

M-Eco

**Tag and Neighbour based Recommender
System for Medical Events**

**WWW2010 Conference
26/04/2010**

Karunakar Reddy & Peter Dolog


AALBORG UNIVERSITY Denmark



Summary

Introduction

Motivation

Context scenarios

Our Approach

Priliminary results

Future Works

Introduction

MEDWORM is assigned to organize and share medical resources through web.

- Tagging

-Medworm is a medical RSS feed provider as well as a search engine built on data collected from RSS feeds.

-Medwrom collects updates from over 6000 authoritative data sources.

- So far Medworm has approximately 2, 62,000 RSS feeds

Medworm is at <http://medworm.com/>

Tag and Neighbours

- Basically **Tag** is **personal** term assigned to a piece of information.

Tags...

- help to **describe** an item;
 - allow items to be found again by browsing or searching;
 - are chosen informally and personally;

- **Neighbours are neighbouring words to the words which are occurrences of a tag from the tag vector in particular document under investigation**

Motivation

When user does search with **tag** on MEDWORM, the search results are might not good due to the quality of tags and their amount is not sufficient.

Query -- **Swine flu**

and no. of records are 15097 as of today

low quality of tags

Neighbors:Vaccine

Tags: swine flu

No Tags

Pages: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

15097 records returned

[Old swine flu vaccine may still protect](#)   
MEMPHIS, April 23 (UPI) -- New evidence indicates the 1976 **swine flu** vaccine may help protect against 2009 H1N1, U.S. researchers say. (Source: Health News - UPI.com)
Source: [Health News - UPI.com](#) - April 24, 2010 Category: [Consumer Health News](#) Source Type: news

[UPDATE: 2009 H1N1 Flu International Situation Update](#)   
This report provides an update to the international situation as of April 23, 2010. The World Health Organization (WHO) continues to report updated 2009 H1N1 **flu**-associated laboratory-confirmed cases and deaths on its Web page. (Source: CDC Swine Flu Updates)
Source: [CDC Swine Flu Updates](#) - April 24, 2010 Category: [Infectious Diseases](#) Source Type: organizations

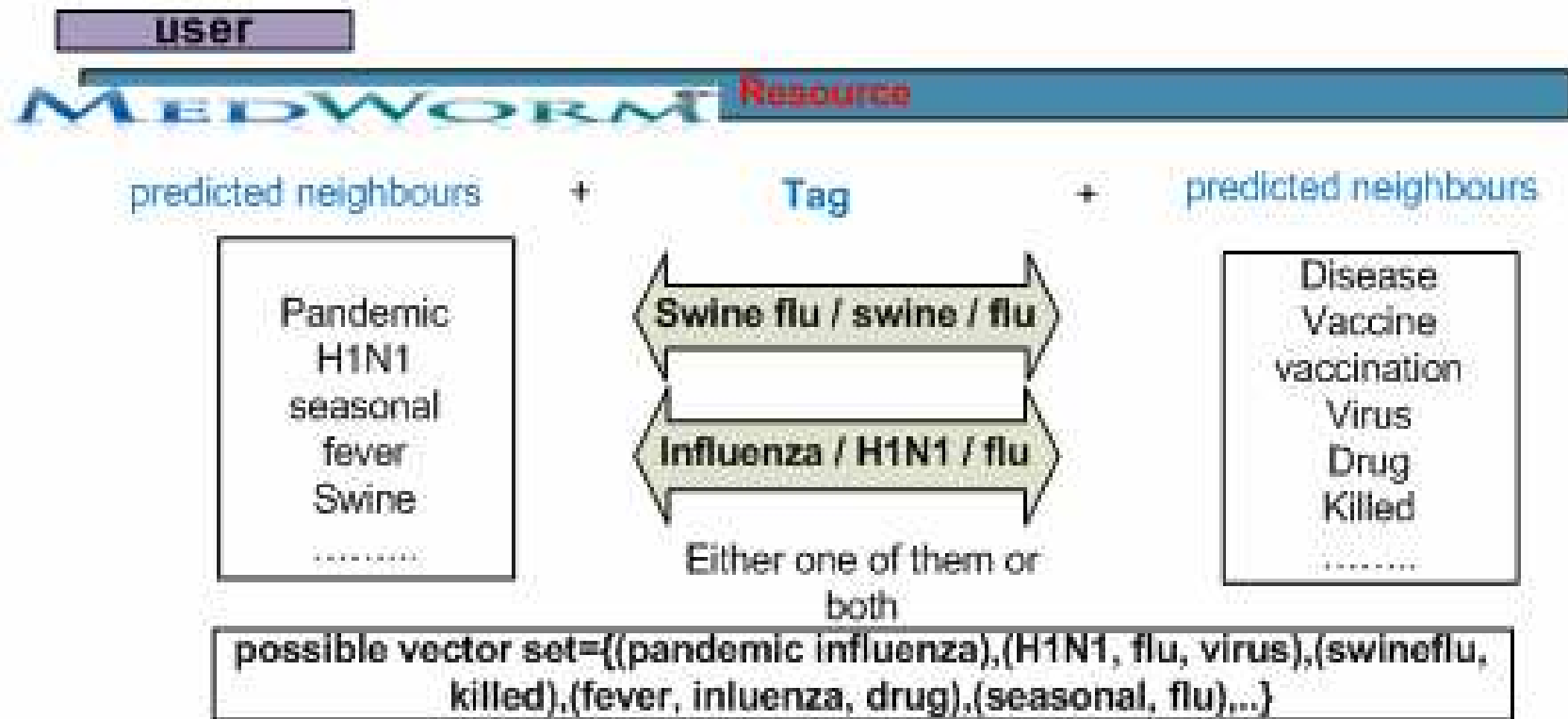
[UPDATE: Weekly FluView Map and Surveillance Report for Week Ending April 17, 2010](#)   
During week 15 (April 11-17, 2010), **influenza** activity decreased in the U.S. 37 (2.1%) specimens tested by U.S. World Health Organization (WHO) and National Respiratory and Enteric Virus Surveillance System (NREVSS) collaborating laboratories and reported to CDC/Influenza Division were positive for **influenza**. (Source: CDC Swine Flu Updates)
Source: [CDC Swine Flu Updates](#) - April 23, 2010 Category: [Infectious Diseases](#) Source Type: organizations

[Urgent Update For GPs Regarding Seasonal Influenza Program For Children](#)   
Due to a spike in the numbers of young children in Western Australia experiencing fever and convulsions following seasonal **flu** vaccinations, Australia's Chief Medical Officer, Professor Jim Bishop, is advising all GPs and immunisation providers to stop giving seasonal **flu** vaccine to children five years and under until a cause is established. Professor Bishop confirmed that this is a precautionary measure while the matter is being urgently investigated by health experts and the Therapeutic Goods Administration (TGA)...
(Source: Health News from Medical News Today)
Source: [Health News from Medical News Today](#) - April 23, 2010 Category: [Consumer Health News](#) Tags: [swine flu](#) Source Type: news

Motivation cont..

However

If you consider **neighbours**, the results could be re-ranked as of user's interests.



Context scenarios

Factors

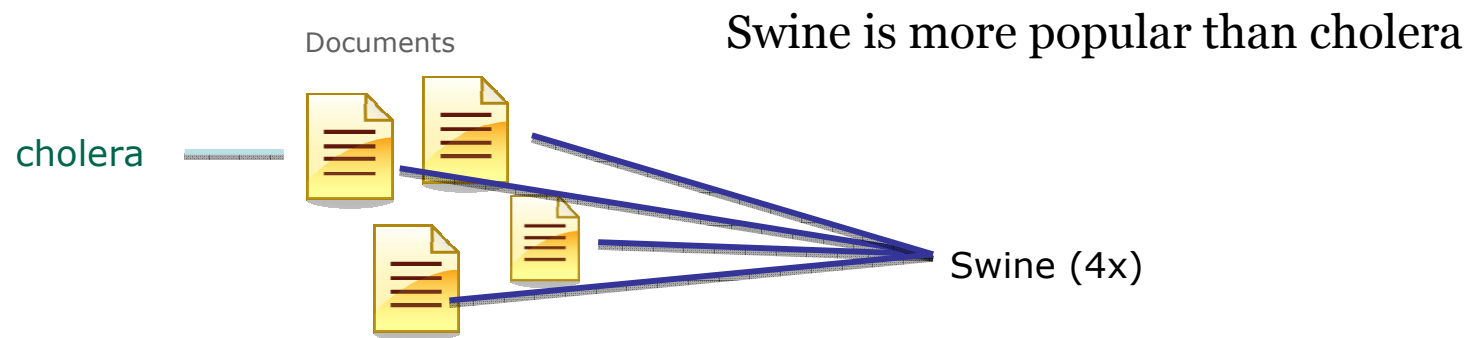
Factor	Data for recommended algorithm
1	Tag popularity
2	Tag Representativeness
3	Tag Similarity(Cosine, semantic similarity)
4	Tag Neighborhood

Aim:

Retrieve most relevant information while doing Tag based search in Medical domain
and
Present the top most relevant documents first.

1. Tag Popularity

No. of occurrences of one tag per total of resources available.



Tag popularity shows in **MEDWORM** with a bigger font size, bold letters and thicker colour.

Tag clouds

'flu . 'free . 'green'? . 'health . 'medicine . *articles . *weight loss
- education - ethics - events - feedback - general - guest -
humour - teaching - technology - tips - training - 2009 - 2010 - abbott -
abstinence - abuse - academic - access - acid - act - action - activism - activity - acupuncture -
acupuncture - add - add new tag - addiction - addiction recovery - adhd - administration - adult - adults
- advertising - advocacy - affairs - age - aging - agriculture - aid - aids - aids, -
alcohol - alcoholics - alcoholics anonymous - alcoholics-anonymous - alcoholism - alli - ambulance - america -
america' - american - americans - amp - analysis - andrew - anger - announcements - anonymous -
anti - antidepressants - anxiety - apnea - apnea - apple - apple ipad - applications - arra - art - arthritis -
articles - artificial - assessment - association - attachment - atypical - audit - autism - autism, -
awards - awareness - baby - back - bad - balance - banking - barack - barack
obama - based - behavior - behavioral - behaviour - belief - beliefs - benefits - benefits, - bextra - big
- bill - bill; - bioinformatics - biology - biotech - biotechnology - bipolar - bipolar disorder - birth -
birth; - black - blog - blog. - bloggers - blog
sugar - blood-sugar - bob - body - book - book review - book-re
brain - brain' - breast - breast car

2. Tag Representativeness

-How much a tag can represent a document it belongs

Swine and **Vaccine** are tags

...and...

are the most frequent terms for Document X

Flue related topic

Tf	Term
9	Swine
6	Vaccine
2	Pandemic

Tags: **swine, vaccine, pandemic, H1N1**

The screenshot shows a web page titled "H1N1 Flu" with a sidebar on the left containing a navigation menu. The main content area has a heading "2009 H1N1 Flu ('Swine Flu') and You" and a sub-heading "What is 2009 H1N1 (swine flu)?". The text describes the virus as a new influenza virus first detected in the US in April 2009. A table of contents on the right lists various topics. Annotations include arrows pointing from the table above to the words "Swine", "Vaccine", and "Pandemic" in the text, and a label "Flue related topic" pointing to the text.

Documents X

3. Tag Similarity

Tag Similarity: {Cosine Similarity, Semantic}

Factor	Data for similarity calculus	Data for cosine similarity
3.	Similarity between tags	Cosine Similarity

Query: swine flu disease

Tags: swine flu, disease, Cholera

Cosine similarity between query set (swine, flu, swine flu, swine flu disease, flu disease, swine disease, disease) and tag vector (swine flu, disease, Cholera, ..)
Swine flu is in query and in tags then score should be high

Tags: swine flu, pandemic, cholera

Tags: H1N1 flu, killed, aids

Swine flu and H1N1 flu are similar tags

4. Neighbour Semantic Similarity

By considering the **neighbor**, the result set could be changed.

30508 records returned

[Swine Influenza and You](#) [email this article to a colleague.](#) [_](#)
[article.](#)

[Save this article to My Clippings.](#) [_](#)

[1 Comment](#) [Discuss or comment on this](#)

What is **swine flu**? Are there human infections with **swine flu** in the U.S.? Is this **swine flu virus** contagious?

What are the signs and symptoms of **swine flu** in people? How serious is **swine flu** infection?...

Source: [CDC Swine Flu Updates](#) - April 24, 2009 Category: [Infectious Diseases](#) Source Type: organizations

Tags: flu, H1N1, killed, birds

neighbour

swine

tag

flu

neighbour

virus

Tags: flu, H1N1, killed, birds

Neighbours: symptoms, swine, virus, infection

Cosine Similarity between tags and neighbors

Our Approach

Hybrid similarity score of the proposed recommender algorithm is

$$HS_{(D_i, D_{ii})} = [(DS_{D_i} \times DS_{D_{ii}}) \times (TS_{(D_i, D_{ii})})] \times NS_{(D_i, D_{ii})}$$

D_i and D_{ii} are a particular documents from a set of documents D .

DS is the document score = $\sum_{i=0}^n \text{Popularity}(\text{Tag } i) \times \sum_{i=0}^n \text{Representativeness}(\text{Tag } i)$

TS is a function for measuring the tag similarity

$\text{SemSim}(s, t) = \text{MDSim}(s, t)$, where s and t are particular tags of set of tags T

$$\text{CosSim}(Q_{D_i}, W_{D_{ii}}) = \frac{\bar{W}(Q_{D_i}, W_{D_{ii}})}{\sqrt{\sum_{t \in T} \bar{W}(Q_{D_i}, W_{D_{ii}})^2}}$$

where T is the set of tags, \bar{W} is the set of words of current page

NS is the neighbour semantic similarity between the tag vectors and neighbours

$$\cos(w(\bar{t}_i), w(\bar{t}_j)) = \frac{w(\bar{t}_i) \cdot w(\bar{t}_j)}{\|w(\bar{t}_i)\| \cdot \|w(\bar{t}_j)\|}$$

Where \bar{t}_i is the tag vector and \bar{t}_j is the neighbourhood vector

Preliminary User interface design

Platform Prototype - Windows Internet Explorer

http://crittechportal1.jrc.it/m-eco/EarlyWarningPortal/PlatformPrototype/tabid/69/Default.aspx

Portal of the Medical EcoSystem working group on risk management and communication

M-Eco Medical EcoSystem

Wednesday, April 21, 2010 | Test User1 | Logout

Home Early Warning Portal Communicators Network Portal **Your profile** Enter keywords

Change profile Feedback Alerts

News Moderation

Dr Chris Kingswood (TSA Head of Research) fielded the following questions at the AGM in November last year: Q: Should someone with TSC have the swine flu vaccine? A: Yes, the benefits of having the vaccine far outweigh any risks. For younger children, vaccines may be given in two doses which are separated by a number of weeks. Source: Medworm Location: Feedback: [+]

Infectious diseases are responsible for millions of deaths worldwide every year. News on outbreaks of infectious diseases like Ebola, Marburg, Lassa fever, which have high death rates, are often the focus of exaggerated press scare stories. Source: Twitter Location: Feedback: [+]

Health agencies learned that information, improvisation are crucial to stop diseases before they spread too far (Source: USATODAY.com Health) Source: Medworm Location: Feedback: [+]

Getting treatment quickly was key to avoiding life-threatening illness, CDC study found Source: HealthDay Related MedlinePlus Pages: H1N1 Flu (Swine Flu), Infections and Pregnancy (Source: MedlinePlus Health News) Source: New paper Location: Feedback: [+]

From Canadian Press DataFile (April 20, 2010) GENEVA -- Drug maker Novartis AG said Tuesday that strong sales particularly of swine flu vaccine helped increase net profit by 49 per cent to \$2.95 billion during the first quarter. Source: Medworm Location: Feedback: [+]

ATLANTA (AP) — The most complete analysis of swine flu deaths among pregnant women last year indicates that quick treatment with Tamiflu or a similar medicine saved the lives of many Source: Medworm Location: Feedback: [+]

Infectious diseases are responsible for millions of deaths worldwide every year. News on outbreaks of infectious diseases like Ebola, Marburg, Lassa fever, which have high death rates, are often the focus of exaggerated press scare stories. Source: Twitter Location: Feedback: [+]

Getting treatment quickly was key to avoiding life-threatening illness, CDC study found Source: HealthDay Related MedlinePlus Pages: H1N1 Flu (Swine Flu), Infections and Pregnancy (Source: MedlinePlus Health News) Source: New paper Location: Feedback: [+]

Add threat

Name

Description

RSS Feed

Create

Threats

- swine flu Sunday, March 28, 2010 10:10 PM disease
- Polioyelitis Monday, March 15, 2010 4:15 PM Polioyelitis
- DengueFever Monday, March 15, 2010 3:40 PM DengueFever
- CholeraInfection Monday, March 15, 2010 3:40 PM CholeraInfection
- Chikungunya Monday, March 15, 2010 3:40 PM Chikungunya
- Avianflu Monday, March 15, 2010 3:40 PM

Done Internet | Protected Mode: On 100%

Future Works

- Evaluation with classifiers
- Apply sentiment analysis

Thank you!

Questions?

Email:

kreddy@cs.aau.dk