Semantic Wiki for Visualization of Social Media Analysis

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

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

Semandex
CONNECTED INTELLIGENCE.

© 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

© 2010 Semandex Networks Inc. All Rights Reserved.

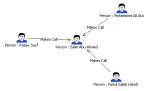
Social media analysis support to operations

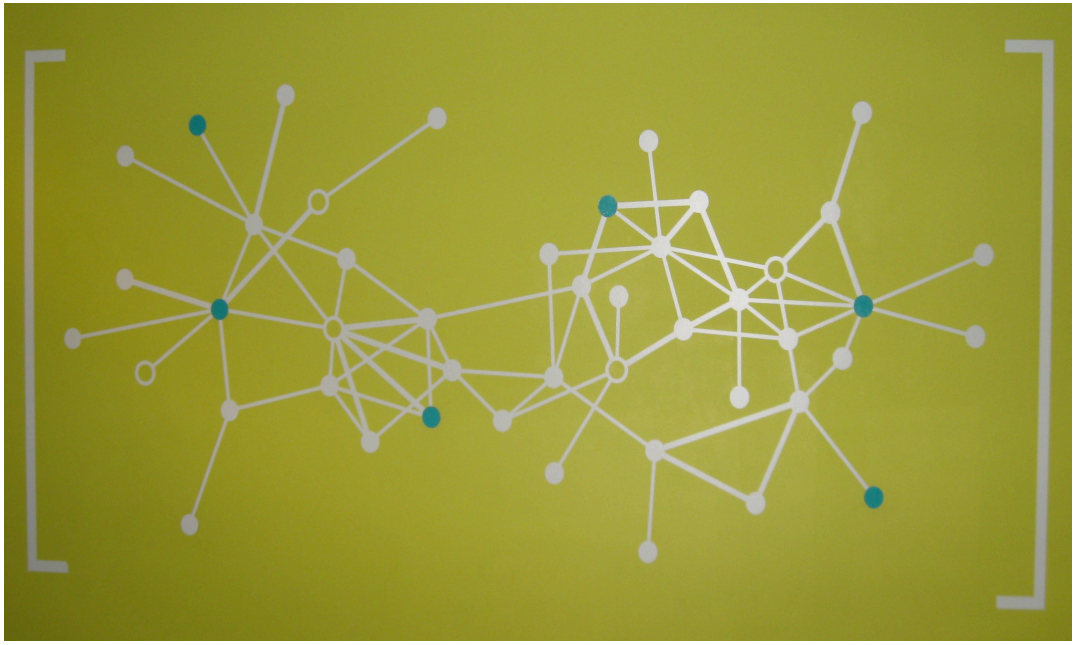
- Target audience assessment
  - Model and extract significance from the social media interactions of persons and organizations of interest.
  - Link entities contained in report, open source articles and other sources encountered by users.
- Tailor message based on accessibility
- The INFORM program 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.)

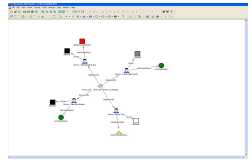
© 2010 Semandex Networks Inc. All Rights Reserved.

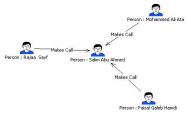
A semantic approach to social media analysis

- 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.









© 2010 Semandex Networks Inc. All Rights Reserved.

Semantic Wiki Homepage

TANGO

Search | All Tango | enter a page name | Search | Browse Pages

Homepage: Dan Reininger

edit your profile | view users | upload a file | help

Run Workflow

Home | Data Loader | TangoApps

My Starred Pages

- CODEIN (Comite de Desarollo Industrial Economic de Nuevo Laredo) (Organization) [x]
- DemoSentimentCensus (Census) [x]
- Group: Nuevo Laredo residents under 50 (Population) [x]
- Jorge Castillo (Person) [x]
- Nuevo Laredo (City) [x]
- Universidad Autonoma de Tamulipas (UAT) (Facility) [x]

Folders

- Nuevo Laredo items of interest
- Broadcast stations
- Neighborhoods in Nuevo Laredo

My Watch List

- CODEIN (Comite de Desarollo Industrial Economic de Nuevo Laredo) (Organization) [x]
- Daniel Ramirez (Person) [x]

Import Webpage

Enter the URL of the webpage to import into Tango.

import

Textploitation Bookmarklet

© 2010 Semandex Networks Inc. All Rights Reserved.

Task: Target audience analysis in a specific area

TANGO

Search | All Tango | enter a page name | Search | Browse Pages

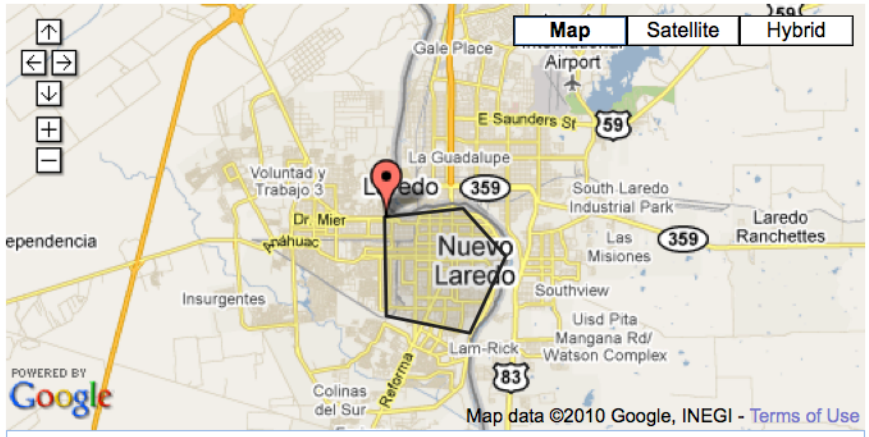
Starred page | Export / Save | Folders... | Delete this page

City: Nuevo Laredo (City Area)

Run Workflow

Place | Location | Event Plot | Similar | Notes

Name
Nuevo Laredo



Clear Polygon Location

Location Of

Kidnapping of Yvette Martinez [r] [x]

Current Activity

Today

Link added to: Dallas Morning News Death in Nuevo Laredo.txt — Mon, Oct 25, 2010

Link added to: STRATFOR Operation Safe Mexico in Nuevo Laredo.txt — Mon, Oct 25, 2010

Link added to: DoS on border crime.txt — Mon, Oct 25, 2010

Link added to: DoS on border crime.txt — Mon, Oct 25, 2010

Link added to: DoS on border crime.txt — Mon, Oct 25, 2010

Governed By

none

add new link

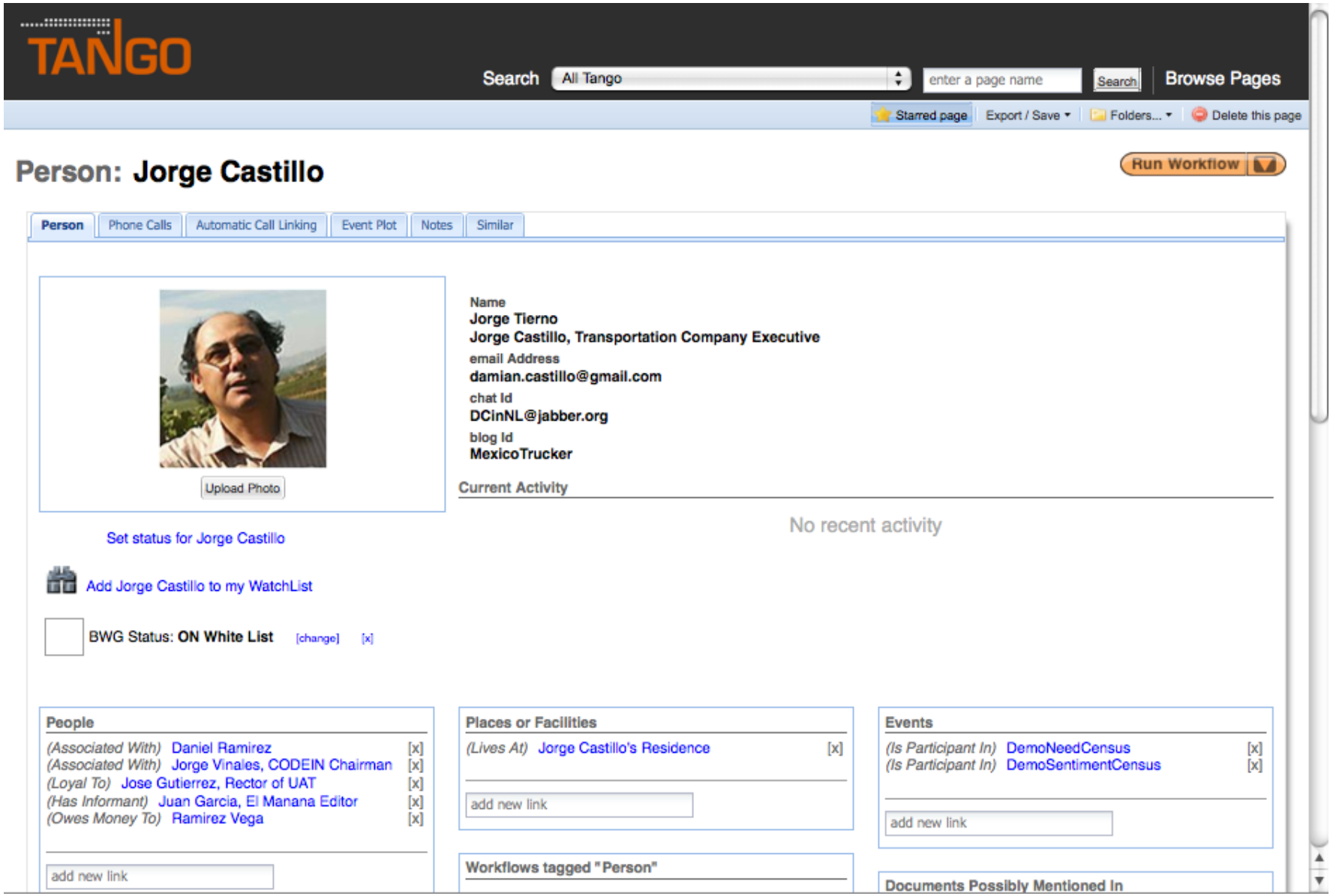
People Associated with this Place

(Associated With) Jorge Castillo [x]

add new link

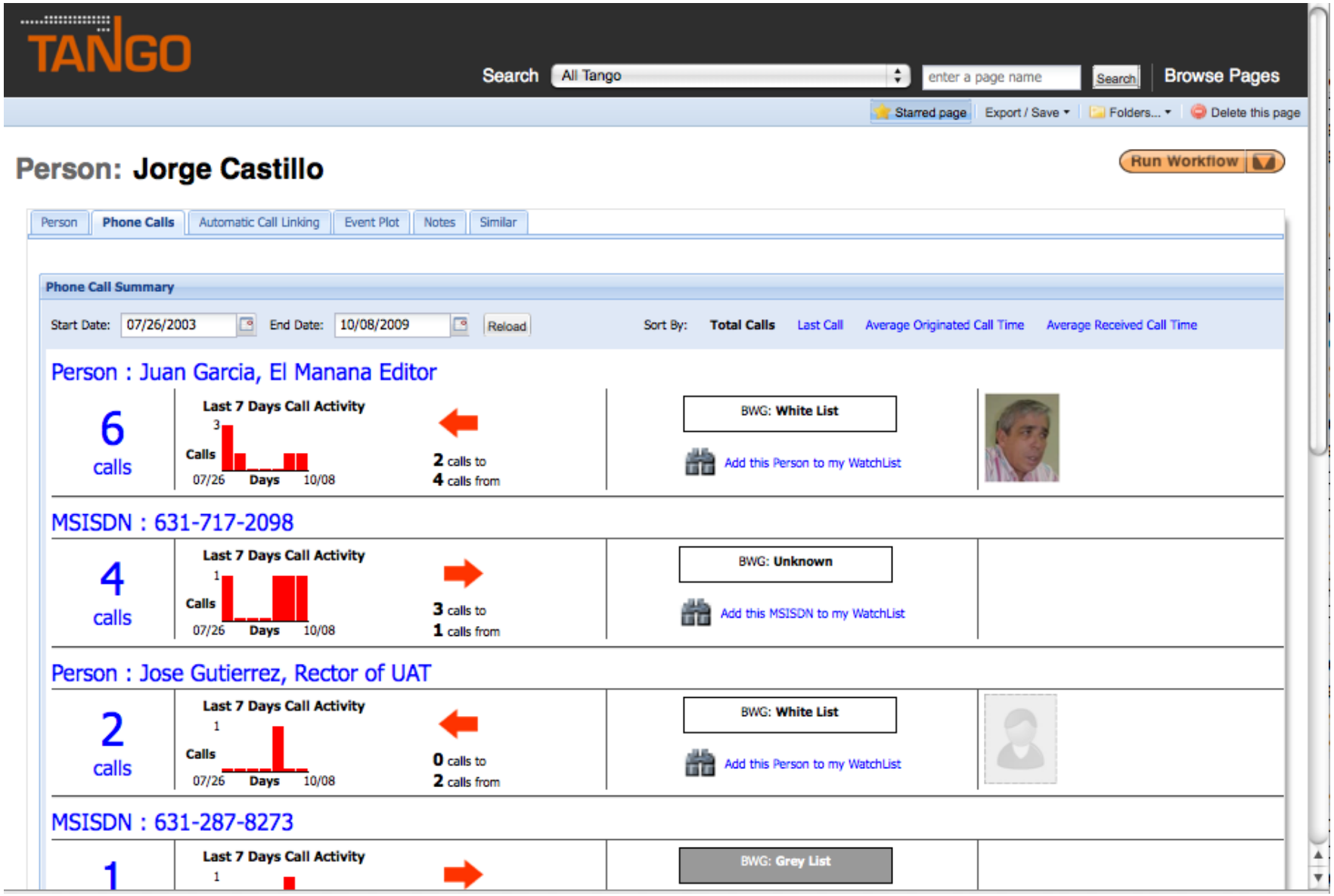
© 2010 Semandex Networks Inc. All Rights Reserved.

Social analysis: From individuals to population groups



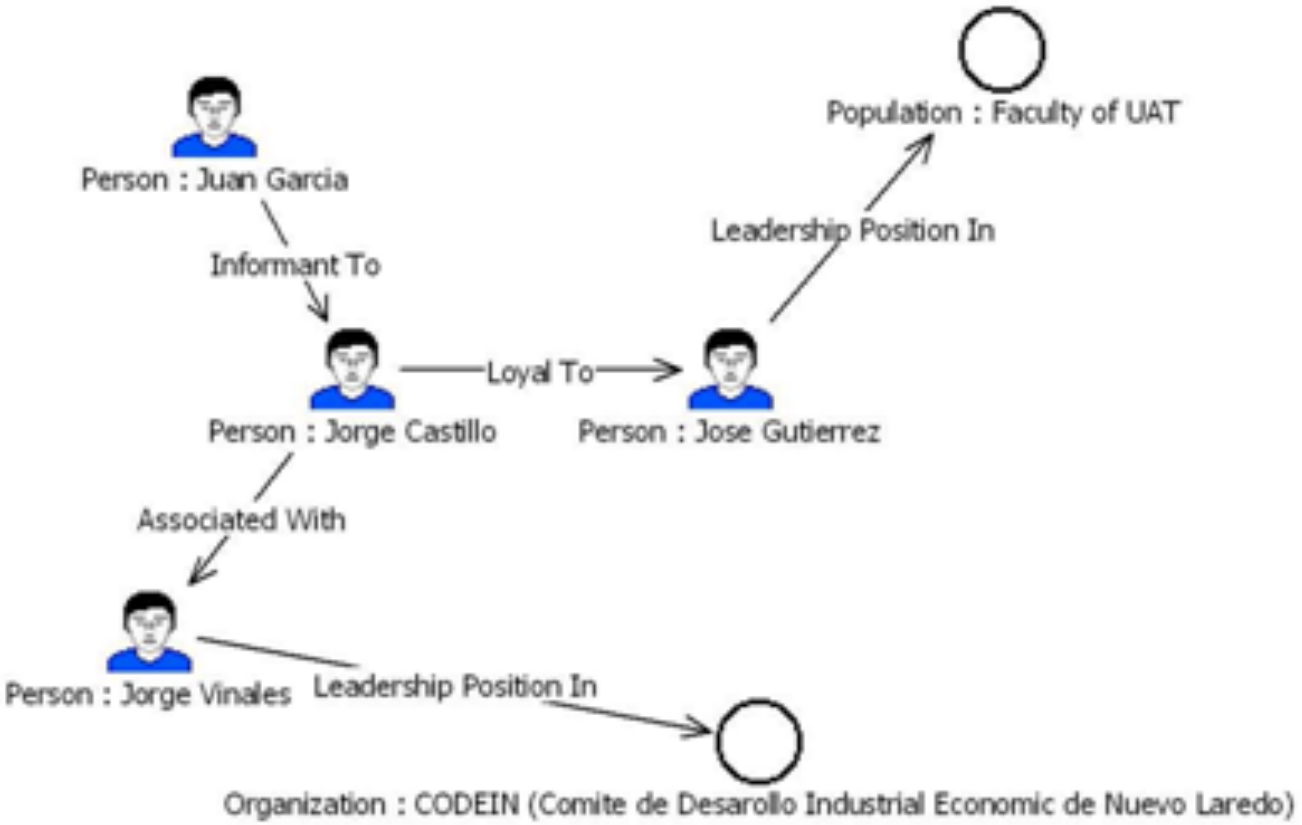
© 2010 Semandex Networks Inc. All Rights Reserved.

A view of one type of social interaction: mobile communications



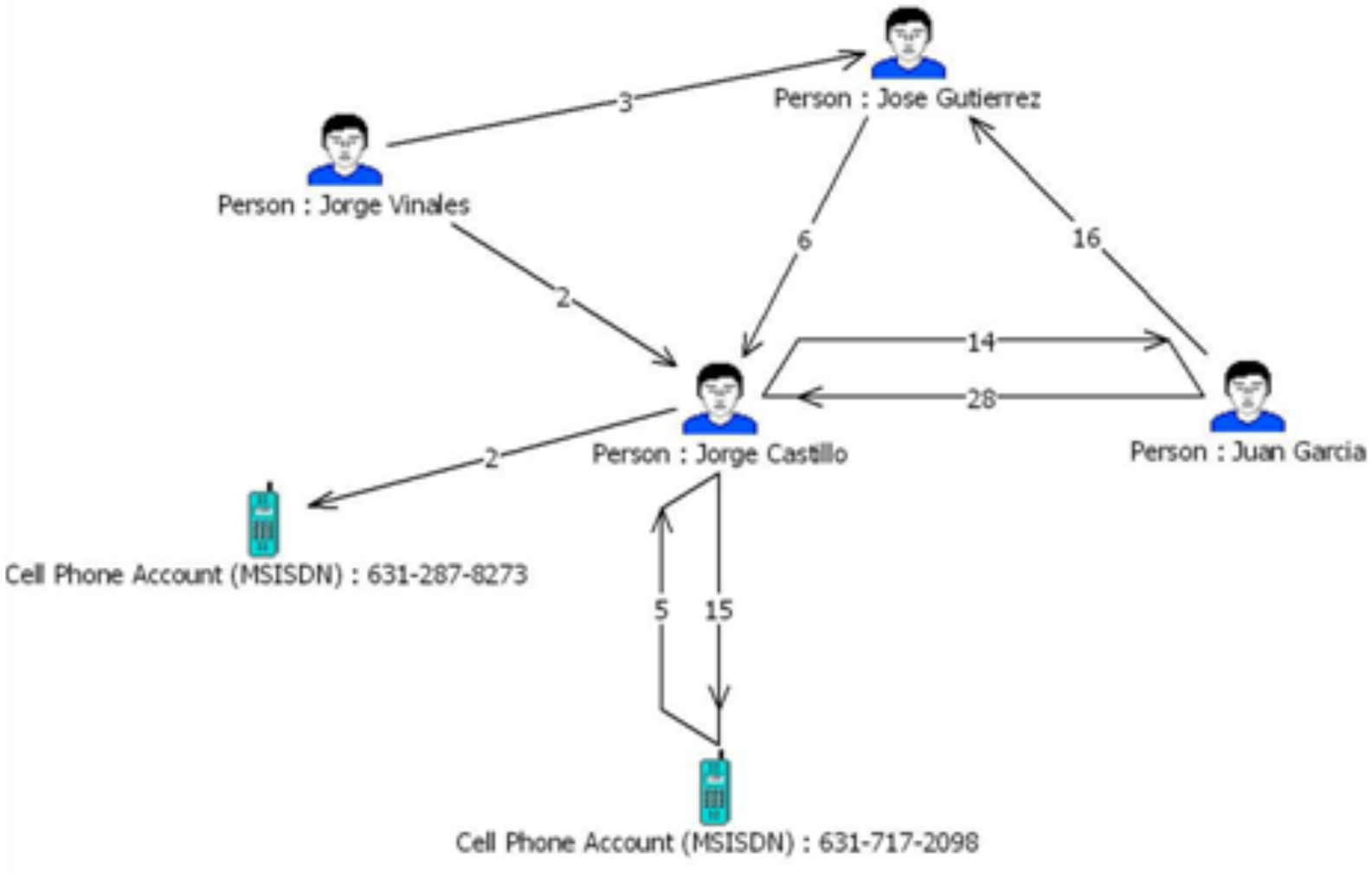
© 2010 Semandex Networks Inc. All Rights Reserved.

We can export a summary of these interactions in graph format



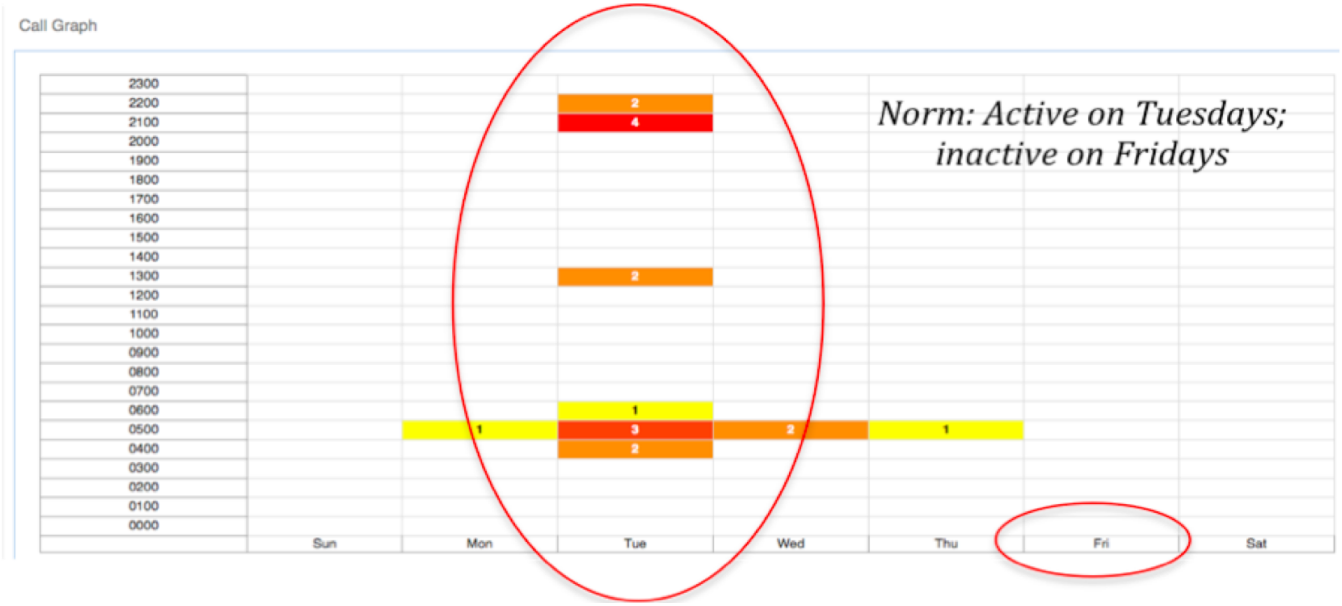
© 2010 Semandex Networks Inc. All Rights Reserved.

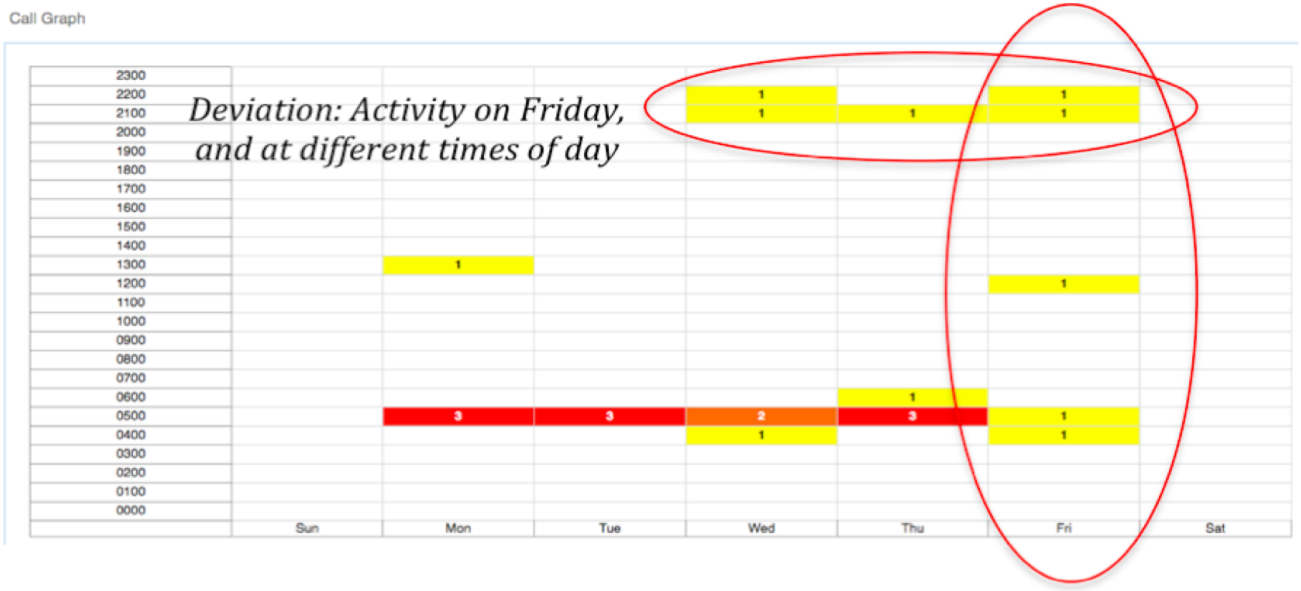
We can tailor how we view the data; for example, by frequency



© 2010 Semandex Networks Inc. All Rights Reserved.

Temporal view identifies changes in trends of social media use



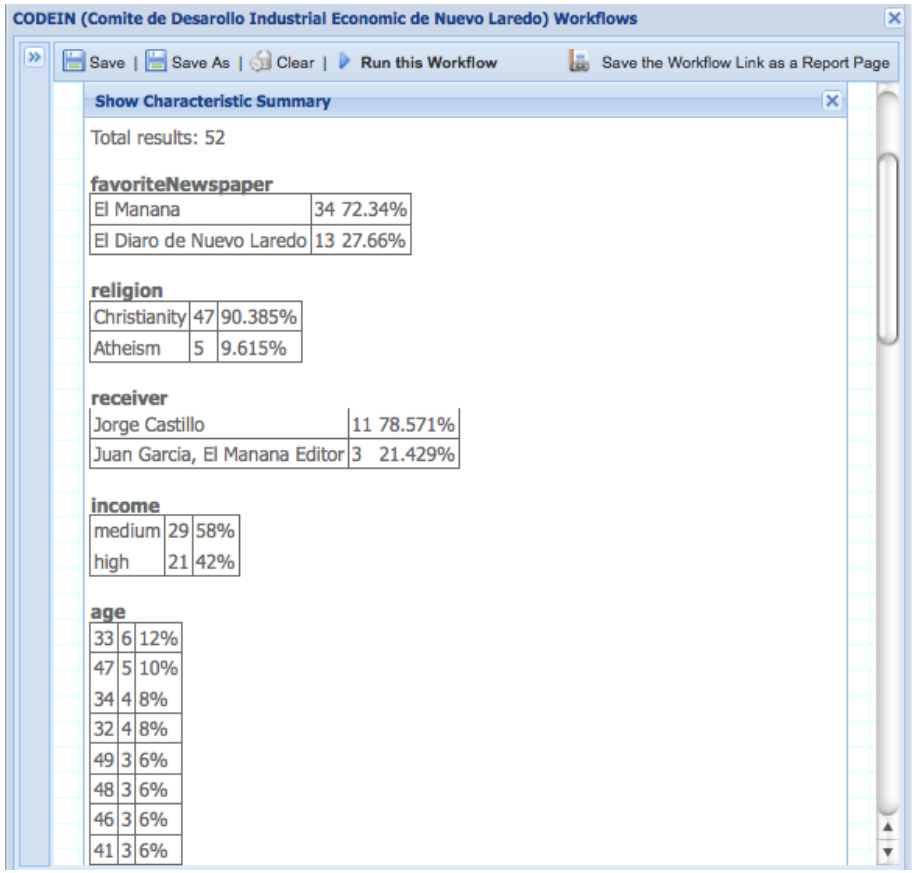


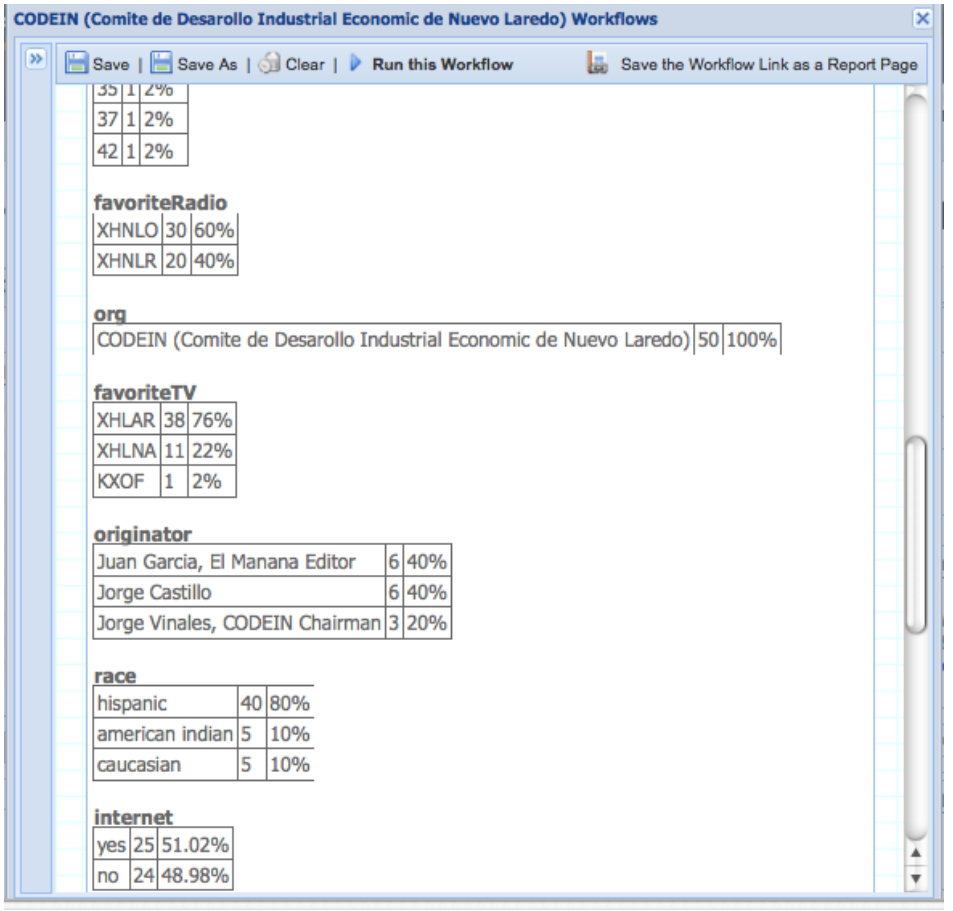
© 2010 Semandex Networks Inc. All Rights Reserved.

We can aggregate data: People are members of an Organization

© 2010 Semandex Networks Inc. All Rights Reserved.

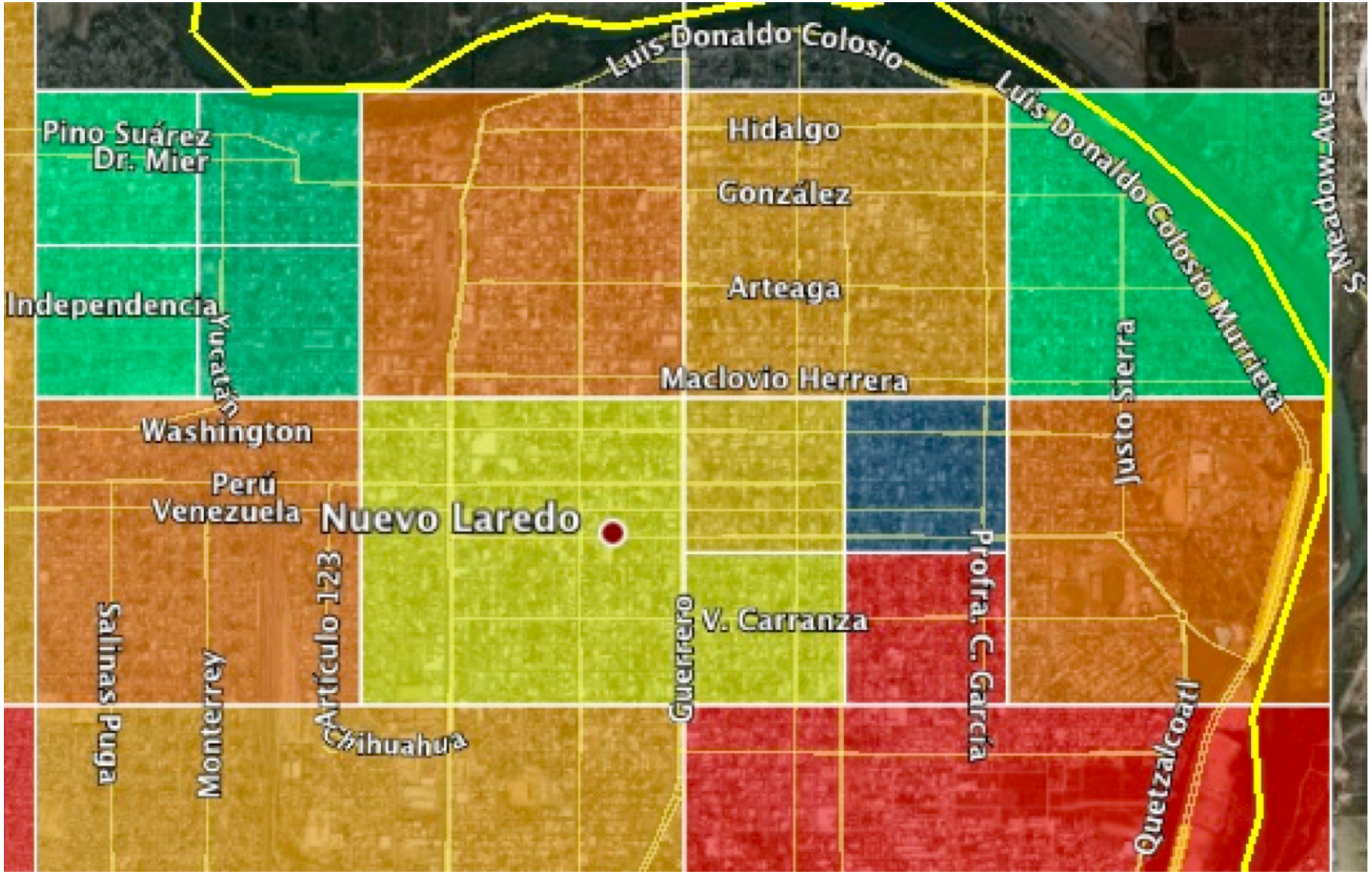
Characteristic Summary of Members of Organization





© 2010 Semandex Networks Inc. All Rights Reserved.

Export geospatial views of sentiment of our target population

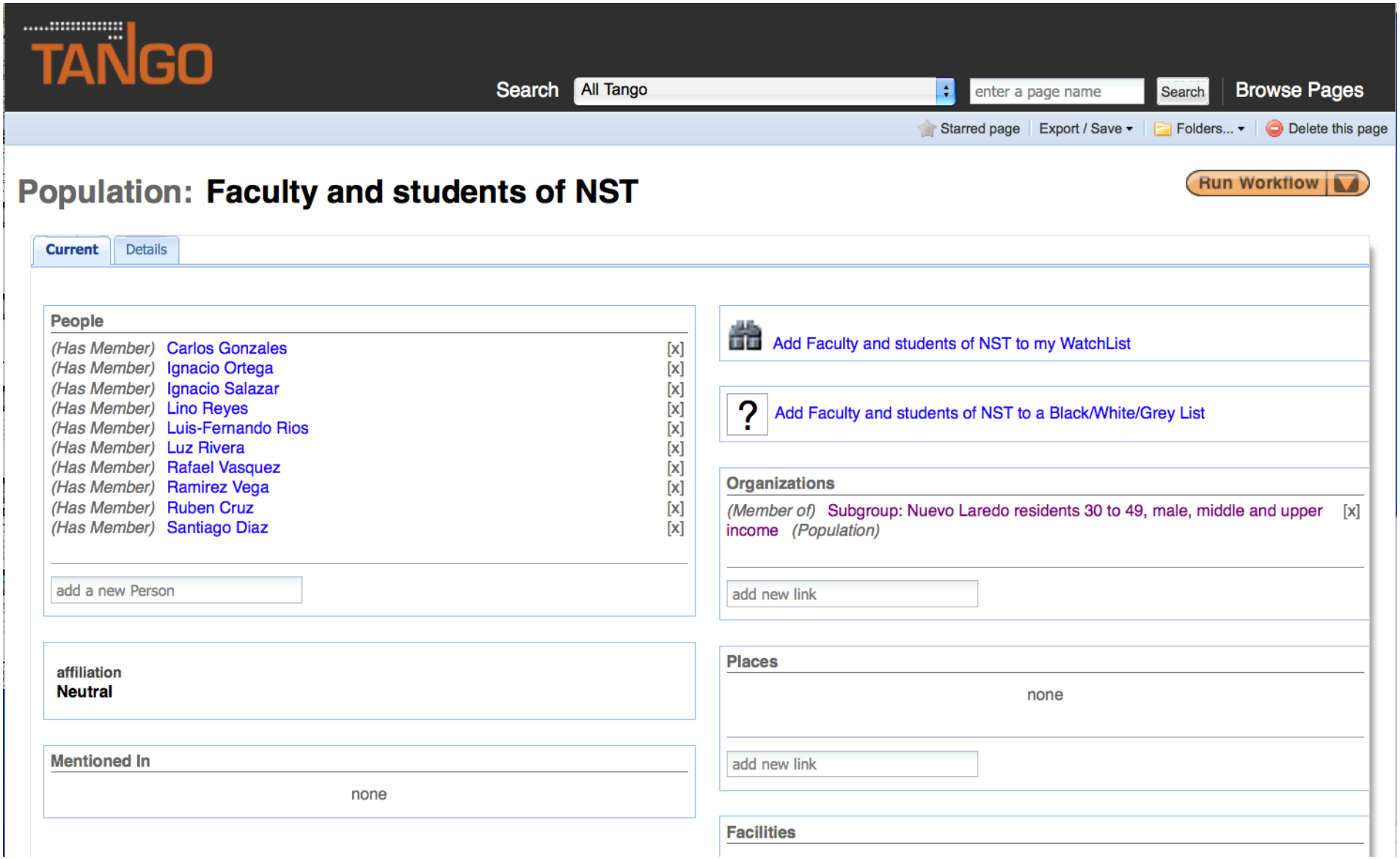


© 2010 Semandex Networks Inc. All Rights Reserved.

Compare characteristics among members of our target group

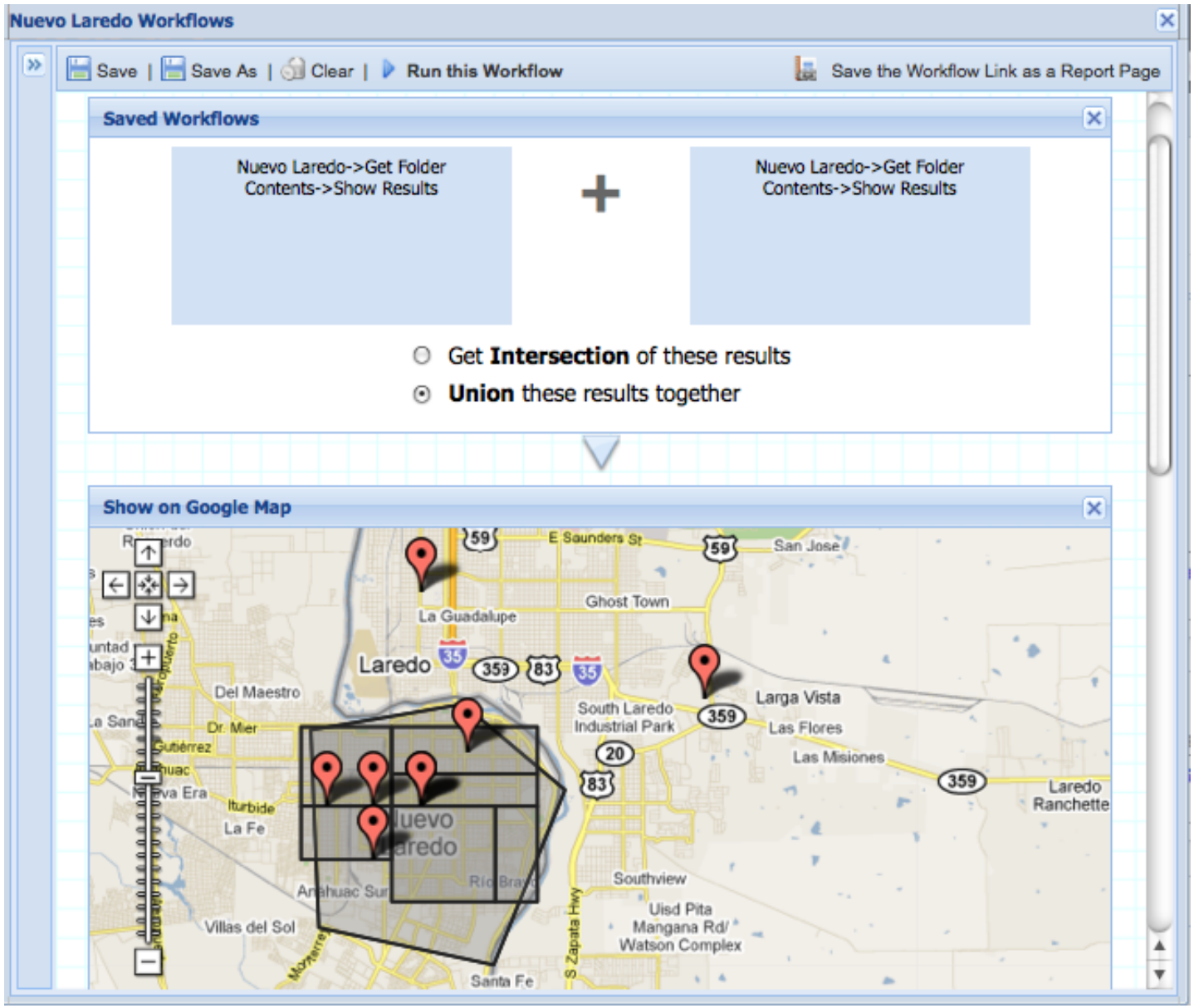
© 2010 Semandex Networks Inc. All Rights Reserved.

Target Audience Sub-group



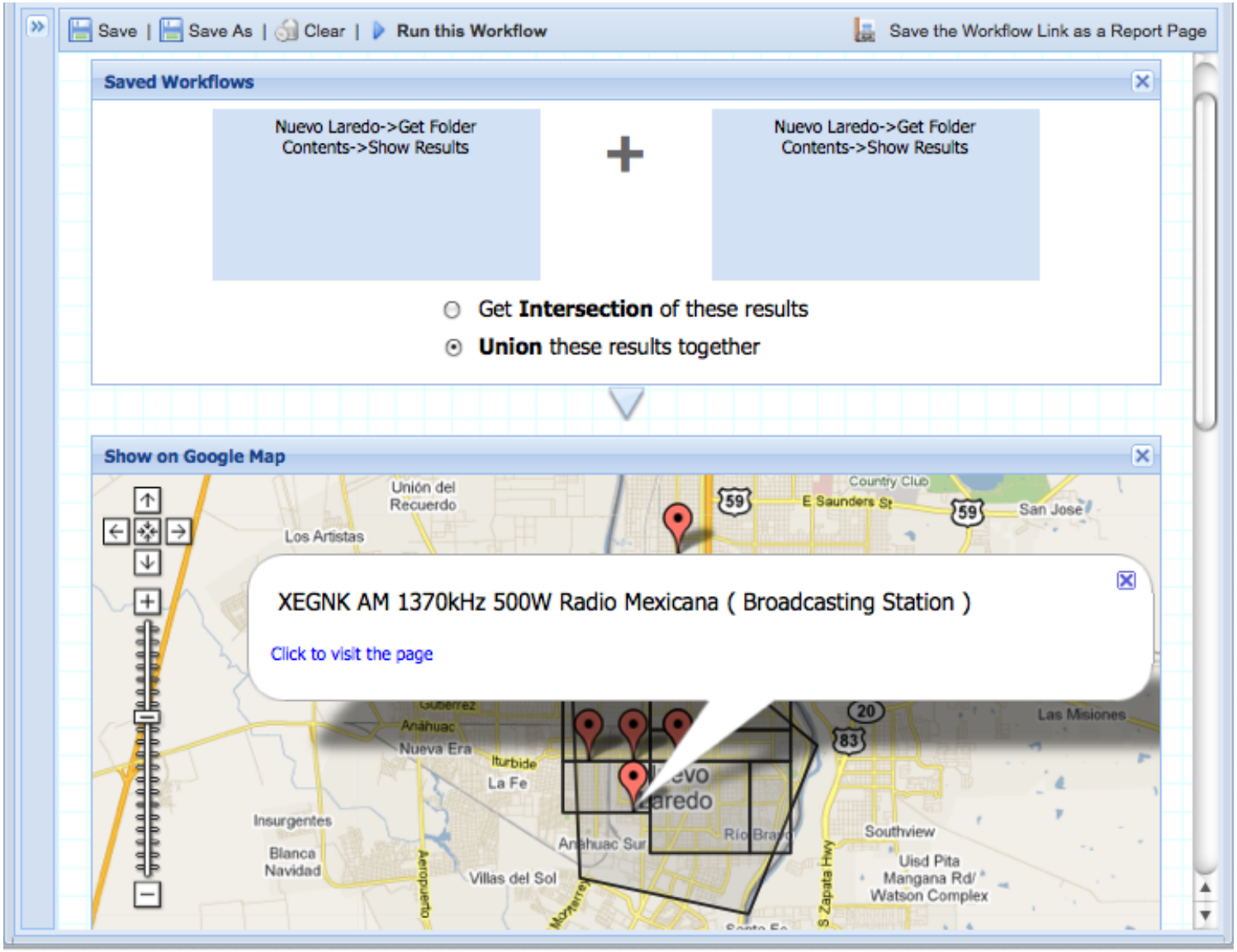
© 2010 Semandex Networks Inc. All Rights Reserved.

Analytic workflow showing neighborhoods & broadcast stations



© 2010 Semandex Networks Inc. All Rights Reserved.

Drill down detail for social media knowledge



© 2010 Semandex Networks Inc. All Rights Reserved.

Ontology Management and Graph Creation

© 2010 Semandex Networks Inc. All Rights Reserved.

Machine assisted text exploitation

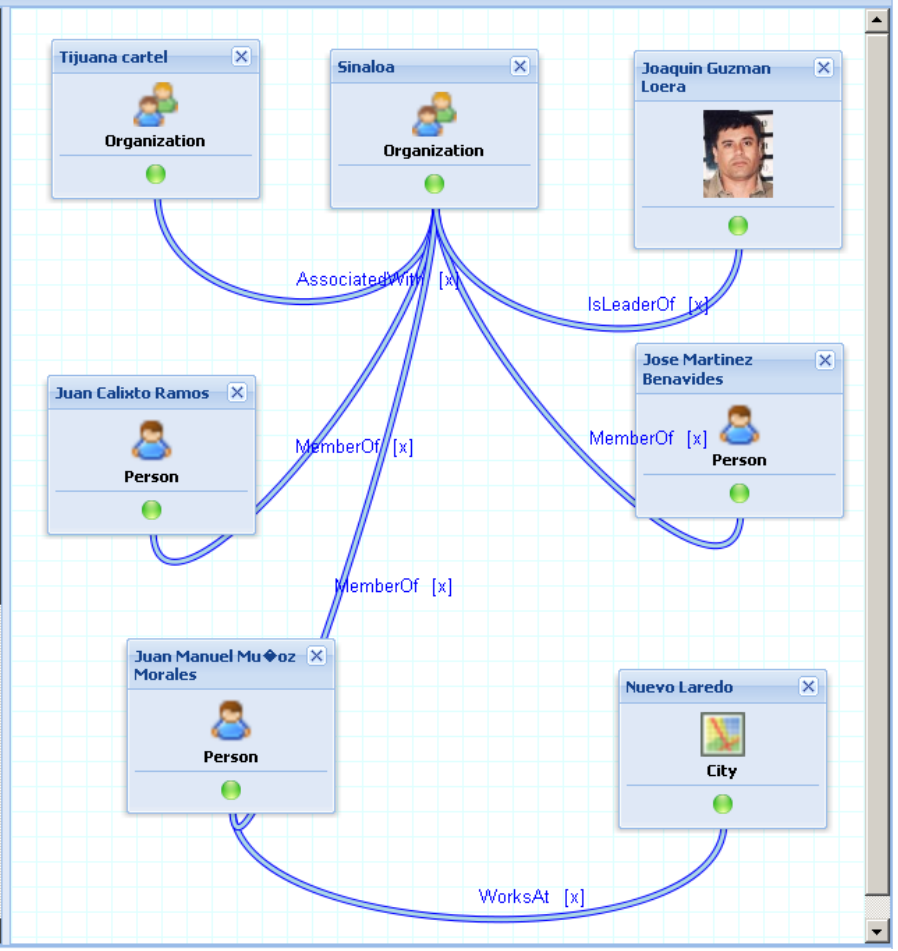
Document: World News: Mexicans get long prison terms in drug case

Run Workflow

Main

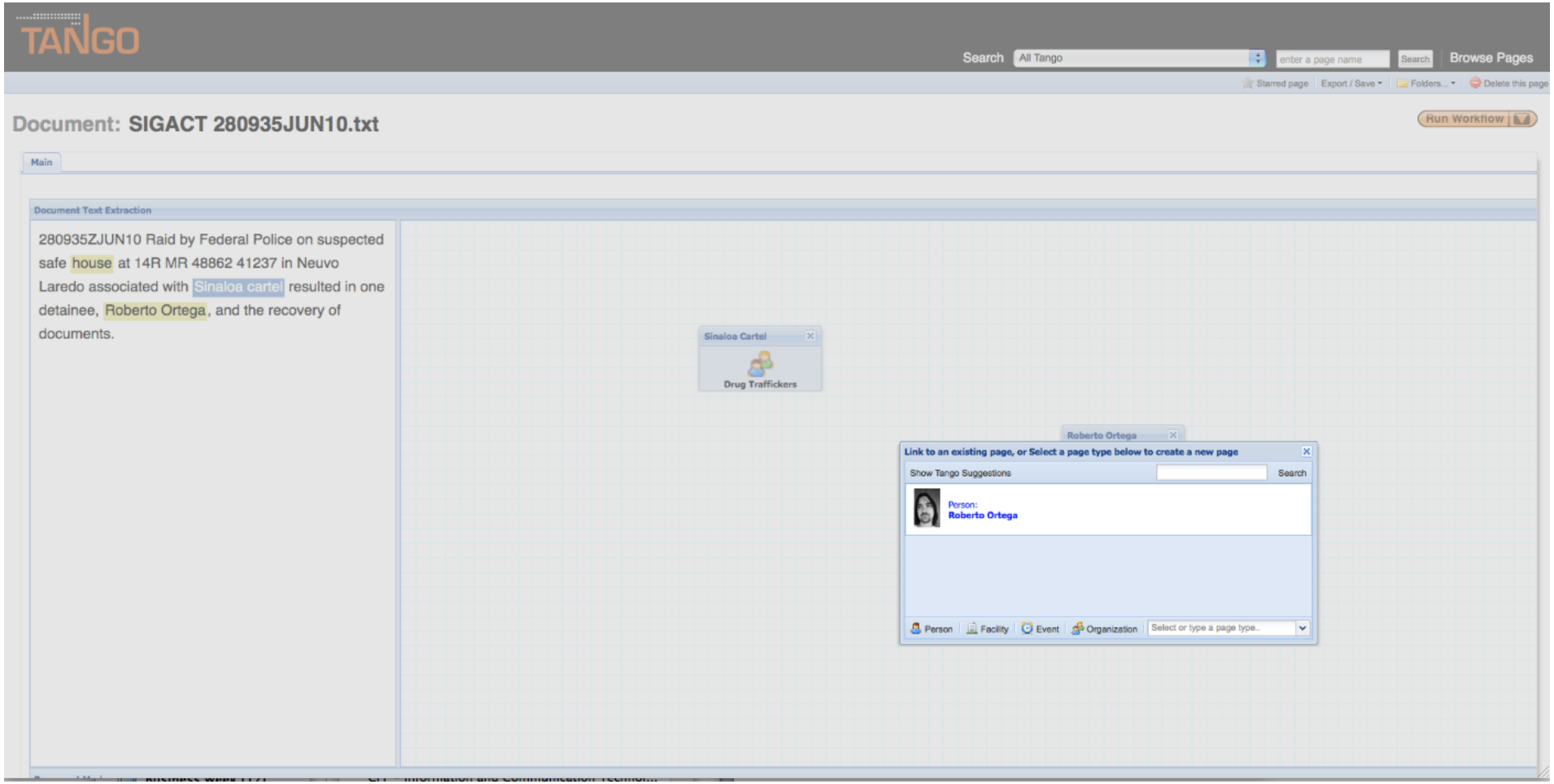
Document Text Extraction

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 Muñoz 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 ($13,800) fine. He was also removed from his post and banned from working as a public servant for nine years. Muñoz was accused of freeing the



© 2010 Semandex Networks Inc. All Rights Reserved.

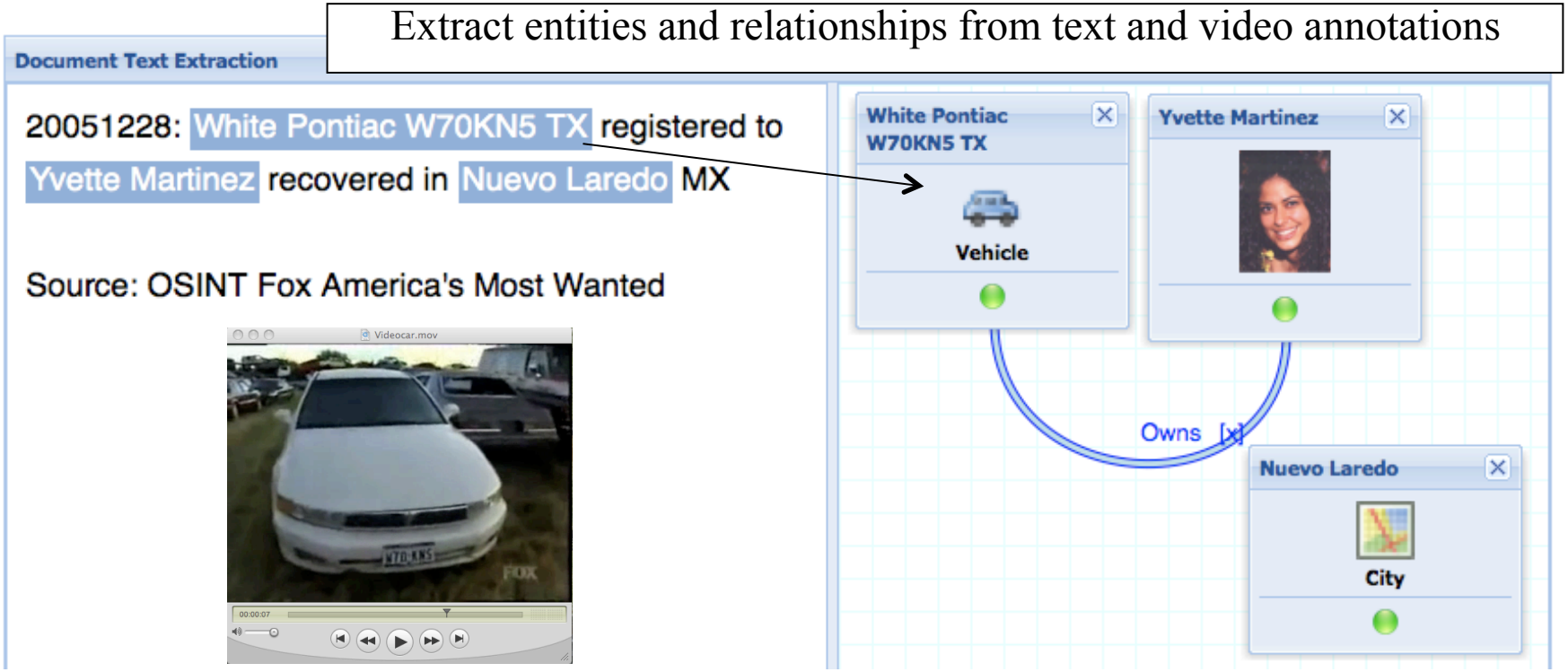
Recommendation Engine suggests matches



© 2010 Semandex Networks Inc. All Rights Reserved.

Machine-assisted semantic annotation of video

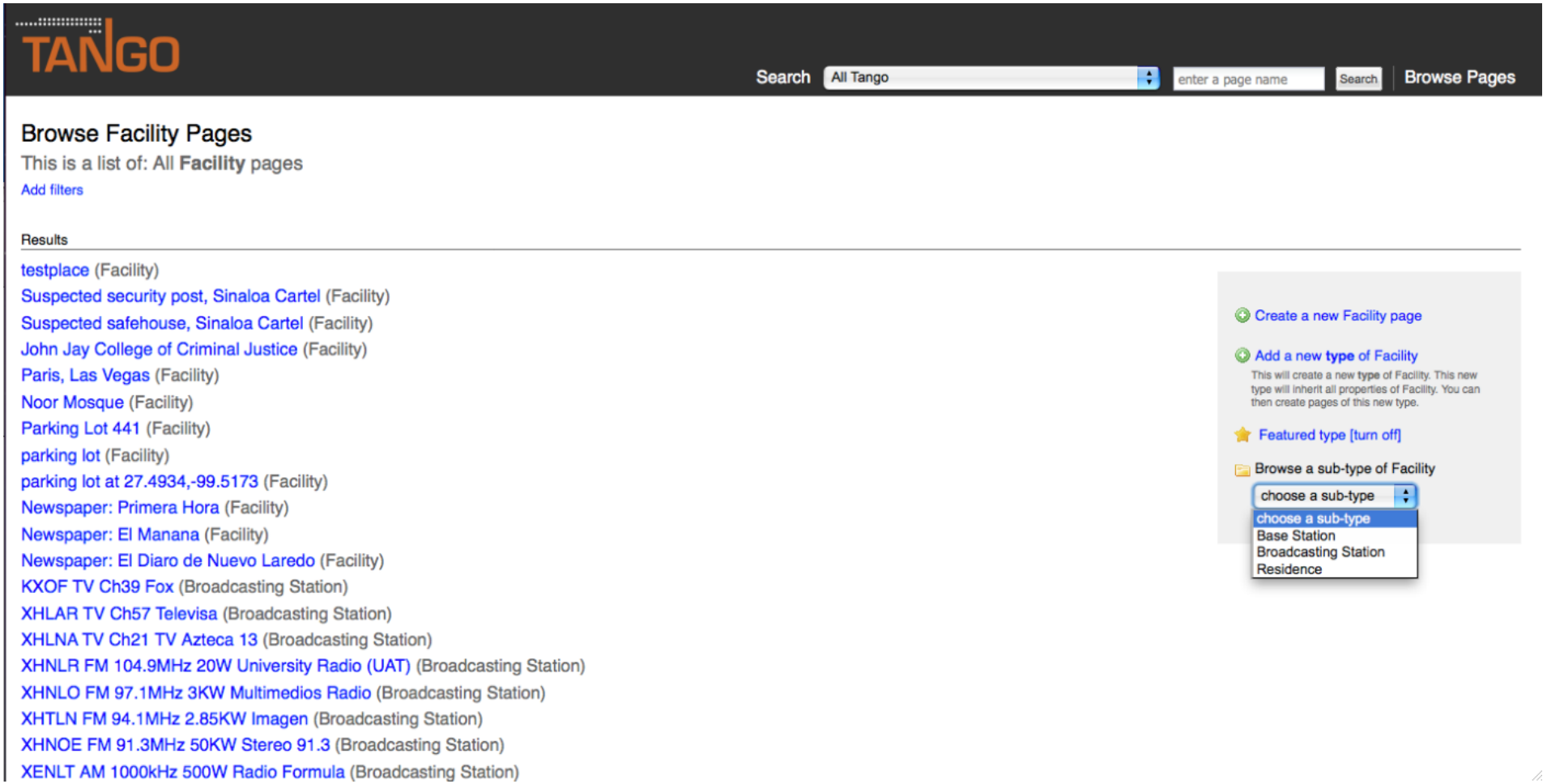
- **Direct linking of video, images and text into the information base**



Semandex
CONNECTED INTELLIGENCE.

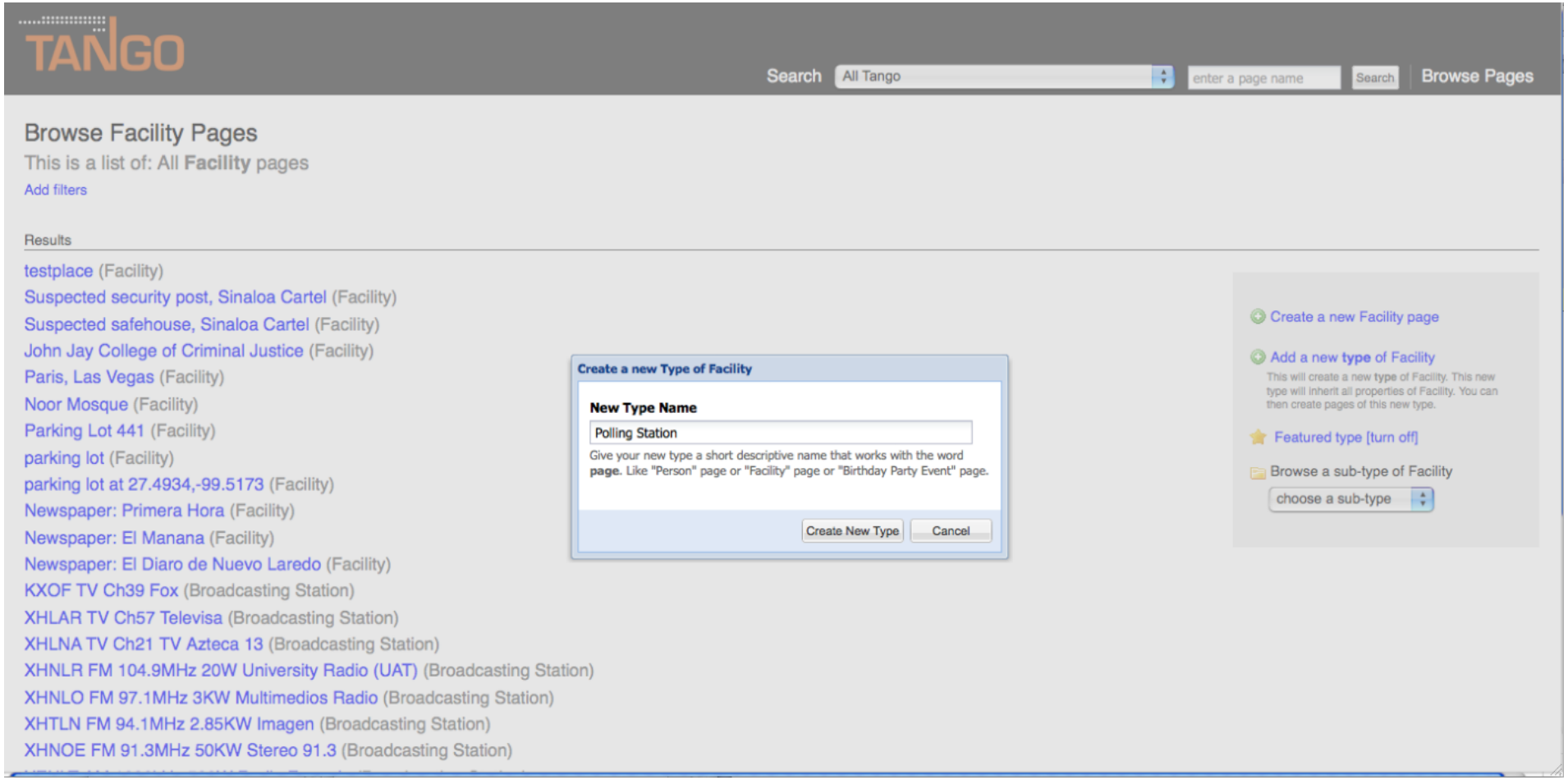
© 2010 Semandex Networks Inc. All Rights Reserved.

Checking existing concept sub-types



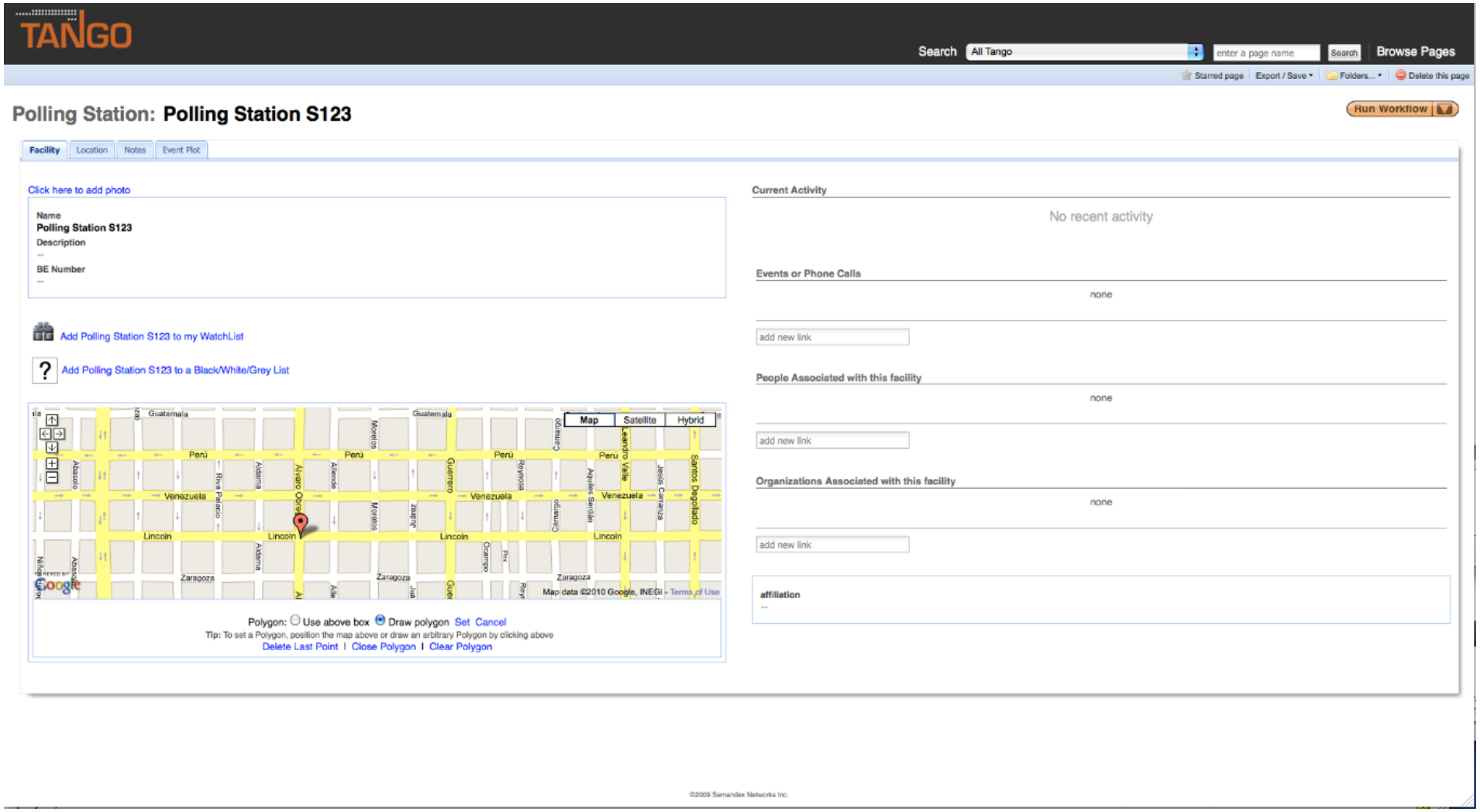
© 2010 Semandex Networks Inc. All Rights Reserved.

Creating new concept sub-types



© 2010 Semandex Networks Inc. All Rights Reserved.

An instance of a new concept sub-type



© 2010 Semandex Networks Inc. All Rights Reserved.

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.

© 2010 Semandex Networks Inc. All Rights Reserved.

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.**

© 2010 Semandex Networks Inc. All Rights Reserved.

Daniel Reininger

dan@semandex.net

www.semandex.net

Semandex
CONNECTED INTELLIGENCE.

© 2010 Semandex Networks Inc. All Rights Reserved.