Administering A 160-Year-Old Database

Knowledge Management At The New York Times

Evan Sandhaus
Lead Architect, Semantic Platforms
Agenda

- Intro
- The Story of Our First Semantic Technology
- Linked Data
- Semantic Markup
- Questions
First Foray Into Semantics?

First Semantic Platform?
Semantics in the early 20th Century

Adolph Ochs
Publisher 1896 - 1935

1913
TITANIC SINKS FOUR HOURS AFTER HITTING ICEBERG; 866 RESCUED BY CARPATHIA, PROBABLY 1250 PERISH; ISMAY SAFE, MRS. ASTOR MAYBE, NOTED NAMES MISSING

The Lost Titanic Being Towed Out of Belfast Harbor.

PARTIAL LIST OF THE SAVED.

Includes Bruce Ismay, Mrs. W. H. Webb, Mr. and Mrs. Lord and 127 others.

CAPT. E. J. SMITH

Captain of the Titanic.
For Business Men

Experts on current economics contribute to The Annalist each week articles on financial and commercial developments throughout the world.

The purpose of The Annalist is to interpret the important news of the day for the benefit of business leaders. Not an Act of Congress, but the potentialities of that Act, form the substance of Annalist comment. Not a coal or a rail strike, but the possibilities of these events in their relation to the welfare of industry and the comfort of the nation, find publication in The Annalist.

Average prices of leading bonds and stocks, food cost of living and Federal Reserve gold holdings are indicated in charts.

The Open Security Market consists of advertisements of reputable brokers quoting prices on unlisted municipal, railroad and other securities.

THE ANNALIST
A Magazine of Finance, Commerce and Economics
Published Mondays by The New York Times Company
TIMES SQUARE, NEW YORK
10 cents a copy. $5.00 a year.
INDIA AND THE MOSLEM AWAKENING

Significance of the historic telegram from the Viceroy of India, which caused a sudden vacancy in Lloyd George's Cabinet and startled the British Empire—The Moslems in India seeking to force the Allies to give Turkey more territory by threatening to overturn British rule—Gandhi's part in the movement

GREAT BRITAIN has made peace with Ireland, but she now faces an even greater danger—the rebellion against her rule in India. Events are moving swiftly in that vast Asiatic possession—all pointing in the same direction, and proving that England's policy in Turkey and the Near East has been the cause of her undoing in India. This was shown only too conclusively by the startling telegram sent by Lord Reading, the present Viceroy of India, on March 7, the publication of which by the Right Hon. Edwin S. Montagu, the British Secretary for India, without the authorization of Lloyd George and during his absence from the Cabinet owing to illness, cost Mr. Montagu his official head. In this message the Viceroy declared that the support which the Moslem cause had received throughout all India and the intensity of Indian feeling over the terms of the Servia Treaty made its immediate revision imperative.

The first Lloyd George knew of this was when he saw it in his morning paper at home. After consultation with other members of the Government, he sent for Mr. Montagu, declined to accept his explanations, and demanded his resignation forthwith, on the ground that he had violated, without cause or necessity, the rule of collective responsibility. This was on March 9. On the following day Mohandas K. Gandhi, Britain's bitterest enemy in India and the leader of the non-co-operation movement, which has injured English prestige in India in the past two years more than the world realizes, was arrested, chiefly it now appears, because of the alliance of his party with the Mohammedans, whose clamors for revision of the Turkish Treaty had become articulate and insistent. On March 18 Gandhi was sentenced to six years' imprisonment without hard labor. The charges against him were admitted. The All-India Congress, however, the organ...

OUR CEMETERIES IN FRANCE

EVERY line that leaves New York carries at least one former member of the A. E. F., and sometimes as many as twenty or twenty-five. They are not only the officers of higher rank, but include many of the men who made up the citizen army. In fact, any former soldier who fought in France, and who can scrape together the money is turning his steps, as well as his thoughts, to the battlefields.

Our cemeteries there, as well as in Belgium and England, will thus necessarily continue for at least a generation to be objects of pilgrimage for Americans. For this reason they should be kept permanently, and given care to maintain a definite memorial, in order that our participation in the war may not be lost sight of, and to justify the actions of those who elected that the bodies of their dead should remain in the soil where they fought and died. This, in most cases, was an act of patriotism and sacrifice. The request for...
The New York Times INDEX

Vol. X. No. 2
April—May—June 1922

Published by
THE NEW YORK TIMES
Times Square, N. Y.

Copyright, 1922, by The New York Times Co.
Complete Sets Available
The New York Times Index
From January, 1913

1913—Covering the Balkan War, China Rebellion, the Mexican Rebellion, and other events in the United States and abroad.

1914-1918—A chronological record of the great World War and momentous happenings in all countries.

1919-1921—League of Nations, tariff and taxation, prohibition, changes in National Administration, reconstruction, Disarmament Conference, aviation.

1922—Irish question, unrest in India, independence of Egypt, ratification of treaties, amateur radio activity, Genoa Conference, Election Campaigns, the latest activities in science, politics and business, besides thousands of subjects and biographies that find their way into the newspaper.

The Index is indispensable for quick reference to dates and facts.

Bound in cloth, $8.00 a year.
Value of New York Times's Index.
(From The Fourth Estate.)

The subject of news indexes is one that
takes an interest at all newspapers
because of the amount of printing that
data and the probability that such
indexes will be of permanent value.

Y. M. C. A. LIBRARIAN VALUES TIMES INDEX

It Will Form Part of the Library
of Every Branch Eventually,
Miss F. R. Petrie Thinks.

FITED TO THEIR FIELD

As a Record of Athletics and Debates It Will Fill a Need Long
Felt by the Association.

TIMES INDEX A HELP IN BUSINESS WORLD

Field of Commerce Covered in an Admirable Way, Says Mr.
Pratt of Chamber Here.

BROAD INFORMATION GIVEN

It Supplements Statistical Data
Now Being Gathered by All
Large Houses and Interests.
Old NYT HQ 10th Floor Reference Library
INDEX TO THE NEW-YORK TIMES.
From September 1851 to September 1858.

The figures represent the number of the Times.

<table>
<thead>
<tr>
<th>A</th>
<th>26, 30, 293</th>
<th>Academy of Medicine (N.Y.) report on Public health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abd-el-Kader murdered for by Lord Londonderry</td>
<td>351</td>
<td>Addison (Laura) appearance in New York</td>
</tr>
<tr>
<td>Visited by Louis Napoleon</td>
<td>352</td>
<td>Her sudden death</td>
</tr>
<tr>
<td>His liberation</td>
<td>362</td>
<td>Biographical Sketch</td>
</tr>
<tr>
<td>Stabs fatally Hamet Ben Aita, a fellow-prisoner</td>
<td>363</td>
<td>Advertising, fortunes made by</td>
</tr>
<tr>
<td>In Paris</td>
<td>368</td>
<td>Adams (John Q) on the Fisheries</td>
</tr>
<tr>
<td>His interview with Louis Napoleon</td>
<td>374</td>
<td>Adams (John Q) letters to Robert Walsh</td>
</tr>
<tr>
<td>Compared to Toussaint L’Ouverture (Ed)</td>
<td>410</td>
<td>(W. J. Q.) Her death and burial</td>
</tr>
<tr>
<td>Invited to England by Lord Londonderry</td>
<td>829</td>
<td>views of J.Q.A. on the fishery question</td>
</tr>
<tr>
<td>Departs for Turkey</td>
<td></td>
<td>on the purchase of Cuba</td>
</tr>
<tr>
<td>History of</td>
<td></td>
<td>371</td>
</tr>
<tr>
<td>Abrahams (George) re-arrested</td>
<td>206, 2085</td>
<td>393</td>
</tr>
<tr>
<td>A. Beckett (Gilbert) his death</td>
<td>1562</td>
<td>427</td>
</tr>
<tr>
<td>A. Beckett (Gilbert) his death</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Descriptors
- Real Estate Brokers
- Tendinitis
- Wolves
- Suits and Litigation
- Pilots
- Banks and Banking
- Brown v Board of Education
- Anasazi Indians
- Brussels Sprouts
- Lakes
- Hypothermia
- Plumbing
- Venus (Planet)
- Burritos
- Muscular Dystrophy
- Eggs
- Exxon Valdez (Tanker)
- Mushrooms
- Caste Systems
- Postal Service
- Sustainable Living
- Redistricting
- Hair
- Ulcers
- Day Laborers

### Locations
- Texas
- Tenafly (NJ)
- Tucson (Ariz)
- Oklahoma
- Mott Haven (NYC)
- St Petersburg (Fla)
- Zermatt (Switzerland)
- Whistler
- Segovia (Spain)
- St-Tropez (France)
- Chiapas (Mexico)
- Gorges du Verdon
- Sharjah
- Dokdo Islands
- Cold Spring (NY)
- Bermuda
- Brooklyn Navy Yard (NYC)
- Ludlow (Vt)
- Reno (Nev)
- New England States (US)
- Kentucky
- Rossville (NYC)
- Yarmouth (Mass)
- Los Roques (Venezuela)
- Lambertville (NY)

### People
- Anthony, Marc
- Weintraub, Jerry
- Tamayo, Rufino
- Singer, Isaac Bashevis
- Wyatt, Edward
- Reed, Pamela
- Lee, Jennifer 8
- Brown, Robbie
- Meester, Leighton
- Pirro, Albert J
- Johnson, Russell
- Hughes, Karen P
- Henry, Kyle
- Buckley, William F Jr
- Bashir, Abu Bakar
- de Mornay, Rebecca
- Prueher, Joseph W
- Demme, Jonathan
- Pine, Chris
- Koteas, Elias
- Lee, Gavin
- Wilkerson, Isabel
- Krusiec, Michelle
- Rockefeller Family
- Gates, Anita

### Organizations
- Flint Telecom Group Inc
- ClearPoint Business Resources Inc
- Tone in Twenty
- Power Medical Interventions Inc
- Liberty Acquisition Holdings Corporation
- Freeport-McMoRan Copper & Gold Inc
- DIAS Holding Inc
- Theravance Inc
- MicroChannel Technologies Corporation
- Vocus Incorporated
- Danvers Bancorp Inc
- Level 3 Communications Inc
- Knightsbridge Tankers Limited
- Natco Group Incorporated
- Sourcefire Inc
- CBS Corp
- Callisto Pharmaceuticals Inc
- IRIS International Inc
- Hanover Insurance Group Inc
- MGP Ingredients Incorporated
- UC Hub Group Inc
- By Design Inc
- Canadian Natural Resources Ltd
- Hartford Financial Services Group Inc
- Timken Company
Why Tag?

**News Alerts**

- Related Searches
  - Housing
  - Tax Credits
  - Law and Legislation
  - Lobbying and Lobbyists

**Widgets**

- NYT > Birds
  - A Natural Treasure That May End Up Without a Country 4/15/08
  - Did Your Shopping List Kill a Songbird? 4/15/08
  - A Man’s Love of Pigeons, and His Fight to Feed Them 4/15/08
  - THE SEASONS 4/15/08
  - The Natural World, Up Close and Personal 4/15/08

**Automate editorial processes**

- Past Coverage
  - Housing Aid Advancing in Congress (April 10, 2008)
  - Chances of Homeowner Relief Losing Momentum in Senate (April 9, 2008)
  - Housing Bill To Get Vote In Senate (April 5, 2008)
  - In Senate, Agreement On Housing (April 3, 2008)

**RSS and Syndication**

- RSS FEEDS ON GLOBAL WARMING
  - Subscribe to an RSS feed on this topic. What is RSS?
    - RSS Global Warming

And to power Topic Pages...
North Korea's Nuclear Program

Updated: May 26, 2009

In October 2006 North Korea became the world's eighth atomic power, conducting an underground nuclear weapons test. Although the country's nuclear program and its development of long range rocket systems has outraged world opinion, it is still unclear whether the country has mastered the ability to deliver a working nuclear weapon.

On May 25, 2009, North Korea announced that it had successfully conducted its second nuclear test, again defying international warnings.
Linked Data
Barack Obama

**Latest Developments:** Updated: Oct. 30, 2009

- President Obama rolled back several changes made by the Bush administration that had weakened the **Intelligence Oversight Board**. Oct. 30

- President Obama paid his first visit to Dover Air Force Base in Delaware to **honor 18 Americans killed in Afghanistan**. Oct. 30

- The president’s speech to a military audience offered few clues about the review of his policy toward Afghanistan and Pakistan. Oct. 27

- President Obama’s new strategy for Sudan is meant to build pressure on the Khartoum government to end the abuses that have killed or displaced millions in Darfur. Oct. 20

- Officials said President Obama could postpone any decision on sending more troops to Afghanistan until the disputed election has been settled. Oct. 19

**OVERVIEW**

Barack Hussein Obama was sworn in as the 44th president of the United States on Jan. 20, 2009. The son of a black man from Kenya and a white woman from Kansas, he is the first African-American to ascend to the highest office in the land.

He is also the first new president since terrorists attacked New York and Washington on Sept. 11, 2001, the first to use the Internet to decisive political advantage, the first to insist on handling a personal smartphone while in the White House. So striking was the unlikelihood of his candidacy that he embraced that aspect, calling himself "a skinny kid with a funny name" and making "change" the theme for his campaign.

**ARTICLES ABOUT BARACK OBAMA**

*Newest First | Oldest First*

**Democrats Strive to Hold Governor Posts**
By DAVID W. CHEN and DAVID M. HALBFINGER
President Obama’s rally in New Jersey for Gov. Jon S. Corzine underscored the White House’s resolve to stave off defeat. Virginia voters will also choose a new governor.
November 2, 2009

**THE CAUCUS**

**If Fox Is Partisan, It Is Not Alone**
By JOHN HARWOOD
Partisan fragmentation throughout America’s news media and their audiences has grown significantly.
November 2, 2009

**OP-ED CONTRIBUTOR**

**Israelis and Obama**
By HENRY SIEGMAN
Obama is not popular in Israel because he is serious about peace.
November 2, 2009

**TELEVISION REVIEWS | ‘POLIWOOD,’ ‘BY THE PEOPLE’**

**Star-Struck Filmmakers in Political Territory**
By ALESSANDRA STANLEY
In politics, the only thing worse than no access is too much access. Two well-meaning documentaries on the anniversary of the election prove the point.
November 2, 2009

**NEWS ANALYSIS**

**With Karzai, U.S. Faces Weak Partner in Time of War**
By DAVID E. SANGER
After Abdullah Abdullah’s withdrawal from the Afghan runoff, President Obama faces the challenge of enabling President Hamid Karzai to regain enough legitimacy to help the U.S. with the war.
Barack Obama

**Latest Developments:** Updated: Oct. 30, 2009

- President Obama rolled back several changes made by the Bush administration that had weakened the **intelligence Oversight Board.** Oct. 30

- President Obama paid his first visit to Dover Air Force Base in Delaware to **honor 18 Americans killed in Afghanistan.** Oct. 30

- The president's speech to a military audience **offered few clues** about the review of his policy toward Afghanistan and Pakistan. Oct. 27

- President Obama's **new strategy for Sudan** is meant to build pressure on the Khartoum government to end the abuses that have killed or displaced millions in Darfur. Oct. 20

- Officials said President Obama could **postpone any decision on sending more troops to Afghanistan** until the disputed election has been settled. Oct. 19

**OVERVIEW**

Barack Hussein Obama was sworn in as the 44th president of the United States on Jan. 20, 2009. The son of a black man from Kenya and a white woman from Kansas, he is the first African-American to ascend to the highest office in the land.

He is also the first new president since terrorists attacked New York and Washington on Sept. 11, 2001, the first to use the Internet to decisive political advantage, the first to insist on handling a personal smartphone while in the White House. So striking was the unlikeliness of his candidacy that he embraced that aspect, calling himself "a skinny kid with a funny name" and making "change" the theme for his campaign.
<table>
<thead>
<tr>
<th>Subject Headings</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locations</td>
<td>2,491</td>
</tr>
<tr>
<td>People</td>
<td>13,027</td>
</tr>
<tr>
<td>Descriptors</td>
<td>3,614</td>
</tr>
<tr>
<td>Organizations</td>
<td>9,909</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>29,041</strong></td>
</tr>
</tbody>
</table>
## City of Overland Park

### City

- **Country**: United States
- **State**: Kansas
- **County**: Johnson

#### Government
- **Mayor**: Carl R. Gerlach

#### Area
- **Total**: 65.3 sq mi (169.3 km²)
- **Land**: 65.2 sq mi (169.0 km²)
- **Water**: 0.1 sq mi (0.3 km²)

#### Elevation
- **1,079 ft (329 m)**

#### Population (2008)
- **Total**: 171,231
- **Density**: 2,627/sq mi (1,014.3/km²)

#### Time zone
- CST (UTC-6)
- Summer (DST): CDT (UTC-5)

#### Area code(s)
- 913

#### FIPS code
- 20-33775[^1]

#### GNIS feature ID
- 0479210[^2]
DBPedia

- 2.6 million entities
- 198,000 persons
- 328,000 places
- 101,000 musical works
- 34,000 films
- 20,000 companies
- 4.9 million links to external data sources
Tremendous database of databases.

- 11 Million Total Topics
- 1.3 Million People
- 800,000 Locations
- 405,926 Business Topics
- 88,563 Government Topics
Geonames

- **HUGE** database of place names.
- 8 million place names
- 2.6 million populated places
- 2.8 alternate place names
- 11 million requests per day
Geonames
Wayne Gretzky

Wayne Douglas Gretzky, CC (born January 26, 1961) is a retired Canadian professional ice hockey player. He is the owner, head of hockey operations, and head coach of the Phoenix Coyotes of the National Hockey League (NHL). Born and raised in Brantford, Ontario, Gretzky honed his skills at a backyard rink and regularly played minor hockey at a level far above his peers.[1] Despite his unimpressive stature, strength, and speed, Gretzky's intelligence and reading of the game were uncanny. He was adept at dodging checks from opposing players, and he could consistently anticipate where the puck was going to be and execute the right move at the right time. Gretzky also became known for setting up behind the net, an area that was nicknamed "Gretzky's office" because of his skills there.[2]

In 1978, he signed with the Indianapolis Racers of the World Hockey Association (WHA), where he briefly played before being traded to the Edmonton Oilers. When the WHA folded, the Oilers joined the NHL, where he established many scoring records and led his team to four Stanley Cup Championships. His trade to the Los Angeles Kings had an immediate impact on the team's performance, leading them to the 1993 Cup finals, and credited with popularizing hockey in the southern United States. Gretzky played briefly for the St. Louis Blues and finished his career with the New York Rangers.

Gretzky captured nine Hart Trophies as the most valuable player, ten Art Ross Trophies for most points in a season, five Lady Byng Trophies for sportsmanship and performance, five Lester B. Pearson Awards, and two Conn Smythe Trophies as playoff MVP.

After his retirement in 1999, he was immediately inducted into the Hockey Hall of Fame, being the last player to have the waiting period waived. He became Executive Director for the Canadian national men's hockey team during the 2002 Winter Olympics, where the team won a gold medal. In 2000, he became part owner of the Phoenix Coyotes.


N Y T G U I D E T O E S S E N T I A L K N O W L E D G E

Gretzky, Wayne

Gretzky, Wayne, b. Brantford, Ontario, Canada, 1961. Hockey player. Wayne Gretzky is widely regarded as the greatest player in the history of ice hockey. In a 20-year NHL career (1979–99), he set all-time records for goals (892), assists (1,487), and total points (2,380). Known as "the great one," the smooth-skating, center topped the league in scoring 10 times, won 11 V.P. honors nine times, and led the Edmonton Oilers to four Stanley Cups. His accolades also include: most points ever in
Mapping Progress

Locations
1,910 of 2,491
76%

Descriptors
498 of 3,614
14%

People
4,978 of 13,027
38%

Organizations
3,081 of 9,909
31%

Total
10,467 of 29,041
36%
For the last 150 years, The New York Times has maintained one of the most authoritative news vocabularies ever developed. In 2009, we began to publish this vocabulary as linked open data.

The Data

The New York Times has published 5,000 people subject headings as linked open data under a CC BY license. We provide both RDF documents and a human-friendly HTML versions.

Browse individual data records:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Or download all of the data records in one SKOS File.

The Effort

The New York Times uses approximately 30,000 tags to power our Times Topics Pages. It is our intention to publish all of these tags as linked open data.

Watch video of the announcement of our linked open data strategy.
Semantic API

- **Semantic API:**
  - Lookup by
    - Name
    - URI
  - Search by name
  - Lookup by article

- **Geo API:**
  - Query by:
    - GeoNames Criteria (e.g. altitude, population ranges).
    - Bounding Box
    - Nearby
What We’ve Learned (So Far)
What We’ve Learned (So Far)

- Linked Data Helps With
  - Adding “ontological context”
  - Automated data alignment
  - Non traditional aggregation

- Areas for improvement
  - Community interaction
  - Data Accuracy
Adding Ontological Context
Adding Ontological Context
The Problem Of Geography
The Problem Of Geography
The Problem Of Geography

New Berlusconi Scandal Could Threaten His Government
By ELISABETTA POVOLEDO
The most recent scandal involving the Italian prime minister’s private life could open the door to a full-fledged government crisis.

Berlusconi Faces Public Discontent Over Italy’s Trash Crisis
By RACHEL DONADIO
Prime Minister Silvio Berlusconi of Italy is facing discontent as trash piles up on the streets of Naples.
October 29, 2010
The Problem Of Geography

HEADS UP

More Doors Open in Rome

A nighttime tour at the Colosseum, one of the sites extending hours and opening areas previously closed.

By KATIE PARLA
Published: September 9, 2010

Rome (Italy) Italy
The Problem Of Geography
But With Linked Data (Geonames)
Auto Aligning Data
Automated Data Alignment

Congress API

NYT Archive
Automated Data Alignment

Congress API

Congressional Bio Id
Automated Data Alignment

Index Subject Headings

NYT Archive
Automated Data Alignment

NYT Archive

Franken, Al (Per)

Index Subject Headings

Congressional Bio Id

Congress API

F000457
Automated Data Alignment

Index Subject Headings

Congressional Bio Id
Automated Data Alignment

Index Subject Headings

Congressional Bio Id
Non Traditional Aggregation
A New News Map
A New News Map
A New News Map

The New York Times

Linked Open Data

Xian (China)

OP-ED COLUMNIST: Failure Is Not An Option
By THOMAS L. FREDMAN
April 29, 2010
China is having a good week in America. Yes it is. I’d even suggest that there is some high-fiving going on in Beijing. I mean, wouldn’t you...read more

As China Surges, It Draws High-Tech Researchers From America
By KEITH BRADISHER
March 18, 2010
XIAN, China – For years, many of China’s best and brightest left for the United States, where high-tech industry was more cutting-edge. But...read more

Map data ©2010 Google, Esri, DeLorme, U.S. Census, NASA, NGA, USGS, NRCan, and OpenStreetMap contributors. Terms of Use
A New News Map
A New News Map
Alumni In The News
Alumni In The News

Ed Wilson
Born: June 10, 1929

Scientists Square Off on Evolutionary Value of Helping Relatives - August 31, 2010
BOOKS ON SCIENCE; A Similar Terrain for Ants and Men - April 20, 2010
Ear to the Ground - April 11, 2010
BOOKS OF THE TIMES; Life Lessons, Taught By Insects - April 06, 2010
E Pluribus Unum - November 23, 2006
Corrections: For the Record - July 19, 2008
Corrections - July 18, 2008
SCIENTIST AT WORK | EDWARD O. WILSON; Taking a Cue From Ants On Evolution Of Humans - July 15, 2008
The Encyclopedia of Life, No Bookshelf Required - February 26, 2008
Scientist Finds the Beginnings of Morality in Primate Behavior - March 20, 2007

Franklin D. Roosevelt
Lawyer
Lawyer
Born: January 30, 1882
Died: April 12, 1945

Book Review - Scorpions - By Noah Feldman - November 07, 2010
OP-ED COLUMNIST; Time for the This Big Dog to Bite Back - September 12, 2010
Citing F.D.R., Putin Hints at 3rd-Term Bid - September 07, 2010
OP-ED COLUMNIST; 1938 In 2010 - September 06, 2010
ECONOMIC VIEW; What Would Roosevelt Do? - August 01, 2010
In Roosevelt Archive, History as He Made It - July 29, 2010
OP-ED CONTRIBUTOR; General Upset; The Other Truman Doctrine - June 23, 2010
OP-ED COLUMNIST; When Greatness Slips Away - June 22, 2010
FROM THE INTERNATIONAL HERALD TRIBUNE; 100, 75, 50 Years Ago - June 01, 2010
OP-ED COLUMNIST; The Old Enemies - May 24, 2010
Standing Tall
Standing Tall

Danny DeVito

Born: 1944-11-17 (Auburndale, Florida)
Profession: Film Director, Actor, Film Producer, Writer
Nationality: United States of America
Gender: Male
Religion: Roman Catholicism

Daniel Michael "Danny" DeVito, Jr. (born November 17, 1944) is an American actor, director and producer, who first gained prominence for his portrayal of "Louie De Palma" on the ABC and NBC TV series Taxi (1978–1983). DeVito and his wife, Rhea Perlman, founded Jersey Films, a production company known for films such as Pulp Fiction, Garden State, and Freedom Writers. He also owns Jersey Television, which produced the Comedy Central series Reno 911...

Articles

MOVIE REVIEW | 'SOLITARY MAN'; Yes, He's a Jerk, but There's Something About Ben
By A. O. SCOTT
May 21, 2010
-- not a moment too soon! -- in Oliver Stone's sequel to "Wall Street," But you don't have to wait that long to see Michael Douglas playing a disgraced businessman. In the meantime there is "Solitary Man," a sharp, small-scale comedy of male misbehavior that turns out to be one of this dreary spring's pleasant cinematic surprises. Not that Mr...More

'Always Sunny,' With Singing
By DAVE ITzkOFF
undefined 16, 2009
As the fall television season begins, and new and returning shows jostle one another to remind easily distracted viewers that they exist, there are many routes series can take to promote themselves. Some buy print advertisements and rent billboards; others use Internet banners and viral videos. But probably only one is sending its cast on the road...More
Areas For Improvement
Areas For Improvement

- **Data Accuracy**
  - DBPedia, Freebase, Geonames and several other data sets, are based on community curated data with all the advantages and disadvantages that entails.
  - What can we trust while still maintaining the organization’s high standard?

- **Community Involvement**
  - Hope was that releasing the data would result in a number of apps around our data.
  - API’s are *definitely* helping.
Future Directions
Future Directions

- Continue mapping subject headings to external data sources.
- Operationalize prototype linked data infrastructure.
- Empower internal and external developers with the launch of three new APIs.
  - Concepts API [Launched]
  - Links API
  - Geo API [Launched]
- Collaborate with other media organizations and broader community to develop a framework for annotating news documents with RDFa.
rNews
Embedded Data For
The News Industry
Why we need Semantic Markup
Diplomatic Breakdown Amid Bieber Fever in Israel

By ISABEL KERSHNER
Published: April 12, 2011

JERUSALEM — The teenage pop idol Justin Bieber became embroiled in a diplomatic imbroglio on Tuesday when it emerged that plans for a meeting between the singer and Prime Minister Benjamin Netanyahu of Israel had been called off, with the sides differing over why.

Mr. Bieber, who was discovered on YouTube and is famous for hits like “Baby,” is in Israel for an outdoor concert on Thursday night.

A spokesman for Mr. Netanyahu said that his office had been approached with the idea of a meeting, and that the prime minister had been “open to that.” The prime minister’s office then suggested including children from communities in southern Israel that have been under intense rocket fire from Gaza in recent days.

“Sadly,” the spokesman said, “that proved impossible,” suggesting that Mr. Bieber’s representatives had turned down the idea of including the children.

Related
Times Topics: Benjamin Netanyahu | Justin Bieber
Diplomatic Breakdown Amid Bieber Fever in Israel

By ISABEL KERSHNER
Published: April 12, 2011

JERUSALEM — The teenage pop idol Justin Bieber became embroiled in a diplomatic quandary on Sunday when it emerged that plans for a meeting between the Prime Minister Benjamin Netanyahu of Israel had been called off, with the sides differing over why.

Mr. Bieber, who was discovered on YouTube and is famous for hits like “Baby,” is in Israel for an outdoor concert on Thursday night.

A spokesman for Mr. Netanyahu said that his office had been approached with the idea of a meeting, and that the prime minister had been “open to that.” The prime minister’s office then suggested including children from communities in southern Israel that have been under intense rocket fire from Gaza in recent days.

“Sadly,” the spokesman said, “that proved impossible,” suggesting that Mr. Bieber’s representatives had turned down the idea of including the children.
are not so obvious to a machine.
The Problem of Structured Data

- Modern Web Sites Built with 3 Tier Architecture
  - **Data Tier**: Database Where Content Lives.
  - **Presentation Tier**: HTML Document that is sent to user.
  - **Logic Tier**: Software that reads from the Data Tier and outputs the Presentation Tier.

![Diagram of 3 Tier Architecture]

- Display Tier
- Logic Tier
- Data Tier
The Problem Of Structured Data: Continued

- Content very well structured on Data Tier, but all of this structure is lost in translation to presentation tier.
The Problem Of Structured Data: Continued

Display Tier

<html>
<head>
  <title>
    New Web Code Draws Concern...
  </title>
</head>
<body>
  <div>
    New Web Code Draws Concern...
  </div>
  <div>
    By TANZINA VEGA
  </div>
  <div>
    October 10, 2010
  </div>
  <div>
    In the next few years, a powerful...
  </div>
</body>
</html>

- Search engines, social networks, aggregators and other sites only see the Display Tier, and cannot leverage the underlying structure of the data.
The Problem Of Structured Data: Continued

- Without structured data search engines, social networks and other sites cannot attractively format links back to our site, potentially decreasing referral traffic.
The Case of the Missing Structured Data

- How do we solve it?
- Four approaches:
  - Microformats
  - RDF/RDFa
  - Microdata
  - JSON
Semantic Markup Standards

- **Microformats**
  - First
  - Simple
  - Rigid

- **RDFa**
  - Official
  - Complex
  - OpenGraph

- **Microdata**
  - Unofficial
  - Flexible
  - Schema.org

- **JSON**
  - Official
  - Developers
  - External
rNews
rNews Defined

- rNews is a data model for embedding machine-readable publishing metadata in web documents and a set of suggested implementations.
- slightly shorter
rNews is a data model

DRAFT - Version 0.7
for embedding machine-readable publishing metadata in web documents
and a set of suggested implementations

RDFa
Today

Microdata
Today

JSON
Maybe?
How Does rNews Fit in With Existing IPTC Standards

NewsML
NITF
XML
feeds
CMS
publisher
rNews
HTML
web
rNews - Timeline

- **September 2010** - rNews proposed to IPTC at fall meeting
- **March 2011** - rNews draft version 0.1 approved by IPTC at summer meeting.
- **March - May 2011** - IPTC solicits feedback on draft standard.
- **June 2011** - IPTC to vote on revised standard at summer meeting
- **October 2011** - rNews 1.0 Approved by IPTC
- **November 2011** - Start of rNews Implementation on nytimes.com
Allies Are Split on Goal and Exit Strategy in Libya

NATO Takes Command

Rebel fighters take cover during a shelling near Ajdabiya, Libya on Thursday.

By STEVEN LEE MYERS
WASHINGTON | March 24, 2011

Having largely succeeded in stopping a rout of Libya’s rebels, the inchoate coalition attacking Col. Muammar el-Qaddafi’s forces remains divided over the ultimate goal — and exit strategy — of what officials acknowledged Thursday would be a military campaign that could last for weeks.

© Copyright 2011The New York Times Company

Disclaimer
Allies Are Split…

NATO Takes Command

Rebel fighters take...

By STEVEN LEE MYERS

WASHINGTON | March 24, 2011

Having largely succeeded...

© Copyright 2011 The New York Times Company

Disclaimer

So the question is...

March 25th, 2011 8:27 am

Chuck
HTML 5 Microdata

<!DOCTYPE HTML>
<html itemscope itemtype="http://schema.org/NewsArticle">
<head>
    <style type="text/css">@import url(css/iptc_times2.css);</style>
    <meta itemprop="dateCreated" content="2011-03-23"/>
    <meta itemprop="description" content="The questions about the command...">
    <meta itemprop="inLanguage" content="en-US"/>
    <meta itemprop="thumbnailUrl" content="http://graphics8.nytimes.com/images/common/icons/t_wb_75.gif"/>
    <meta itemprop="genre" content="Current"/>
    <meta itemprop="id" content="1248069687395"/>
    <meta itemprop="version" content="2"/>
    <meta itemprop="publishingPrinciples" content="http://www.nytco.com/press/ethics.html"/>
    <meta itemprop="wordCount" content="879"/>
</head>
<body>
    <div style="height:900px" class="article">
        <div class="a_column">
            <div itemprop="headline" class="headline">Allies Are Split on Goal and Exit Strategy in Libya</div>
            <div itemprop="alternativeHeadline" class="rider">NATO Takes Command</div>
            <div itemprop="associatedMedia" itemscope itemtype="http://schema.org/ImageObject">
                <img itemprop="URL" class="image" src="img/libya_sample_reuters.jpg"/>
                <div class="image_credit">Credit:
                    <span itemprop="creator" itemscope itemtype="http://schema.org/Person">
                        <span itemprop="name">Goran Tomasevic</span>
                    </span>
                </div>
            </div>
        </div>
    </div>
</body>
Allies Are Split on Goal and Exit Strategy in Libya

NATO Takes Command

Credit: Goran Tomasevic/Reuters

Rebel fighters take cover during a shelling near Ajdabiyah, Libya on Thursday.
The Way to rNews
The Way To rNews

- June: Genesis of rNews - Evan at SemTech 2010
- November 5 - Rome: chartered
- *internal discussions about NYT draft*
- March 9 - Dubai: rNews 0.1
- *lots of feedback, changes and additions*
- June 9 - Berlin: rNews 0.5
- June 28: rNews 0.6
- September 6: rNews 0.7 [aligned w/ schema.org]
- October 7 - Vienna: rNews 1.0
Engaging Our Community
Engaging Our Community
Engaging Our Community
Engaging Our Community
Feedback we incorporated...
Feedback we incorporated...

Person \[\text{editor} \rightarrow \text{NewsItem}\]
What is Schema.org?

This site provides a collection of schemas, i.e., html tags, that webmasters can use to markup their pages in ways recognized by major search providers. Search engines including Bing, Google and Yahoo! rely on this markup to improve the display of search results, making it easier for people to find the right web pages.

Many sites are generated from structured data, which is often stored in databases. When this data is formatted into HTML, it becomes very difficult to recover the original structured data. Many applications, especially search engines, can benefit greatly from direct access to this structured data. On-page markup enables search engines to understand the information on web pages and provide richer search results in order to make it easier for users to find relevant information on the web. Markup can also enable new tools and applications that make use of the structure.

A shared markup vocabulary makes easier for webmasters to decide on a markup schema and get the maximum benefit for their efforts. So, in the spirit of sitemaps.org, Bing, Google and Yahoo! have come together to provide a shared collection of schemas that webmasters can use.

We invite you to get started!
**Thing > CreativeWork**

The most generic kind of creative work, including books, movies, photographs, software programs, etc.

<table>
<thead>
<tr>
<th>Property</th>
<th>Expected Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>properties</td>
<td></td>
<td></td>
</tr>
<tr>
<td>about</td>
<td>Thing</td>
<td>The subject matter of the content.</td>
</tr>
<tr>
<td>accountablePerson</td>
<td>Person</td>
<td>Specifies the Person that is legally accountable for the CreativeWork.</td>
</tr>
<tr>
<td>aggregateRating</td>
<td>AggregateRating</td>
<td>The overall rating, based on a collection of reviews or ratings, of the item.</td>
</tr>
<tr>
<td>alternativeHeadline</td>
<td>Text</td>
<td>A secondary title of the CreativeWork.</td>
</tr>
<tr>
<td>associatedMedia</td>
<td>MediaObject</td>
<td>The media objects that encode this creative work. This property is a synonym for encodings.</td>
</tr>
<tr>
<td>audio</td>
<td>AudioObject</td>
<td>An embedded audio object.</td>
</tr>
<tr>
<td>author</td>
<td>Person or Organization</td>
<td>The author of this content. Please note that author is special in that HTML 5 provides a special mechanism for indicating authorship via the rel tag. That is equivalent to this and may be used interchangeably.</td>
</tr>
<tr>
<td>awards</td>
<td>Text</td>
<td>Awards won by this person or for this creative work.</td>
</tr>
<tr>
<td>comment</td>
<td>UserComments</td>
<td>Comments, typically from users, on this CreativeWork.</td>
</tr>
<tr>
<td>contentLocation</td>
<td>Place</td>
<td>The location of the content.</td>
</tr>
<tr>
<td>contentRating</td>
<td>Text</td>
<td>Official rating of a piece of content—for example, 'MPAA PG-13'.</td>
</tr>
<tr>
<td>contributor</td>
<td>Person or Organization</td>
<td>A secondary contributor to the CreativeWork.</td>
</tr>
<tr>
<td>copyrightHolder</td>
<td>Person or Organization</td>
<td>The party holding the legal copyright to the CreativeWork.</td>
</tr>
<tr>
<td>copyrightYear</td>
<td>Number</td>
<td>The year during which the claimed copyright for the CreativeWork was first asserted.</td>
</tr>
</tbody>
</table>
This class contains derivatives of IPTC rNews properties. rNews is a data model of publishing metadata with serializations currently available for RDFa as well as HTML5 Microdata. More information about the IPTC and rnews.org can be found at