrase like “from <place>” and “to <place>”’. These verbs include “transfer”, “send”, “travel”, “move”, and so forth.

We often found that either the source or destination location was mentioned explicitly, but not both. When only the source or destination location was stated, partial disagreements between annotators usually occurred. Disagreement occurred when one annotator selected only a source (or destination) location that appeared nearby in the same sentence, while another annotator also attempted to infer the destination (or source) location from the discourse. Disagreement also occurred when both annotators attempted to infer missing locations, but the results of inference were not the same.

2) Location information via inference
Locations to be selected as event locations may be stated directly or can be inferred from context or discourse. When they were stated directly, the annotators were capable of choosing event locations with 100% agreement. When event locations were not mentioned explicitly, however, disagreement between annotators could occur. Without such explicit information at hand, we often found that while one annotator tried to infer the most specific locations according to what was available in the news content, another annotator tended to select locations with a higher level of administration, such as a location at the country or province level, whenever there was uncertainty. The following are examples of these situations:

(1) Dr. Ruth Improso denied that the number of typhoid fever cases in Bunawan town, particularly in Barangay (village) Libertad, went up to as high as 500.

For reporting events, annotators generally selected the location of the reporting agency as the event location. When the annotators did not have this information, however, there could be disagreement. In the above example, one annotator selected the fine-grained location reported to have experienced the outbreak, which was Bunawan. Another annotator, however, concluded that there was not enough information to assume that the “denied” event occurred in Bunawan and selected the Philippines as the event location, which is more general.

(2) Mekong Delta provinces are in the grip of a dengue outbreak with 38% more patients year on year. Measles is also afoot in northern Lai Chau Province. Deputy Minister of Health Trinh Quan Huan announced news of the outbreaks recently, saying that measures were underway to prevent further spread.

In the above example, while one annotator selected Mekong Delta provinces and Lai Chau as the locations of the “were underway” event, another annotator doubted
whether the measures were underway only in these affected provinces, and decided to select Vietnam, which is more general, instead.

(3) Iam Tanthong, a patient from Ban Paa Ngiw in Wiang Pa Pao district, said she had been admitted to hospital complaining of serious abdominal pain.

When the location of a hospital was unknown to the annotators, it could lead to inconsistent annotation. We found that while one annotator selected the hometown of the patient as the location of the “admitted” event, another one decided to select the most certain location, which was the city or country where the situation had occurred.

3) Disagreement in interpreting the location of an event

This kind of situation did not occur very often, but the annotators could sometimes be misled by unclear passages, as in the following example:

(13) So far, there's no hint of an outbreak in Canada. But Canadian health officials are watching what happens in the U.S. They may just start testing birds here to find out if they're carrying the virus. Because if they've got it, mosquitoes will pick it up, and then, people will be next.

While one annotator considered the events “start testing”, “will pick up”, and “will be” related to a hypothetical situation in Canada, another annotator chose the U. S. as the event location.

Zone temporal attribute annotation: annotation results

For the temporal attribute annotation task, we used the same measurement calculation as that for location agreement analysis, i.e., the percentage agreement

Zone temporal attribute annotation: agreement analysis

In agreement analysis for temporal attribute annotation, we considered the temporal attribute to be the same only when all of the temporally related attributes of each
event were consistently marked up by the two annotators. The agreement statistics for temporal attributes are listed in table 8.

For strict analysis, only events that were marked up with exactly the same set of temporal attributes were counted as agreement in annotation. In loose analysis, we considered any pair of annotations with the same VAL attribute as indicating agreement.

From the results, we can see that the agreement on temporal attributes was very promising for both pairs of annotators, even with strict analysis. This indicates that temporal annotation was less confusing for human annotators than location annotation, and that our instructions for temporal annotation were reproducible.

**Zone temporal attribute annotation: error analysis**

In order to locate the causes of disagreement, we investigated the annotated documents. We observed that disagreements mostly occurred when temporal information was not stated directly but could be inferred from the discourse.

There is an idiosyncrasy in news reports in which the first paragraph is used as a summary of the news story. In the case of an interview with a person in charge, the time of the interview is usually given in this first paragraph and then omitted in the rest of the story. Moreover, in this paragraph, interviewed people are usually referred to by a short description, such as “doctors from several hospitals” or “senior health officials”. This usually caused disagreement between annotators, especially in long articles. Each annotator might have judged differently whether each person appearing in the story was the same person or was part of a group mentioned in the first paragraph. This led to inconsistency between annotators in selecting temporal attributes. Figure 9 shows a capture of text representing one example of this situation.
Gerunds and infinitives were another source of disagreement in temporal annotation. The surface forms of these expressions are tenseless. Without tense, which is a basic signal of temporal information, annotators sometimes had different opinion about an event’s time.

Another situation for which we found quite often disagreements occurred when there was a temporal expression in a relative clause, as in the following example:

(1) *It had reports of 39 deaths from the outbreak of suspected acute hemorrhagic fever which began in January.*

There could be disagreement in that one annotator felt that the “had reports” event occurred in the same period as the beginning of the outbreak, while another annotator thought that the “had reports” event could have occurred at any time after the beginning of the outbreak.

Different judgment of the time span or length of an event was another cause of disagreement, as in the example below:

(2) *On Christmas day, a 24-year-old woman from Jakarta also died from the virus after buying a live chicken from a market.*

In the above example, while annotator A viewed “buying” as an event that occurred before Christmas day, annotator B considered both “died” and “buying” to have occurred on the same day, i.e., Christmas day.

**Discussion of the current scheme and future work**

In our spatiotemporal zoning scheme, we classify news content into four classes according to spatiotemporal characteristics, as well as the information that the content conveys. The scheme provides a means of anchoring events to their locations, as well as their approximate times of occurrence. In the current design, however, there are some issues that we had not previously considered.
The first issue concerns events relating to the spatial movement of an entity (e.g., “transfer”, “send”). Currently, we do not distinguish between the source and destination locations. This information can be critical, however, for detecting international travel health threats. For the next stage of our scheme, we plan to include this information when detecting verbs in a certain group.

The second issue is the polarity of events. This information is necessary in judging whether an outbreak event occurred. However, the sentiment analysis is a very complex task. Moreover, it expresses issues which are to some extent disjoint to those which influence spatiotemporal semantics. Therefore, in the current scheme, we did not consider positive or negative sentiments expressed in a sentence.

For the next stage, we plan to develop an automatic system capable of annotating news reports according to our proposed scheme. We have already started a preliminary experiment on automatic recognition of event locations. Given the present level of human agreement we have observed, our group’s previous study of automatic zoning [26] and the promising results for temporal relation identification [18, 27], we believe that an automatic spatiotemporal zoning system can be achieved.

Conclusions
In this paper, we started with the limitations of current Web-based health surveillance systems in enhancing the spatial and temporal granularity for outbreak detection. We then proposed a novel scheme to overcome these limitations by means of spatiotemporal zone analysis. In order to provide a means to enable outbreak detection in a better granularity in both spatial and temporal, we introduced an annotation scheme for outbreak news reports, collected data annotated according to this scheme, and gave a comprehensive analysis of the characteristics of the data. We also conducted experiments to study the reliability of the annotation scheme. The
results are quite promising, showing that the proposed scheme is reliable and can be learned by different annotators. The inter-annotator scores are more than 0.9 in average for event type and temporal attributes annotation, with a slightly degradation in annotating spatial attribute. However, we found that the annotators usually have high agreements on location annotation at the country, state or province levels, which are the geocoded level normally employed in the health surveillance systems. We also discussed the problems and difficulties that cause disagreements between annotators. Although we present only the framework in this paper, we believe that an automatic system can be developed by applying this scheme. Current advances in natural language processing technologies, as well as the availability of tools and resources, can provide a methodology to tackle each sub-problem in spatiotemporal zoning. In the next stage of our work, we plan to develop such an automatic zone annotation system and evaluate it with news articles reporting outbreaks. Although we have focused mainly on analysis of news articles, we expect that our approach can be applied to other types of outbreak research materials, such as official reports and ProMED-mail.

**Competing interests**
In the past five years, we have never received reimbursements, fees, funding, or salary from any organization that may gain or lose financially from the publication of this manuscript. We do not hold any stocks or shares in any organization that may financially gain or lose from the publication of this manuscript. Moreover, we do not hold or apply for any patents relating to the content of the manuscript. We haven’t received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to this manuscript’s content. Besides, we also do not have any other financial competing interests.
Authors' contributions
This work was directed by NC. HC carried out the framework design and analysis with the technical support and comments from NC. AK participated in the framework design and provided linguistic support. HC carried out the annotation experiments. All authors contributed during the whole length of the framework development and writing the paper. All authors read and approved the final manuscript.

Acknowledgements
We gratefully acknowledge the kind support of Mukda Suktarachan and Chotika Tunleng for participating in the annotation experiments. We would like to thank Mike Conway for proof-reading some parts of the paper. We would also like to express our gratitude for the helpful comments from the anonymous reviewers.

References


24. **BioCaster text mining project.** [http://biocaster.nii.ac.jp](http://biocaster.nii.ac.jp)


28. **Google Maps** [http://maps.google.com](http://maps.google.com)


Figures

Figure 1 - Various locations with different roles in outbreak news reports
Location names occur in the news reports are not always the location of the outbreak. In the text captures illustrated in the figure, Japan, Caribbean countries, and Africa are referred to as a location where HTLV-1 usually occurs, while South Africa and U.S. are the countries that provide the medical assistance to the affected country.

Figure 2 - Text capture of spatiotemporal zoning in a news report
Text is marked-up with spatiotemporal zone according to the annotation guideline. The first zone is report zone consists of one event, which is reported. This event occurred in Yei County, Central Equatorial, in Sudan from 1 September to 8 November 2006. These spatial and temporal information are represented in the zone’s Location_id, STime, and ETime attributes, respectively. The second zone also consists of one event, which is crossed. This event is annotated as occurred in Yei County, in the last week of October 2006, according to information available in the news report.

Figure 3 - Temporal metadata in the spatiotemporal zoning scheme
Temporal metadata consists of anchor date, which is the publication date of the news articles, and the normalized form of each temporal expression occurs in text. The
normalized forms and the original forms of temporal expressions are linked to each other via time_id.

Figure 4 - Distribution of the number of sentences, including partial sentences
This chart represents the distribution of the number of sentences in our corpus. In corpus set 1, most of the news articles contain 6 to 20 sentences, while in corpus set 2, the highest proportion are the articles that contain 6-10 sentences.

Figure 5 - Distribution of the number of events to be annotated
This chart shows the distribution of the number of events in each document in our corpus. The majority of the documents in overall consist of 6 to 25 events.

Figure 6 - Distribution of news articles in the corpus in terms of the publication date
This chart shows the distribution of news articles in our corpus in terms of the publication date. The corpus consists of news articles whose publication dates range from 1996 to 2007. However, the majority of the news articles were published from the middle of 2005 to the end of 2006.

Figure 7 - Distribution of outbreak situations reported in our corpus, classified in terms of outbreak-affected country
This figure represents the outbreak affected countries reported in news articles in our corpus. The map illustration was created by using Google Maps API [28] for the visualized purpose of location distribution. The chart on the top-left corner of the figure shows the number of documents that report the situation in each country. Remark that, in our corpus, although most of the articles reported the outbreak within one country, there are also some documents that reported the outbreak situations in many countries.

Figure 8 - Interface of the spatiotemporal zone annotation tool
Figure 9 - Example of co-referring of events
Text capture shown in the figure exemplifies the situation which multiple clauses refer to the same real-world event. In the text example, the phrase “Medical Service director-general Dr Chatri Banchuen said”, “Chatri added”, “hospital director Dr. Jessa Chokedumrongsuk said”, “hospital director Dr. Vinit Pua-pradit said”, and “the doctor claimed” are parts of the event previously mentioned in the clause “doctors at several hospitals said yesterday”.

Tables

Table 1 - Zone attribute for spatiotemporal annotation

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Represents zone’s ID Number</td>
<td>Number</td>
</tr>
<tr>
<td>EVENT TYPE</td>
<td>Indicates the type of events in a zone</td>
<td>“Event_Information”, “Event_Hypothetical”, “Event_Report”, “Event_Normal”</td>
</tr>
<tr>
<td>LOCATION</td>
<td>Specifies the geographical location where the events in the zone happened</td>
<td>The location attribute can be any politically or geographically defined location. To be specific, any location expression that is considered annotatable according to the specification used in the BioCaster system [29] can be a location attribute.</td>
</tr>
<tr>
<td>Temporal attribute</td>
<td>ANCHOR_VAL</td>
<td>Indicates a reference time used for interpretation of other temporal attributes</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>VAL</td>
<td>Indicates a relative time with regard to the value of ANCHOR_VAL, at which, according to the available textual information, the events in the zone hold true or happened</td>
<td>PRESENT_REF for a present event, PAST_REF for a past event, and FUTURE_REF for a future event</td>
</tr>
<tr>
<td>STIME</td>
<td>Indicates the (approximate) starting time of the events in the zone</td>
<td>The normalized form of a temporal expression reported in the text, or “PAST”, “PRESENT”, or “FUTURE”</td>
</tr>
<tr>
<td>ETIME</td>
<td>Indicates the (approximate) ending time of the events in the zone</td>
<td></td>
</tr>
<tr>
<td>STIME_DIR</td>
<td>Relative direction/orientation between the value of STIME and the events in the zone</td>
<td>“AS_OF”, “BEFORE”, “AFTER”</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>ETIME_DIR</td>
<td>Relative direction/orientation between the value of ETIME and the events in the zone</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2 - Constituent types qualifying for an independent zone annotation**

The square brackets in the examples indicate zone boundaries. Here, A and B are considered as clauses representing events that are either spatiotemporally different or different in type.

1) A sequence of sentences

2) A sentence

3) Coordinating clauses
   For example: [A] [and B]

4) Subordinate clause [30]
   - Subordinate clause introduced with a subordination marker (e.g., that, whether, and if)
     For example: [A said that [B].]
   - Subordinate clause introduced with a word functioning as the head of the constituent.
     For example: [A] [when B.]
5) Non-finite clause [30]

- Infinitive

For example: [A [to B].]

Note: If an event expressed by an infinitive cannot be located in time, such as an event expressing a purpose, goal, or intention, we do not consider it to qualify for being in an independent zone.

- Gerund

For example: [A, [saying …]].

<table>
<thead>
<tr>
<th>Table 3 - Data statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corpus</td>
</tr>
<tr>
<td>Set1</td>
</tr>
<tr>
<td>Set2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4 - Proportions of events classified by each annotator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corpus</td>
</tr>
<tr>
<td>Set1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Set2</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5 - Krippendorff’s diagnostics for category distinction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone class</td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>Reporting</td>
</tr>
<tr>
<td>Hypothetical</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>Information</td>
</tr>
</tbody>
</table>

Table 6 – Confusion matrix between annotators A and B on Set1 and between annotators A and C on Set2

<table>
<thead>
<tr>
<th>Annotator</th>
<th>Normal</th>
<th>Reporting</th>
<th>Hypothetical</th>
<th>Information</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>543</td>
<td>6</td>
<td>11</td>
<td>27</td>
<td>587</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>247</td>
<td>0</td>
<td>0</td>
<td>264</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0</td>
<td>51</td>
<td>6</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>0</td>
<td>1</td>
<td>158</td>
<td>172</td>
</tr>
<tr>
<td></td>
<td>579</td>
<td>253</td>
<td>63</td>
<td>191</td>
<td>1086</td>
</tr>
<tr>
<td>C</td>
<td>436</td>
<td>3</td>
<td>2</td>
<td>12</td>
<td>453</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>242</td>
<td>0</td>
<td>0</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>36</td>
<td>13</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>0</td>
<td>2</td>
<td>138</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>460</td>
<td>245</td>
<td>40</td>
<td>163</td>
<td>908</td>
</tr>
</tbody>
</table>

Table 7 - Agreement statistics for temporal attributes annotation

<table>
<thead>
<tr>
<th>Annotators</th>
<th>Normal</th>
<th>Reporting</th>
<th>All classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A and B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strict</td>
<td>0.92</td>
<td>0.97</td>
<td>0.94</td>
</tr>
<tr>
<td>Loose</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>A and C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strict</td>
<td>0.95</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>Loose</td>
<td>0.99</td>
<td>0.95</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table 8 - Agreement statistics for location attribute annotation

<table>
<thead>
<tr>
<th>Annotators</th>
<th>Normal</th>
<th>Reporting</th>
<th>Hypothetical</th>
<th>Information</th>
<th>All classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A and B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strict</td>
<td>0.82</td>
<td>0.84</td>
<td>0.80</td>
<td>0.70</td>
<td>0.806</td>
</tr>
<tr>
<td></td>
<td>Loose</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>A and C</td>
<td>Strict</td>
<td>0.75</td>
<td>0.78</td>
<td>0.58</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Loose</td>
<td>0.99</td>
<td>0.98</td>
<td>0.89</td>
<td>1</td>
</tr>
</tbody>
</table>
Health officials in Nunavut will begin asking all expecting and nursing mothers to undergo screening for a rare and untreatable virus that has appeared in the territory.

Human T-cell Lymphotropic Virus, Type 1 (HTLV-1) occurs mostly in Japan, Caribbean countries and Africa, affecting between 15 million and 25 million people worldwide.

Doctors say most people who contract it will show no symptoms, but in about five per cent of cases, it can lead to cancers of the blood and diseases affecting the nervous system. The development of those conditions can take 10 to 20 years.

It's rarely seen in Canada, but as many as 20 people in Nunavut have tested positive for it.

Since January 2006 diarrhea in Botswana has claimed the lives of 446 babies.

Botswana is struggling to contain the disease and is currently getting assistance and expertise from South Africa, the Centers for Disease Control in the U.S., the World Health Organization and other international health organizations. Over 11 000 children have already been affected by the disease.
From 1 September to 8 November 2006, 16 deaths of meningococcal disease have been reported in Greater <NAME cl="LOCATION" id=2>Yei</NAME> County, Central <NAME cl="LOCATION" id=3>Equatorial</NAME> State of South <NAME cl="LOCATION" id=4>Sudan</NAME>.

The epidemic threshold was crossed in this county during the last week of October.
Figure 3
Distribution of the number of sentences

Figure 4
Figure 5

Distribution of the number of events

Number of documents

Number of events

Set1
Set2
Set1+Set2
The Ministry of Health in Cambodia confirmed today that a 20-year-old woman from Kampot province, who died on 19 April in a hospital in Viet Nam, was the country's fourth reported case of avian influenza. The woman tested positive for avian influenza at the Pasteur Institute in Ho Chi Minh City.
The patients suffering botulism poisoning who were rushed to Bangkok from Nan remain in a critical condition, doctors at several hospitals said yesterday. While the 17 people are stable, they are not yet out of danger, Medical Service director-general Dr Chatri Banchuen said.

..., Chatri added.

..., hospital director Dr. Jessa Chokedumrongsuk said.

..., hospital director Dr. Vinit Pua-pradit said.

..., the doctor claimed.