Semantic Wiki for Visualization of Social Media Analysis

D. Reininger, D. Ihrie & P.R. Bullard

Semantic Technology for Intelligence, Defense, and Security (STIDS) 2010

© 2010 Semandex Networks Inc. All Rights Reserved.
Outline

- Social media analysis support to operations
- A semantic approach to social media analysis
- Implementation and results
- Conclusions and future directions
Social media analysis support to operations

- Target audience assessment
  - Model and extract significance from the social media interactions of persons and organizations of interest.
  - Link entities contained in report, open source articles and other sources encountered by users.

- Tailor message based on accessibility

The INFORM program created a web-based application with these capabilities to support Information Operations.

- Review data from open source media, combined with data from additional sources, to support assessment of social groups, subgroups, and individuals within a population.

- Understand how information influences perception and public opinion of populations and how these influences are transmitted and communicated (e.g., media, peer groups, one-to-one, etc.)
A semantic approach to social media analysis

- Define a core semantic schema of concepts and relations.
- Build a semantic graph from heterogeneous data sources.
- Use a web-based wiki for data capture, analysis and display.
Task: Target audience analysis in a specific area

City: Nuevo Laredo (City Area)

Name: Nuevo Laredo

Location Of: Kidnapping of Yvette Martinez

Clear Polygon Location

Map data ©2010 Google, INEGI - Terms of Use

Current Activity

Today
- Link added to: Dallas Morning News Death in Nuevo Laredo.txt
- Link added to: STRATFOR Operation Safe Mexico in Nuevo Laredo.txt
- Link added to: DoS on border crime.txt
- Link added to: DoS on border crime.txt
- Link added to: DoS on border crime.txt

Governed By
- none

People Associated with this Place

(Associated With) Jorge Castillo

add new link

add new link
Social analysis: From individuals to population groups

Person: Jorge Castillo

Name
Jorge Tierno
Jorge Castillo, Transportation Company Executive

email Address
damian.castillo@gmail.com

chat id
DCinNL@jabber.org

blog id
MexicoTrucker

Current Activity

People
(Associated With) Daniel Ramirez
(Associated With) Jorge Viales, CODEIN Chairman
(Loyal To) Jose Gutierrez, Rector of UAT
(Has Informed) Juan Garcia, El Manana Editor
(Owes Money To) Ramirez Vega

Places or Facilities
(Lives At) Jorge Castillo's Residence

add new link

Events
(IS Participant In) DemoNeedCensus
(IS Participant In) DemoSentimentCensus

add new link

Workflows tagged "Person"

Documents Possibly Mentioned In
A view of one type of social interaction: mobile communications

Person: Jorge Castillo

Phone Call Summary

Person: Juan Garcia, El Manana Editor

6 calls

MSISDN: 631-717-2098

4 calls

Person: Jose Gutierrez, Rector of UAT

2 calls

MSISDN: 631-287-8273

1 call

© 2010 Semandex Networks Inc. All Rights Reserved.
We can export a summary of these interactions in graph format.
We can tailor how we view the data; for example, by frequency.
Temporal view identifies changes in trends of social media use

Norm: Active on Tuesdays; inactive on Fridays

Deviation: Activity on Friday, and at different times of day
We can aggregate data: People are members of an Organization

Organization: CODEIN (Comité de Desarrollo Industrial Economic de Nuevo Laredo)

People:
- Cesar Castillo
- Crespin Castro
- Crespin Hernandez
- Crespin Pena
- Cruz Perez
- Daniel Cruz
- Daniel Ramirez
- DeMario Ramos
- Doroteo Diaz
- Doroteo Reyes
- Eber Rios
- Elbano Flores
- Fieraz Mendez

Add CODEIN (Comité de Desarrollo Industrial Economic de Nuevo Laredo) to my WatchList

BWG Status: ON White List [change] [x]

Organizations:
- Does Business With: City Government of Nuevo Laredo (Government Organization)
- Hostile Towards: Sinaloa Cartel (Drug Traffickers)
- Member of: Nuevo Laredo residents 30 to 49, male, middle and upper income (Population)

add new link
## Characteristic Summary of Members of Organization

**CODEIN (Comité de Desarrollo Industrial Economic de Nuevo Laredo) Workflows**

**Show Characteristic Summary**

Total results: 52

<table>
<thead>
<tr>
<th>favoriteNewspaper</th>
<th>El Manana</th>
<th>34 72.34%</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Diario de Nuevo Laredo</td>
<td>13 27.66%</td>
<td></td>
</tr>
</tbody>
</table>

**religion**

<table>
<thead>
<tr>
<th>Christianity</th>
<th>47</th>
<th>90.385%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atheism</td>
<td>5</td>
<td>9.615%</td>
</tr>
</tbody>
</table>

**receiver**

<table>
<thead>
<tr>
<th>Jorge Castillo</th>
<th>11 78.571%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juan García, El Manana Editor</td>
<td>3 21.429%</td>
</tr>
</tbody>
</table>

**income**

<table>
<thead>
<tr>
<th>medium</th>
<th>29</th>
<th>58%</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>21</td>
<td>42%</td>
</tr>
</tbody>
</table>

**age**

<table>
<thead>
<tr>
<th>33</th>
<th>6</th>
<th>12%</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td>34</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>8%</td>
</tr>
<tr>
<td>49</td>
<td>3</td>
<td>6%</td>
</tr>
<tr>
<td>48</td>
<td>3</td>
<td>6%</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
<td>6%</td>
</tr>
<tr>
<td>41</td>
<td>3</td>
<td>6%</td>
</tr>
</tbody>
</table>

**favoriteRadio**

<table>
<thead>
<tr>
<th>XHNLQ</th>
<th>30</th>
<th>60%</th>
</tr>
</thead>
<tbody>
<tr>
<td>XHNLR</td>
<td>20</td>
<td>40%</td>
</tr>
</tbody>
</table>

**org**

| CODEIN (Comité de Desarrollo Industrial Economic de Nuevo Laredo) | 50 | 100% |

**favoriteTV**

<table>
<thead>
<tr>
<th>XHNLQ</th>
<th>30</th>
<th>60%</th>
</tr>
</thead>
<tbody>
<tr>
<td>XHNLR</td>
<td>20</td>
<td>40%</td>
</tr>
</tbody>
</table>

**originator**

| Juan García, El Manana Editor | 6 | 40% |
| Jorge Castillo               | 6 | 40% |
| Jorge Viales, CODEIN Chairman| 3 | 20% |

**race**

| Hispanic         | 40 | 80% |
| American Indian  | 5  | 10% |
| Caucasian        | 5  | 10% |

**Internet**

<table>
<thead>
<tr>
<th>yes</th>
<th>25</th>
<th>51.02%</th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>24</td>
<td>48.98%</td>
</tr>
</tbody>
</table>
Export geospatial views of sentiment of our target population
Compare characteristics among members of our target group

**Person: Jorge Castillo**

<table>
<thead>
<tr>
<th>Name</th>
<th>income</th>
<th>org</th>
<th>religion</th>
<th>favoriteNewspaper</th>
<th>favoriteRadio</th>
<th>favoriteTV</th>
<th>race</th>
<th>age</th>
<th>internet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daniel Ramirez (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Elbanco Flores (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Renau Castillo (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Crispin Castro (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Doroteo Diaz (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Crispin Pena (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Carlos Ortega (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Francisco Vasquez (Person)</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ernesto Gutierrez (Person)</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Target Audience Sub-group

Population: Faculty and students of NST

People

<table>
<thead>
<tr>
<th>Has Member</th>
<th>Name</th>
<th>[x]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Carlos Gonzales</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ignacio Ortega</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ignacio Salazar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lino Reyes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Luis-Fernando Rios</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Luz Rivera</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rafael Vasquez</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ramirez Vega</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ruben Cruz</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Santiago Diaz</td>
<td></td>
</tr>
</tbody>
</table>

Mentioned In

none

Organizations

(Member of) Subgroup: Nuevo Laredo residents 30 to 49, male, middle and upper income (Population)

Places

none

Facilities

add new link

Add Faculty and students of NST to my WatchList

Add Faculty and students of NST to a Black/White/Grey List
Analytic workflow showing neighborhoods & broadcast stations
Drill down detail for social media knowledge

![Diagram showing saved workflows and Google Map with location of XEGNK AM 1370kHz 500W Radio Mexicana (Broadcasting Station)]
Ontology Management and Graph Creation
Mexico City, May 30 (EFE) - Five members of Mexico's powerful Sinaloa drug cartel - one of them a public prosecutor - received prison sentences ranging from 23 to 32 years, the attorney general's office said Wednesday. The AG's office said in a statement that the defendants were sentenced by a court in Toluca, the capital of the central state of Mexico. They were convicted of organized crime, attempted homicide and possession of military-grade weapons. Two of the men - Jose Martinez Benavides and Juan Calixto Ramos Vazquez - were sentenced to 24 years in prison and ordered to pay a fine of roughly 20,000 pesos ($1,800), while two others, Jose Silverio Martinez Gonzalez and Erick Rolando Perez Perez, were given 30 years and 32 years in prison, respectively. In the same case, Juan Manuel Munoz Morales, a former chief prosecutor in the northern border city of Nuevo Laredo, received a prison sentence of 23 years and was ordered to pay a 150,000-peso ($13,500) fine. He was also removed from his post and banned from working as a public servant for nine years. Munoz was accused of freeing the
Recommendation Engine suggests matches
Machine-assisted semantic annotation of video

- Direct linking of video, images and text into the information base

Extract entities and relationships from text and video annotations

20051228: White Pontiac W70KN5 TX registered to Yvette Martinez recovered in Nuevo Laredo MX

Source: OSINT Fox America's Most Wanted
Checking existing concept sub-types

Browse Facility Pages
This is a list of: All Facility pages

Add filters

Results

testplace (Facility)
Suspected security post, Sinaloa Cartel (Facility)
Suspected safehouse, Sinaloa Cartel (Facility)
John Jay College of Criminal Justice (Facility)
Paris, Las Vegas (Facility)
Noor Mosque (Facility)
Parking Lot 441 (Facility)
parking lot (Facility)
parking lot at 27.4934,-99.5173 (Facility)
Newspaper: Primera Hora (Facility)
Newspaper: El Manana (Facility)
Newspaper: El Diario de Nuevo Laredo (Facility)
KXOF TV Ch39 Fox (Broadcasting Station)
XHLAR TV Ch57 Televisa (Broadcasting Station)
XHJNA TV Ch21 TV Azteca 13 (Broadcasting Station)
XHNL X 94.9 MHz 20W University Radio (UAT) (Broadcasting Station)
XHNLO FM 97.1 MHz 3KW Multimedios Radio (Broadcasting Station)
XHTLN FM 94.1 MHz 2.85KWImagen (Broadcasting Station)
XHNPE FM 91.3 MHz 50KW Stereo 91.3 (Broadcasting Station)
XENLT AM 1000kHz 500W Radio Formula (Broadcasting Station)
Creating new concept sub-types

Browse Facility Pages
This is a list of: All Facility pages
Add filters

Results
- testplace (Facility)
- Suspected security post, Sinaloa Cartel (Facility)
- Suspected safehouse, Sinaloa Cartel (Facility)
- John Jay College of Criminal Justice (Facility)
- Paris, Las Vegas (Facility)
- Noor Mosque (Facility)
- Parking Lot 441 (Facility)
- parking lot (Facility)
- parking lot at 27.4934, 99.5173 (Facility)
- Newspaper: Primera Hora (Facility)
- Newspaper: El Manana (Facility)
- Newspaper: El Diario de Nuevo Laredo (Facility)
- KXOF TV Ch39 Fox (Broadcasting Station)
- XHLAR TV Ch57 Televisa (Broadcasting Station)
- XHLNA TV Ch21 TV Azteca 13 (Broadcasting Station)
- XHNLR FM 104.9MHz 20W University Radio (UAT) (Broadcasting Station)
- XHNLO FM 97.1MHz 3KW Multimedios Radio (Broadcasting Station)
- XHTLN FM 94.1MHz 2.85KW Imagen (Broadcasting Station)
- XHNOE FM 91.3MHz 50KW Stereo 91.3 (Broadcasting Station)
An instance of a new concept sub-type
Results

- **Schema of 10 basic concepts**
  - Person, Organization, Place, Event
  - File, Characteristic, Watchlist, Time, Location

- **262 sub-types created during modeling phase**
  - Facility is a type of Place; Broadcasting station is a type of Facility
  - Communication event is a type of Event; Tweet is a type of communication event

- **Only 63 types (24%) actually used during the practical application phase**
  - 204 files in four formats
  - 4310 concept instances in a 102 MB semantic database
  - Semantic graph represented interrelations of 585 events and 196 persons.
Conclusions and future directions

- The INFORM project integrated disparate social media data sources using automated and machine assisted techniques that promote data validity and collaborative accessibility.

- When presented with a myriad of choices in modeling data, users often use commonly recognized entity types:
  - The set of commonly recognized entity types is not known a priori.

- Next generation approach seeks to combine the the structuring power of semantics and the complete extensibility of tagging.