Exposing Latent Information in Folksonomies for Reasoning

January 14, 2010

Sponsored by

Defense Advanced Research Projects Agency (DOD)

Controlling DARPA Office: IPTO

ARPA Order AW79-00

Issued by U.S. Army Aviation and Missile Command Under

Contract No. W31P4Q-09-C-0382

20100201291

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
14-01-2010	Final Report	4/14/2009-12/23/2009

4. TITLE AND SUBTITLE
Exposing Latent Information in Folksonomies for Reasoning

5a. CONTRACT NUMBER
W31P4Q-09-C-0382

5b. GRANT NUMBER
SB082-032

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Ulicny, Brian E. (PI) (VIStology, Inc.)
Kogut, Paul A. (Lockheed Martin)
Heintzelman, Norris H. (Lockheed Martin)
Leung, Yui H. (Lockheed Martin)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
VIStology, Inc. 5 Mountainview Drive, Framingham, MA 01701

8. PERFORMING ORGANIZATION REPORT NUMBER
FolkEvents Final Report

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr. Daniel Oblinger
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
3701 North Fairfax Drive
Arlington, VA 22203-1714
(703) 526-4170

10. SPONSOR/MONITOR'S ACRONYM(S)
DARPA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
A. Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Work involved subcontractor Lockheed Martin Information Systems and Global Services, King of Prussia, PA

14. ABSTRACT

User tagging by means of authority-free folksonomies has become standard for making online videos, photos, bookmarks and blog posts discoverable. We have prototyped the design of a system that identifies and exposes the ontological structure that is latent in the folksonomic tags used in open, publicly available sites such as Flickr (photos), YouTube (video) and Blogger (blogs) and demonstrated the ability to perform higher order processing using the induced structure. Specifically, the technology being developed allows the system to identify photos, videos or blog posts about the same event by means of the tags applied in one medium (e.g. photos) in order to identify tagged items in the same medium or other media (videos or blog posts). We have used large-scale topic hierarchies such as Wikipedia (dbPedia) and lexical resources such as the Library of Congress' Thesaurus for Graphic Materials in order to tokenize and disambiguate the tags into What, Where and When elements, and shown that these can be used to correlate event depictions with the same precision, and better recall, than non-semantic methods such as clustering.

15. SUBJECT TERMS

events; information retrieval; folksonomy; ontology; multimedia information retrieval; geotagging; inference

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Brian Ulicny
U	U	U	SAR	18	19b. TELEPHONE NUMBER *(Include area code)* 508 788 5088

Reset

Abstract

User tagging by means of authority-free folksonomies has become standard for making online videos, photos, bookmarks and blog posts discoverable. We have proposed to design and develop a system that identifies and exposes the ontological structure that is latent in the folksonomic tags used in open, publicly available sites such as Flickr (photos), YouTube (video) and Blogger (blogs) and demonstrates the ability to perform higher order processing using the induced structure. Specifically, the technology being developed will allow the system to identify photos, videos or blog posts about the same event by means of the tags applied in one medium (e.g. photos) in order to identify tagged items in the same medium or other media (videos or blog posts). We proposed to use large-scale topic hierarchies such as Wikipedia (dbPedia) and lexical resources such as the Library of Congress' Thesaurus for Graphic Materials in order to tokenize and disambiguate the tags into unambiguous topics. These will then be mapped onto *what, when, and where* slots, and used to identify depictions of the same event in other media using a formal reasoner, VIStology's BaseVISor, a forward-chaining OWL/RDF reasoner using OWL 2 RL, a tractable set of axioms for reasoning about knowledge encoded as OWL triples.

Table of Contents

List of Figures

1 Summary

In this study, we show that depictions of the same event may be more easily identified within and across various media through semantic processing of folksonomy tags (Spiteri, 2007) assigned to the depictions independently by various users, using no standardized vocabulary. The semantic processing outlined here identifies depictions of the same event by, first, processing folksonomy tags in order to identify tags that depicted locations (Where), activity types (What) and event dates (When), using large-scale semantic resources such as dbPedia and the Library of Congress's Thesaurus for Graphic Materials. Inference is used to determine synonymous terms (using the SKOS vocabulary), and to compute temporal and geospatial distances. The semantic approach described yields greater recall of identical events, compared with the non-semantic method, clustering depictions by tag sets.

2 Introduction

Online digital photography and video sharing sites, as well as blogging and micro-blogging platforms like Twitter, have become an important outlet for sharing depictions of local events with a worldwide audience. Sites like Blogger (blog posts), Flickr (photos) and YouTube (videos) have become important repositories of depictions of such social disruptions as demonstrations, protests, riots, fights, and accidents (Ulicny et al, 2010). Users make depictions of these events discoverable primarily by means of 'tagging'. Tagging a media object (text, video, photo, etc.) is providing a set of keywords that the user feels would describe that depiction for others seeking depictions of that event, or as a way for the user to navigate their photos and videos themselves.[1] These tag phrases are not selected from a pre-authorized set of phrases for depicting events or other things. They are simply whatever the user thinks would be useful for someone looking for depictions of the photo or video's content (Lerman and Jones, 2007). As such, they display a high level of variation. In some cases, a tagging site will suggest tags for a particular item, either based on what others have tagged the same item (e.g. bookmark URLs in the case of social-bookmaring site del.icio.us) or through semantic analysis of the text of the item (e.g. Twine.com).

In many cases, the choice of tags is determined by the thing depicted: people are tagged with their names; scenes are tagged with their location and visible features and setting (e.g. "Paris", "Eiffel Tower", "rain", "evening"). Events make an interesting case for processing tagged items, because events that happen in public often do not have an obvious description. For example, the following set of tags (Table 1) come from a set of 54 YouTube videos linked to by various bloggers discussing a large protest march in Kuala Lumpur in November, 2007, one of the largest protests in Malaysia in recent years.

Table 1 YouTube tags for Kuala Lumpur Protest, November 10, 2007

1. politics, BERSIH, Reformasi, Malaysia, KL, Kuala, Lumpur
2. Kuala, Lumpur, gathering, Bersih
3. Malaysia, Kuala, Lumpur, demonstrations, rally, Hamish, McDonald, Al, Jazeera, Aljazeera, grassroots, outreach
4. BERSIH, demonstration, Malaysia, rally, 10, november, dataran, merdeka
5. BERSIH, malaysia, 10, November, SPR, demo, demonstrasi, protest, pkr, pas, dap, ngos, dataran, merdeka

In general, different YouTube users almost never tag videos of the same event in the same way. Further, each set of tags is too precise for use as a query to find videos of the same event: almost all of them contain too many keywords. Thus, since using all of the tags as keywords would direct the system to look for the Boolean conjunction of those terms, the recall associated with each set of tags as a whole is very low, where recall is the number of relevant videos returned as a fraction of the entire set of videos depicting that event. The key to identifying videos of the same event here is noticing that the term 'Bersih' is common to nearly all of the tag sets (except the second one). BERSIH (meaning 'clean' in Malay) was the nickname of the sponsoring organization of the rallies, the Coalition for Clean and Fair Elections.

[1] We take events to be specified in terms of the time and location in which they occurred, the participants, and the type of activity that occurred. (Davidson, 1967)

Conversely, many users will often apply the same tag set to all of the photos or videos they upload at one time. This is called 'bulk tagging' and is usually facilitated by the photo or video-sharing site's software as a convenience to users. Bulk tagging has the effect of producing misleading tag-tag correlations, if the content creator and date is not taken into account. That is, if a user bulk-tags 100 photos or videos with 'foo' and 'bar', then if these 100 photos or videos represent a large percentage of all the photos tagged with 'foo' and 'bar', the correlation of these two tags overall will be overestimated, because the bulk upload by a single user is not being distinguished from 100 independent events of photo/video tagging.

In an example that we will use as a running illustration of our approach, on June 5, 2008, two climbers – a professional stuntman named Alain Robert, aiming to draw attention to global warming, and a second, amateur climber named Renaldo Clarke, aiming to draw attention to malaria – independently climbed the façade of the recently built, fifty-story tall New York Times Building in New York City and were arrested after completing their climbs.

Table 2 depicts some of the tag sets used to depict photos of the event on Flickr, found by querying [climbing new york times building] using Flickr's whole text query option. These terms might appear anywhere on the page for this photo, not just in tags.

Table 2 Flickr tags of NY Times Building climbing event

1. nyc, newyorkcity, ny, spiderman, nytimes, newyorktimesbuilding, renaldoclarke,
2. ny, nyt, timesbuilding, nyab,
3. building, climb, copycat, spiderman, timessquare, activist, newyorktimes, daredevil, portauthority, 8thavenue, 050508, nytbuilding,
4. new, york, nyc, building, climbing, times, scaling, renaldoclarke
5. nyc, newyorkcity, sculpture, ny, newyork, statue, harlem, manhattan, gothamist, civilrights, higherground, civilrightsmovement, adamclaytonpowelljr, adamclaytonpowell, adamclaytonpowelljrstateofficebuilding, branlycadet, adamclaytonpowelljrstateofficebuildingplaza,
6. gothamist, climber, newyorktimes, alainrobert, renaldoclarke,

The fifth entry was applied to a photo of a statue of politician Adam Clayton Powell depicting him climbing a slope. It was returned because the search matched the surrounding text on the photo page, including the tags.

Table 3 shows some of the tags associated with YouTube videos depicting the same event. Some of the YouTube videos are footage shot by independent people who happened to have a video camera. Other videos on YouTube are captured from TV or other websites and uploaded to YouTube.

Table 3 YouTube tags for NY Times Climbing event

1. International, spiderman, newyork, June5, 2008, man, climbing, building, times, square
2. BBC, News, Spiderman, scales, NY, building
3. Man, Climbs, NYC, Skyscraper, June, 2008, Times, Square, Green, Building
4. NY, times, Climbing, climber, new, york, midtown, ny, building
5. Renaldo, Clarke, Daredevil, NYTimes, New, York, Times, Building, scales, climbs, Top, documentary, climber, Alain, Robert, scale, NYC
6. climbs, new, york, times, man, building, new, york, city, manhattan
7. skyscraper, climber, men, scale, nyc
8. Renaldo, Clarke, Computer, expert, not, Alain, Robert, French, Spiderman, Human, fly, Climbs, man, climbing, New, York, Times, Building
9. colors, project, save, earth, times, new, york
10. climbing, new, york, times, building, dare, devil, escalating

Table 4 shows some of the tags associated with blog posts describing the same event. Because blog posts are themselves textual items, fewer blog posts are tagged in addition, especially those by non-professional bloggers.

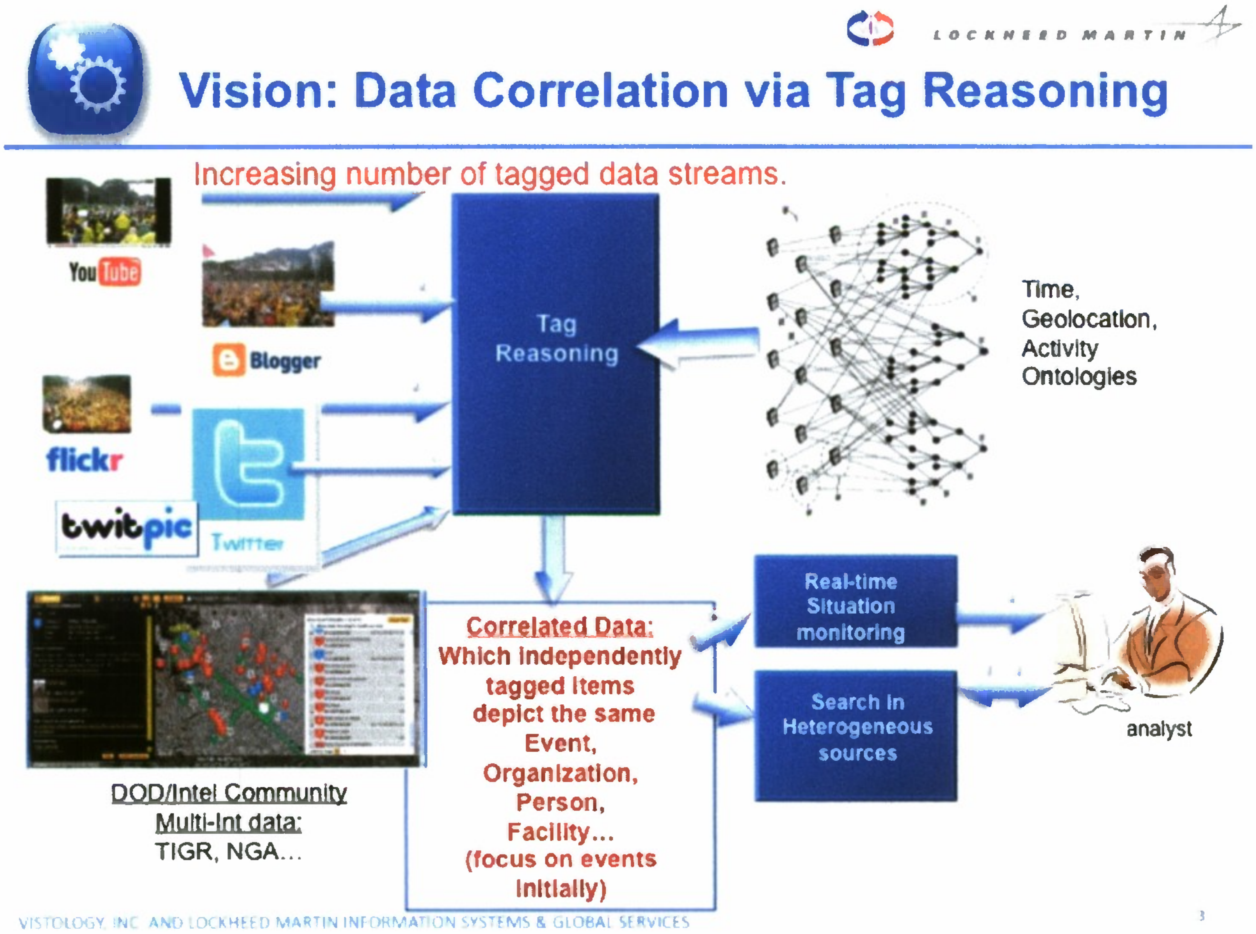
Table 4 Technorati Blog Posts Tags about NY Times Climbing event

1. Offbeat News, alain robert, building, new york, scaling, spider man
2. New York Times Building, Rey Clarke, Stunts, The Spiderman
3. Geniuses, it just happened, new york times
4. alain robert, new york times,
5. government, global warming

Blog post tags are often used more for assigning categories to blog posts in order to provide a category-based form of navigating a blog's content than they are used for describing the content of the blog post itself.

3 Methods, Assumptions, and Procedures

In this SBIR Phase 1 project, we prototyped a system that identifies and exposes the ontological structure that is latent in the folksonomic tags used in open, publicly available sites such as Flickr, YouTube and Blogger and demonstrates the ability to perform higher order processing of this data (Kokar et al, 2004) using the induced structure. Specifically, when mature, the technology developed (Figure 1) will allow the system to identify photos, videos or blog posts depictions of the same event, and later, other types of entities, by means of the tags applied in one medium (e.g. photos) in order to identify tagged items, perhaps in other mediums (videos or blog posts), depicting the same event (e.g. the same protest, traffic accident, or other incident).



Vision: Data Correlation via Tag Reasoning

Increasing number of tagged data streams.

Tag Reasoning

Time, Geolocation, Activity Ontologies

Correlated Data: Which independently tagged items depict the same Event, Organization, Person, Facility... (focus on events initially)

Real-time Situation monitoring

Search in Heterogeneous sources

analyst

DOD/Intel Community Multi-Int data: TIGR, NGA...

VISTOLOGY, INC AND LOCKHEED MARTIN INFORMATION SYSTEMS & GLOBAL SERVICES

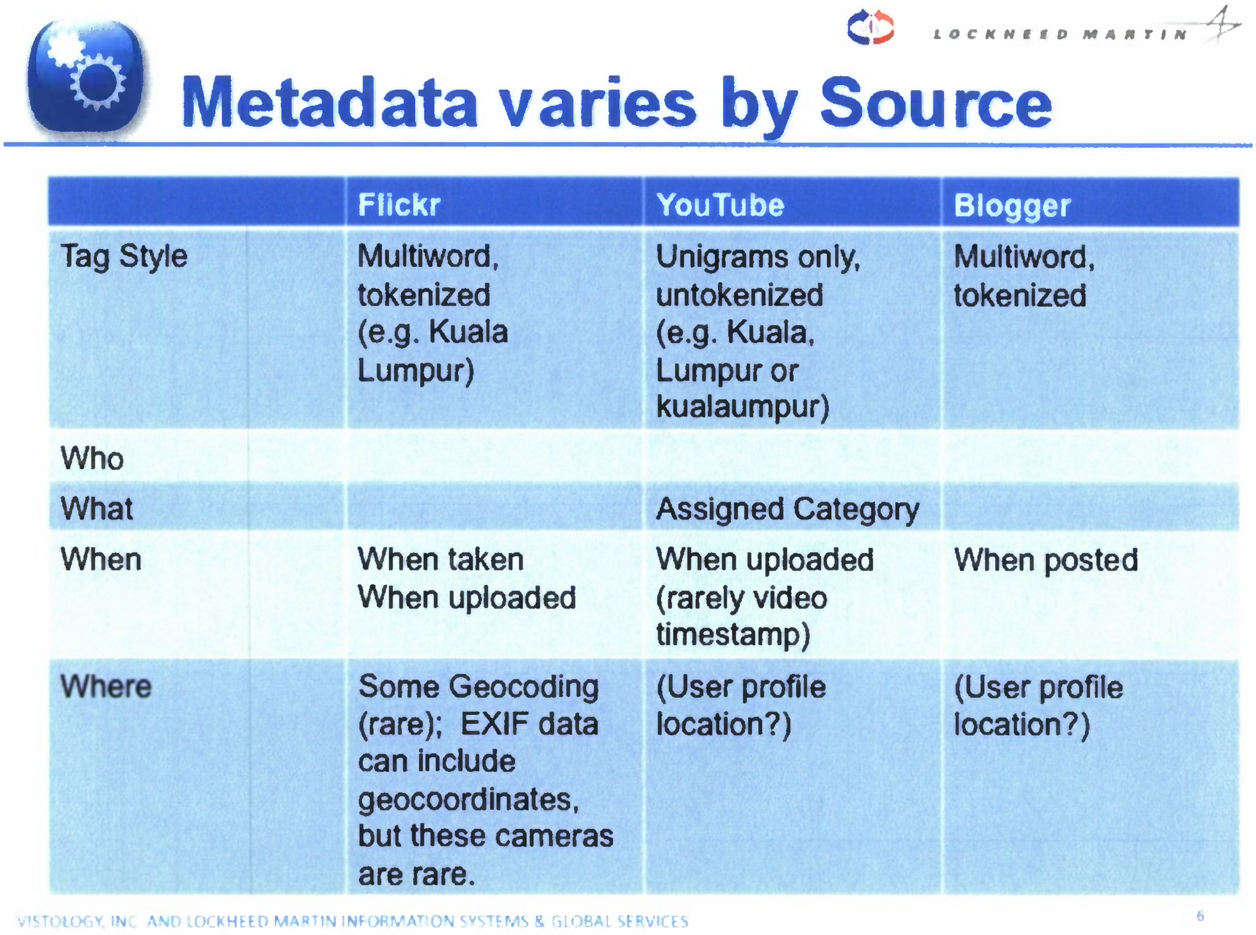
3

Figure 1 Event Depiction Correlation via Tag Reasoning

The prototype system we have developed as a feasibility study for reasoning about depictions of the same event across different folksonomically-tagged depictions and different media types consists of the following components: a tag tokenizer; several tag classifiers and coders; and an event identity reasoner. We outline our progress with respect to each component. We assume that the tags that have been assigned to depictions are assigned in an attempt to make a depiction identifiable, and that the tagging has not been done in a deliberately misleading way. No attempt has been made, in this study, to identify untrustworthy tag assigners or otherwise detect adversarial tagging. Similarly, no attempt has been made to analyze the content of photos or videos directly, via their color distribution or soundtrack, only via their tags.

3.1 Tag Tokenization:

Tags are based on ordinary natural language terms, but it may be difficult to understand where a term begins and ends if it includes whitespaces (i.e. a multiword term like 'New York'. Multiword terms may be concatenated together into a single tag, e.g. 'newyorkcity' or 'new+york+city'. Recovering natural language terms from tags is important for mapping the tags to ontology nodes. In Flickr and Blogger, the tokenization of tags can be recovered directly from the data, so multiword tags can be identified directly. YouTube, however, does not support multiword tags, so the term "new york" will appear as two tags ("new" and "york"). As such, YouTube users often use compound tags without whitespace (e.g. "newyork" or "new_york").



LOCKHEED MARTIN

Metadata varies by Source

	Flickr	YouTube	Blogger
Tag Style	Multiword, tokenized (e.g. Kuala Lumpur)	Unigrams only, untokenized (e.g. Kuala, Lumpur or kualaumpur)	Multiword, tokenized
Who			
What		Assigned Category	
When	When taken When uploaded	When uploaded (rarely video timestamp)	When posted
Where	Some Geocoding (rare); EXIF data can include geocoordinates, but these cameras are rare.	(User profile location?)	(User profile location?)

VISTOLOGY, INC. AND LOCKHEED MARTIN INFORMATION SYSTEMS & GLOBAL SERVICES

6

Figure 2 Depiction Metadata by Source

We have developed scripts that allow us to interact with Flickr, YouTube and Blogger to identify photos/videos/blog posts containing particular tags. We have collected a corpus of approximately 200 videos, photos, and blog posts from YouTube, Flickr and Blogger depicting:

1. The Bersih rally for election reform held in Kuala Lumpur on November 10, 2007. The rally, sponsored by the Coalition for Clean and Fair Elections ('Bersih' means 'clean' in Malaysian) means was held to press for electoral reforms in the lead-up to the Malaysian General Election in March, 2008, in which the ruling party lost the supermajority in parliament that it had enjoyed since the nation's independence, 50 years ago. The rally involved tens of thousands of Malaysians and police, over the distance of a few miles in central Kuala Lumpur over several hours.
2. The illicit climbing of the façade of the New York Times building in Manhattan by two climbers on the same day June 5, 2008, two climbers – a professional stuntman named Alain Robert, aiming to draw attention to global warming, and a second, amateur climber named Renaldo Clarke aiming to draw attention to malaria – independently climbed the façade of the recently built, fifty-story tall New York Times Building in New York City and were arrested after completing their climbs.
3. Various photos, videos, and blog posts containing at least one tag of 'protest' or 'demonstration'

In the following Flickr tagset about climbing the New York Times building in New York City on June 5, 2008, we have the following tags:

nyc, newyorkcity, ny, spiderman, nytimes, newyorktimesbuilding, renaldoclarke,

Obviously, several of these tags are multiword expressions in which the whitespace has been suppressed. Flickr's API returns both a "raw" (includes whitespace) form of tags and the normalized form (downcased, no whitespace). Uniformly, multiword tags are entered as successive unigram tags. Taggers never enter "building, times, ny" when they mean "ny times building".

In other instances, we see multiword terms separated as individual tags, as in this set of tags for a YouTube video about the same event:

NY, times, Climbing, climber, new, york, midtown, ny, building

Here, "NY" and "times" presumably form the multiword expression "NY times" (sic).

An analysis of the 997 unique user-supplied (as opposed to camera-supplied) tags in this corpus shows that the ability to decompound tags is not crucial. Terms are only infrequently combined into a single tag (e.g. 'urbanclimber', 'dataranmerdeka', 'newjersey'). Our initial understanding had been that Flickr tags required large-scale decompounding, but this is not the case. Flickr's API provides multi-word tags in both compounded ("newyork") form, in which text is downcased and whitespace removed, and decompounded ("New York") form. None of the systems we used failed to provide a way to retrieve user-supplied tags in the order the user wrote them, rather than, say, alphabetical order. Re-ordering the tags would make it more difficult to recover "New York Times Building" from "building, new, times, york", since all orderings would have to be examined.

3.2 Tag Classification

Once the tags have been tokenized, we map the tags to categories in a topic hierarchy automatically. We have used several named entity recognizer applications and application programming interfaces (APIs) in order to assign tags to categories. The applications/APIs used were:

- AeroText – named entity markup (People, Organizations, Locations, Times)
- Lockheed Martin's Wikipedia/dbPedia entity matcher
- Thesaurus for Geographic Materials (T4GM.info)
- Flickr metadata, including EXIF (Exchangeable Image File), camera-supplied metadata.
- YouTube metadata

Lockheed Martin's dbPedia tokenizer attempts to find dbPedia entries in a sliding window of tokens in the tag set of a depiction. dbPedia (www.dbpedia.org) is an encoding of Wikipedia's structured data, extracted and encoded in RDF (Auer et al., 2008). Matching terms are identified, along with any synonymous terms, which are encoded as "redirect" items. In addition, categories and geocoordinates are returned, if they exist.

Aerotext, a named entity recognition system developed by Lockheed Martin and now marketed by Rocket Software, was used to identify persons, locations, organizations, facilities, and other named entities in the tag sets. Aerotext was primarily employed here to identify temporal expressions. Secondarily, terms that were identified by Aerotext as named entities did not need to be looked up in dbPedia, unless they were identified as a type (e.g. Facility) that potentially had geocoordinates.

In addition to the unigram tag set the video provider specifies, YouTube provides, for each video, an optional user-assigned Category, an upload time for the video (not the time when the video was filmed), and a user location (via the user's profile information, if provided). We have not made any attempt to use the profile location information to geolocate videos or disambiguate user video tags. Details are specified in the YouTube API documentation[2].

[2] http://code.google.com/apis/youtube/overview.html

Flickr's API allows one to retrieve both multiword tag tokens and compounded versions of the same tags (e.g. "newyork"). In addition, Flickr returns both when a photo was taken and when it was uploaded. In addition to profile location information, which we have not made use of, Flickr photos can contain explicit geo-location information provided by the user in the form of location tags. Additionally, cameras that provide geocoordinates of the camera when a photo was taken will have that information in their EXIF data, which can be retrieved via the Flickr API. Cameras that geocode the position where pictures were taken are still rather rare; we did not encounter any such photos in our data.

Blogger provides multiword-tokenized tag sets, and blog post timestamps via its API. Tagging of blog posts is comparatively much less common than tagging of photos and videos. The text of a blog post is indexed by blog search engines like Google Blog Search or Technorati, so tagging a blog post may seem redundant to bloggers. Only 24% of blog posts in the 2006 Weblogging Ecosystem corpus provided by Blogpulse contain any tags at all (Berendt & Hanser, 2007).

For activity categorization, we have developed an activity tagger using the Thesaurus for Graphic Materials API provided at T4GM.info. T4gm.info is a Linked Data rendering in RDFa of the Library of Congress' Thesaurus for Graphic Materials implemented and maintained by Bradley P. Allen. The *Thesaurus for Graphic Materials* (Library of Congress, 2007) is a vocabulary for indexing visual materials by subject, genre and format. The thesaurus includes more than 7,000 subject terms and 650 genre/format terms to index photographs and other pictorial items. T4gm.info uses the SKOS (Simple Knowledge Organizaing System) standard to express lexical relations (e.g. broader term, narrower term, alternate term, related term) among the thesaurus terms. In our tagger based on T4GM.info, tags are first stemmed, and then candidate terms in the Thesaurus are returned via the T4GM.info API, which does prefix matching. Matches on head terms (e.g. "march" matches "march, protest" or "protest march", but not "marching band") are then determined algorithmically, and the complete RDFa encoding of the SKOS-coded thesaurus data is returned. Leveraging the Thesaurus for Graphic Materials via the T4GM.info API in this way provided a useful way to identify terms and their synonyms (e.g. rally, protest, demonstration) in tag sets that had a much higher recall than the event terms identified by OpenCalais in an evaluation done early in the project.

3.3 Event Depiction Similarity Computing

We determine whether two events depictions are similar, initially, by determining if they coincide in "What", "When" and "Where" they depict, according to their tags. That is, two depictions of events depict the same event if tags expressing when the event took place, where they took place, and the type of event depicted, are close. This is illustrated in Figure 3. Specifically, the inference engine infers that:

(Event Depiction Identity Rule)

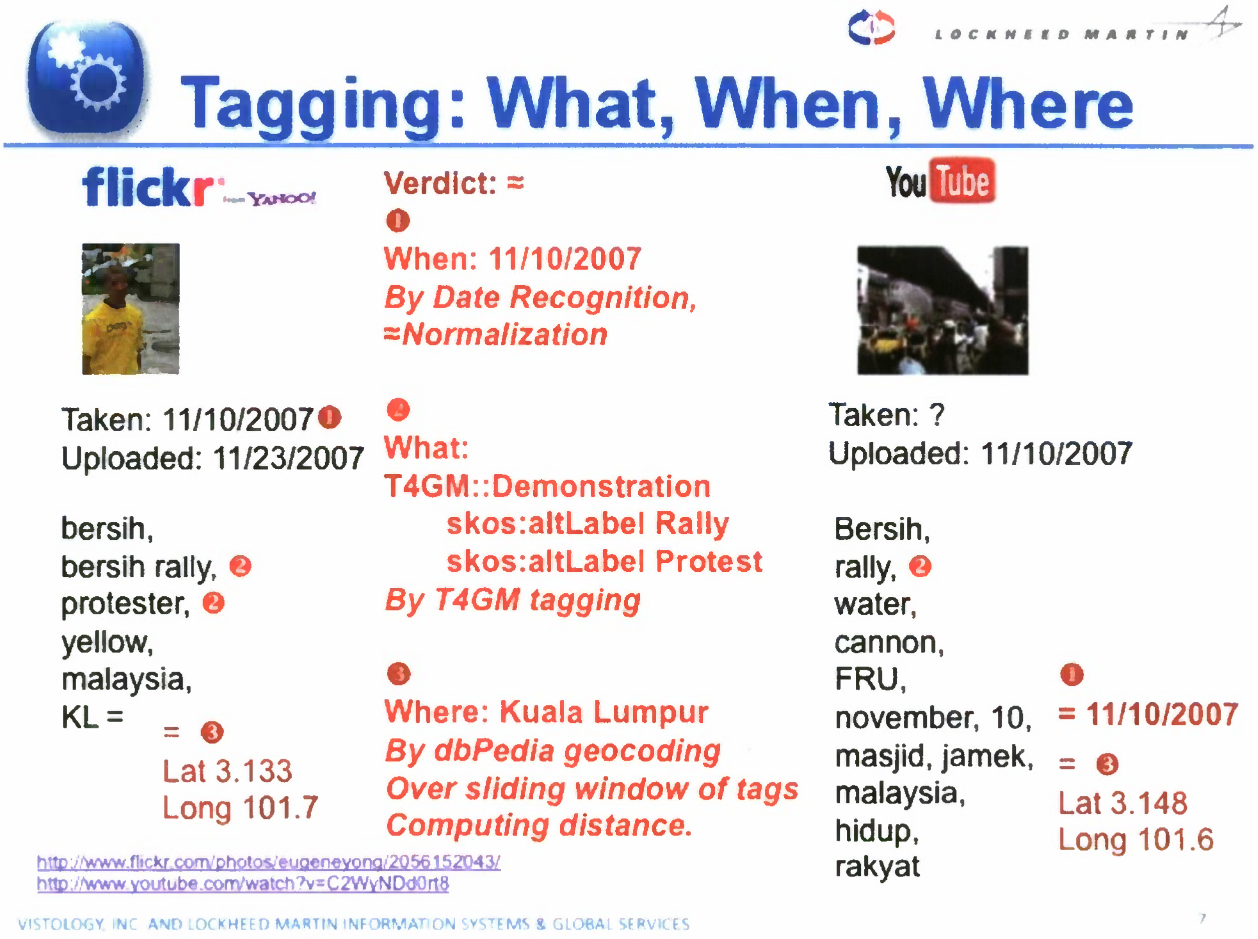
Depiction D1 and depiction D2 depict the same event E if:

D1 and D2 contain the same or synonymous T4GM Activity, Event or similar category

D1 and D2 have a recording time on the same day, plus or minus one calendar day,

D1 and D2 have a geolocation within 20 miles of one another.

We present an example of this reasoning in Figure 3.



LOCKHEED MARTIN

Tagging: What, When, Where

flickr from YAHOO!

Taken: 11/10/2007❶
Uploaded: 11/23/2007

bersih,
bersih rally, ❷
protester, ❷
yellow,
malaysia,
KL =
= ❸
Lat 3.133
Long 101.7

Verdict: ≈

❶
When: 11/10/2007
By Date Recognition,
≈Normalization

❷
What:
T4GM::Demonstration
skos:altLabel Rally
skos:altLabel Protest
By T4GM tagging

❸
Where: Kuala Lumpur
By dbPedia geocoding
Over sliding window of tags
Computing distance.

YouTube

Taken: ?
Uploaded: 11/10/2007

Bersih,
rally, ❷
water,
cannon,
FRU,
november, 10,
masjid, jamek,
malaysia,
hidup,
rakyat

❶
= 11/10/2007

= ❸
Lat 3.148
Long 101.6

http://www.flickr.com/photos/eugeneyong/2056152043/
http://www.youtube.com/watch?v=C2WyNDd0rt8

VISTOLOGY, INC. AND LOCKHEED MARTIN INFORMATION SYSTEMS & GLOBAL SERVICES

7

Figure 3 Assigning Tags to Event What, When, Where

Figure 3 shows two event depictions, a photo on Flickr and a video on YouTube. The photo metadata indicates that the photo was taken on 11/10/2007 (item 1, at left). We only consider the day the photo was taken, plus or minus one day, to allow for incorrect clocks and/or unset time zone specifications. One of the tags is "KL" (item 3, at left). The dbPedia tagger we developed recognizes this as an acronym for "Kuala Lumpur", the site of the Bersih rally, and assigns it the geo-coordinates 3.133 latitude and 101.7 longitude. The depiction is assigned the activity type T4GM::Demonstration, with synonymous terms "protest" and "rally" because of the tags "Bersih rally" and "protester" (item 2, at left).

Similarly, the YouTube video is assigned the depiction day ("When") of 11/10/2007, because of the partial date in the tag set "November, 10", and the upload time "11/10/2007". The system thus infers that the full date of the event is 11/10/2007 (item 1, at right). Although in this case the upload date was the same as the depiction date, this is not always (or even usually) the case. The system incorporates the knowledge that the time of recording an event cannot be later than its upload time. The video is assigned the T4GM::Demonstration category category by our T4GM.info-based tagger because of the tag "rally" (item 2, right). The geocoordinates 3.148 (latitude) and 101.6 (longitude) are assigned to the depiction because the dpPedia tagger recognizes "Masjid Jamek" as the name of a mosque in Kuala Lumpur, which has been assigned geocoordinates in dbPedia (item 3, right).

Because these two depictions coincide in terms of the What (activity type, here the T4GM category "Demonstration"), the Where (two geo-coordinates that are less than 20-miles apart (our default radius), and the When (the same day, November 10, 2007), the system takes them to be depictions of the same event. The computation of distances between geo-coordinates is done using a rule with a procedural attachment in BaseVISor, VIStology's forward-chaining inference engine (Matheus et al, 2006). Similar rules are used to compute temporal distances. BaseVISor is also used here to infer absolute dates (11/10/2007) from partial dates (November 10), in the user-supplied tags, using rules about the upload times vs. recording times.

4 Results and Discussion

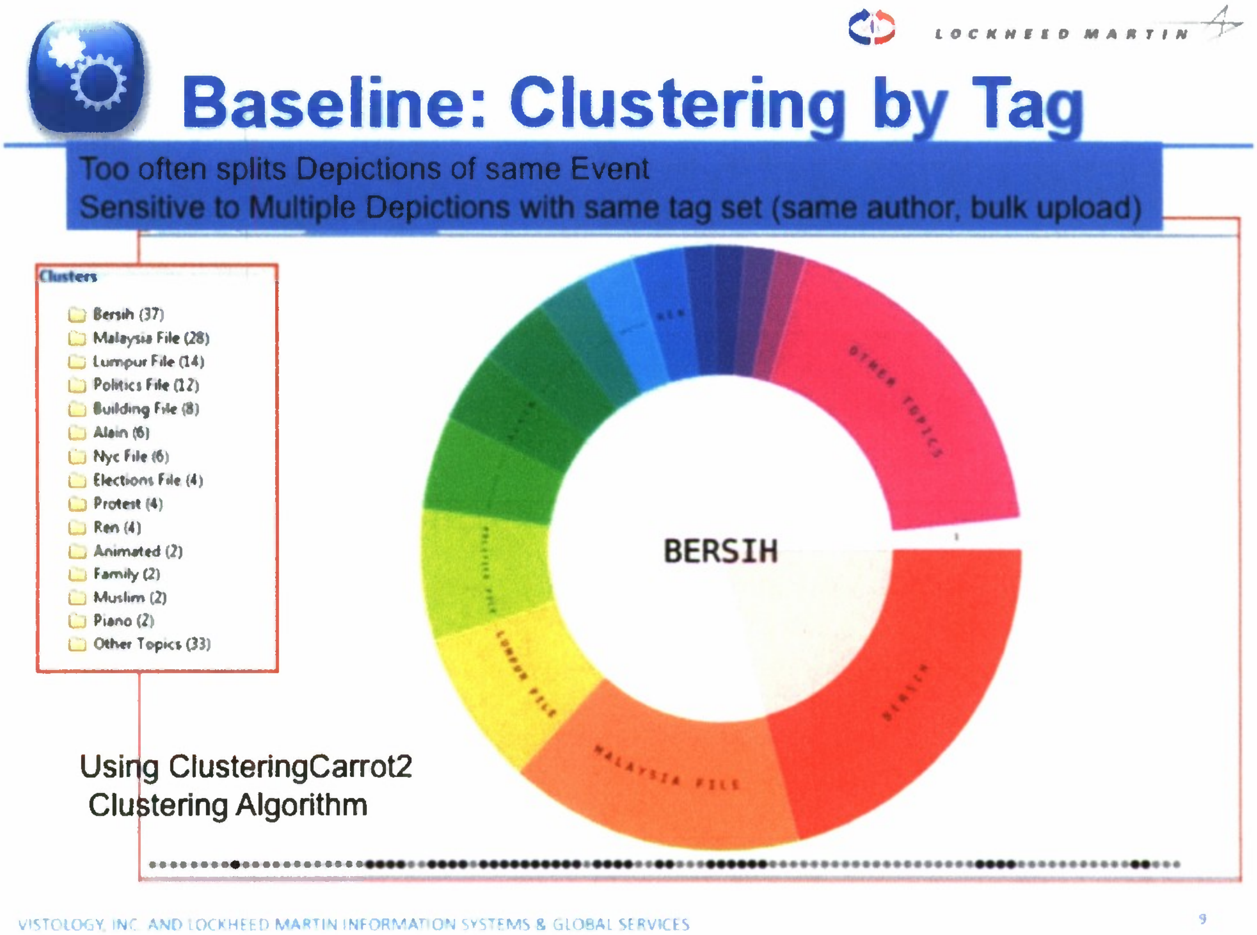
For purposes of evaluation, we compared our What/Where/When event depiction correlation with depiction correlation based on clustering tag sets. We used a set of 207 depictions from Flickr, YouTube and Blogger, described above in Section 3, as a test set.

First, using the Carrot 2 clustering engine (Stefanowski & Weiss, 2003; see also carrot2.org), we clustered 164 Flickr and YouTube-only depictions based only on their tag sets. The clusters are depicted graphically in Figure 4.

Table 5 YouTube and Flickr Clusters by Size

Cluster Label	Number of Depictions (164 total)
Bersih	37
Malaysia	28
Lumpur	14
Politics	12
Building	8
Alain	6
NYC	6
Elections	4
Protest	4
Ren	4
Animated	2
Family	2
Muslim	2
Piano	2
Unclustered	33

The use of a clustering algorithm over the tag sets, rather than the semantics-based approach described above, results in a splitting the depictions into finer bins than is warranted. In these results, for example, we see that a cluster set based on the tag "Bersih" and a cluster set based on the tag "Malaysia" are disjoint, even though these depict the same Bersih rally in Kuala Lumpur, the capital of Malaysia. Using the largest cluster for the Bersih event, the clustering method produces 100% precision, but 35.2% recall (based on the largest cluster). Similarly, the climbing related depictions are split into clusters for "building", "Nyc", "Alain (Robert)" (one of the climbers) and "Ren(aldo Clarke)" (the other climber), yielding precision of 100% but recall of 44.4%. A large number of depictions (33) are not assigned to any cluster.



LOCKHEED MARTIN

Baseline: Clustering by Tag

Too often splits Depictions of same Event
Sensitive to Multiple Depictions with same tag set (same author, bulk upload)

Clusters

- □ Bersih (37)
- □ Malaysia File (28)
- □ Lumpur File (14)
- □ Politics File (12)
- □ Building File (8)
- □ Alain (6)
- □ Nyc File (6)
- □ Elections File (4)
- □ Protest (4)
- □ Ren (4)
- □ Animated (2)
- □ Family (2)
- □ Muslim (2)
- □ Piano (2)
- □ Other Topics (33)

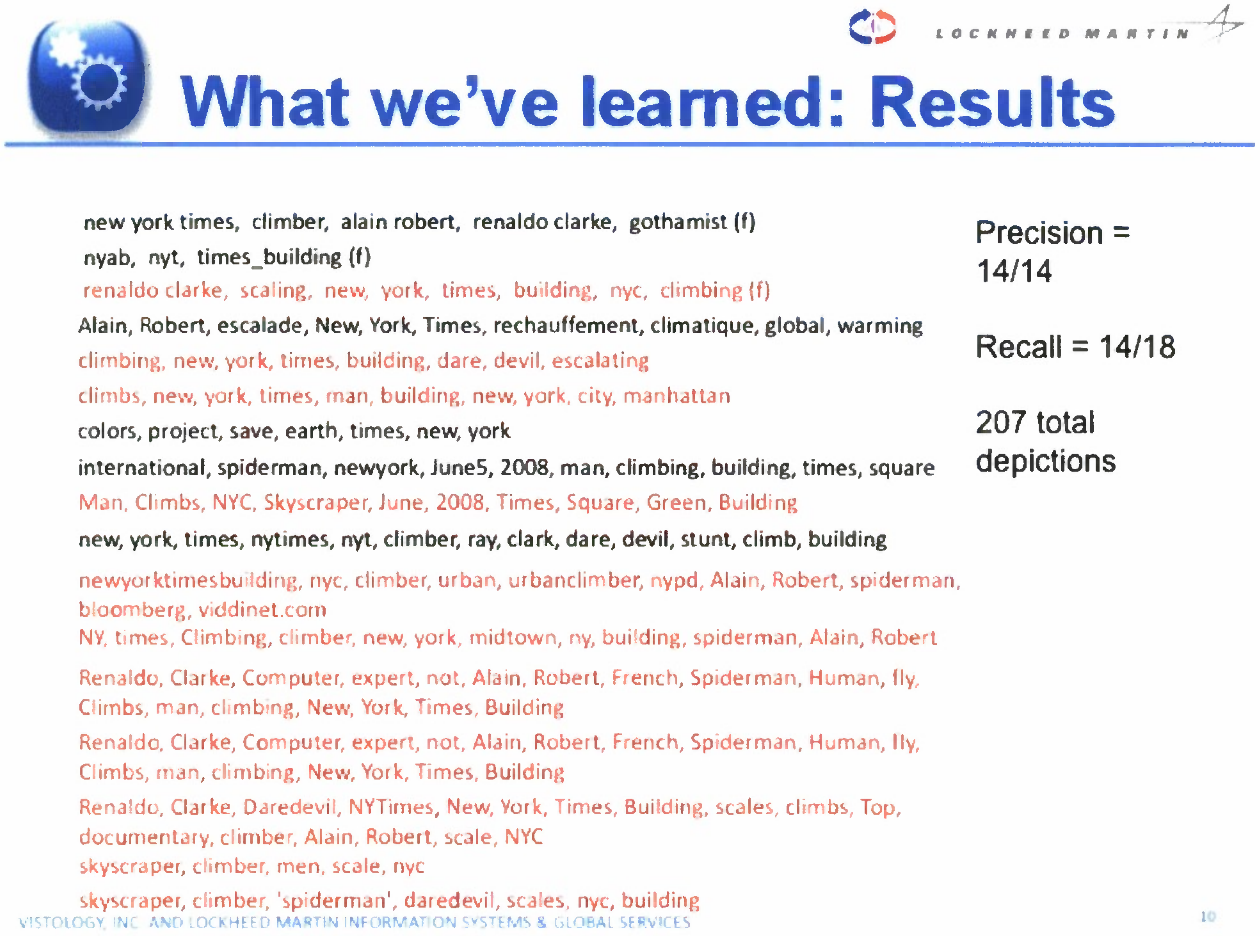
Using ClusteringCarrot2
Clustering Algorithm

VISTOLOGY, INC. AND LOCKHEED MARTIN INFORMATION SYSTEMS & GLOBAL SERVICES

9

Figure 4 Clustering Depictions by Tags

Using the What/Where/When semantics-based method outlined above, we assigned 14 of 18 depictions of the New York Times building climbing event as depictions of the same event, yielding a precision of 100% and a recall of 78% (Figure 5). For the Bersih rally depictions, the event depiction correlation method results in 100% precision and 62% recall. The semantic method failed to correlate 39 depictions of the Bersih rally event that were tagged with just the term "Bersih march" (Figure 6). The term 'march' failed to match with the T4GM category "Demonstration", with synonyms "protest", "rally", and so on, so those depictions, all bulk tagged, are not inferred to be depict the same event as the others.



What we've learned: Results

new york times, climber, alain robert, renaldo clarke, gothamist (f)

nyab, nyt, times_building (f)

renaldo clarke, scaling, new, york, times, building, nyc, climbing (f)

Alain, Robert, escalade, New, York, Times, rechauffement, climatique, global, warming

climbing, new, york, times, building, dare, devil, escalating

climbs, new, york, times, man, building, new, york, city, manhattan

colors, project, save, earth, times, new, york

international, spiderman, newyork, June5, 2008, man, climbing, building, times, square

Man, Climbs, NYC, Skyscraper, June, 2008, Times, Square, Green, Building

new, york, times, nytimes, nyt, climber, ray, clark, dare, devil, stunt, climb, building

newyorktimesbuilding, nyc, climber, urban, urbanclimber, nypd, Alain, Robert, spiderman, bloomberg, viddinet.com

NY, times, Climbing, climber, new, york, midtown, ny, building, spiderman, Alain, Robert

Renaldo, Clarke, Computer, expert, not, Alain, Robert, French, Spiderman, Human, fly, Climbs, man, climbing, New, York, Times, Building

Renaldo, Clarke, Computer, expert, not, Alain, Robert, French, Spiderman, Human, lly, Climbs, man, climbing, New, York, Times, Building

Renaldo, Clarke, Daredevil, NYTimes, New, York, Times, Building, scales, climbs, Top, documentary, climber, Alain, Robert, scale, NYC

skyscraper, climber, men, scale, nyc

skyscraper, climber, 'spiderman', daredevil, scales, nyc, building

Precision = 14/14

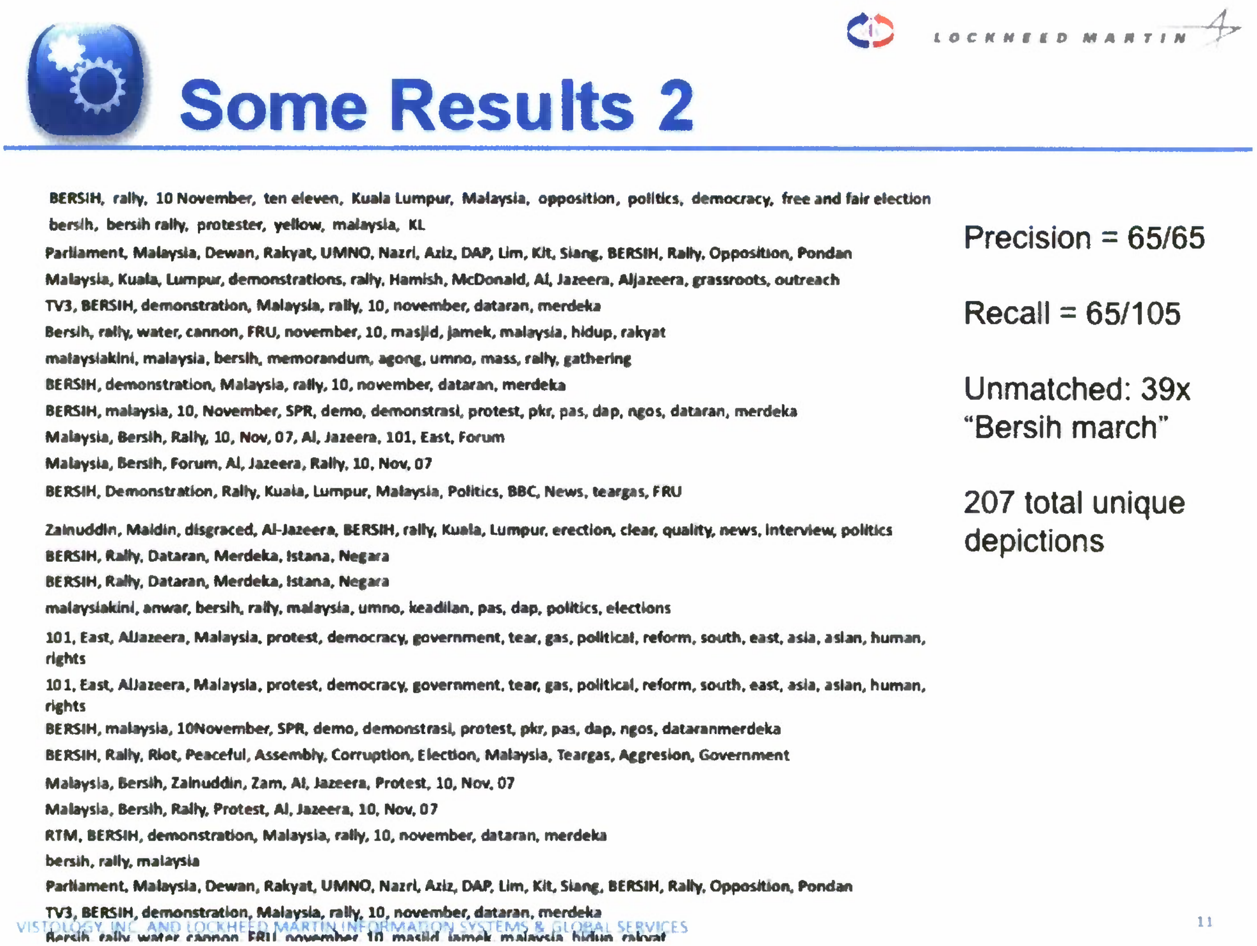
Recall = 14/18

207 total depictions

VISTOLOGY, INC. AND LOCKHEED MARTIN INFORMATION SYSTEMS & GLOBAL SERVICES

10

Figure 5 Results (1): Correlating Depictions of New York Times Building climbing event. Tagsets in red are correlated. They separate tagsets for four depictions not correlated. Flickr tagsets are marked (f); the rest are from YouTube.



LOCKHEED MARTIN

Some Results 2

BERSIH, rally, 10 November, ten eleven, Kuala Lumpur, Malaysia, opposition, politics, democracy, free and fair election
bersih, bersih rally, protester, yellow, malaysia, KL
Parliament, Malaysia, Dewan, Rakyat, UMNO, Nazri, Aziz, DAP, Lim, Kit, Siang, BERSIH, Rally, Opposition, Pondan
Malaysia, Kuala, Lumpur, demonstrations, rally, Hamish, McDonald, Al, Jazeera, Aljazeera, grassroots, outreach
TV3, BERSIH, demonstration, Malaysia, rally, 10, november, dataran, merdeka
Bersih, rally, water, cannon, FRU, november, 10, masjid, jamek, malaysia, hidup, rakyat
malaysiakini, malaysia, bersih, memorandum, agong, umno, mass, rally, gathering
BERSIH, demonstration, Malaysia, rally, 10, november, dataran, merdeka
BERSIH, malaysia, 10, November, SPR, demo, demonstrasi, protest, pkr, pas, dap, ngos, dataran, merdeka
Malaysia, Bersih, Rally, 10, Nov, 07, Al, Jazeera, 101, East, Forum
Malaysia, Bersih, Forum, Al, Jazeera, Rally, 10, Nov, 07
BERSIH, Demonstration, Rally, Kuala, Lumpur, Malaysia, Politics, BBC, News, teargas, FRU
Zainuddin, Maidin, disgraced, Al-Jazeera, BERSIH, rally, Kuala, Lumpur, erection, clear, quality, news, interview, politics
BERSIH, Rally, Dataran, Merdeka, Istana, Negara
BERSIH, Rally, Dataran, Merdeka, Istana, Negara
malaysiakini, anwar, bersih, rally, malaysia, umno, keadilan, pas, dap, politics, elections
101, East, AlJazeera, Malaysia, protest, democracy, government, tear, gas, political, reform, south, east, asia, asian, human, rights
101, East, AlJazeera, Malaysia, protest, democracy, government, tear, gas, political, reform, south, east, asia, asian, human, rights
BERSIH, malaysia, 10November, SPR, demo, demonstrasi, protest, pkr, pas, dap, ngos, dataranmerdeka
BERSIH, Rally, Riot, Peaceful, Assembly, Corruption, Election, Malaysia, Teargas, Aggresion, Government
Malaysia, Bersih, Zainuddin, Zam, Al, Jazeera, Protest, 10, Nov, 07
Malaysia, Bersih, Rally, Protest, Al, Jazeera, 10, Nov, 07
RTM, BERSIH, demonstration, Malaysia, rally, 10, november, dataran, merdeka
bersih, rally, malaysia
Parliament, Malaysia, Dewan, Rakyat, UMNO, Nazri, Aziz, DAP, Lim, Kit, Siang, BERSIH, Rally, Opposition, Pondan
TV3, BERSIH, demonstration, Malaysia, rally, 10, november, dataran, merdeka
Bersih, rally, water, cannon, FRU, november, 10, masjid, jamek, malaysia, hidup, rakyat

Precision = 65/65

Recall = 65/105

Unmatched: 39x
"Bersih march"

207 total unique
depictions

VISTOLOGY INC. AND LOCKHEED MARTIN INFORMATION SYSTEMS & GLOBAL SERVICES 11

Figure 6 Results (II) of Event Correlation of Bersih Rally event depictions

5 Conclusions

In this study, we have demonstrated that depictions of the same event can be more easily identified within and across various media through semantic processing of folksonomy tags assigned to the depictions independently by various users, using no standardized vocabulary. The semantic processing outlined here identified depictions of the same event by first processing folksonomy tags in order to identify tags that depicted locations (Where), activity types (What) and event dates (When). Inference was used to determine synonymous terms (using the SKOS vocabulary), compute temporal and geospatial distances, and normalize partial dates. The semantic approach yielded greater recall of identical events, compared with the clustering method.

Nevertheless, although What/Where/When semantics provides a good first approximation of event depiction identity, much improvement could be made. Our method relies on the existence of terms corresponding to tags in semantic resources such as thesauri. Having identified a set of correlated event depictions using What/Where/When semantics, significant terms that are not in these semantic resources (e.g. "Bersih", the name of the organization organizing the rally depicted) could potentially be used to extend the set of correlated event depictions even when they don't coincide in What, Where, and When. In future work, we plan to investigate such techniques. Semantic processing provides a better starting point for increased recall in identifying event depictions than non-semantic methods such as clustering.

6 References

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., and Ives, Z. (2008). Dbpedia: A nucleus for a web of open data. ISWC 2007. pages 722-735.

Berendt, B., & Hanser, C. (2007). Tags are not metadata, but "just more content" - to some people. International Conference on Weblogs and Social Media 2007. Boulder, CO.

Blogger.com. Blogger APIs. http://code.google.com/apis/blogger/ (retrieved January, 2010).

Carrot 2. Clustering Carrot Search Engine. http://www.carrot2.org (retrieved January, 2010).

Davidson, D. (1967). The logical form of action sentences. Chapter 6. *Essays on Actions and Events*, 1(9):105-149.

Flickr.com, Flickr API Documentation. http://www.flickr.com/services/api/ (retrieved January, 2010).

Kokar, M., Matheus, C., Letkowski, J., Baclawski, K., Kogut, P., (2004). Association in Level 2 Fusion. In Proc of SPIE Conference on Multisensor, Multisource Information Fusion, Orlando, FL., April 2004

Lerman, K. and Jones, L. (2007). Social browsing on Flickr. ICWSM, 2007.

Matheus, C., Dionne, B., Parent, D., Baclawski, K., and Kokar, M.. (2007) *BaseVISor: A Forward-Chaining Inference Engine Optimized for RDF/OWL Triples.* In Digital Proceedings of the 5th International Semantic Web Conference, ISWC 2006, Athens, GA, Nov. 2006.

OpenCalais. OpenCalais Documentation. http://www.opencalais.com/documentation/opencalais-documentation (retrieved January, 2010).

Rocket Software. Rocket Aerotext. http://www.rocketsoftware.com/products/rocket-aerotext (retrieved January, 2010).

SKOS (2008). Simple Knowledge Organization System Reference Editors' Draft 1 October 2008 $Revision: 1.73 $.")
http://www.w3.org/2006/07/SWD/SKOS/reference/20081001/

Spiteri, L.F. (2007) "The structure and form of folksonomy tags: The road to the public library catalog" Information Technology and Libraries

Stefanowski, J., & Weiss, D. (2003). Carrot^2 and Language Properties in Web Search Results Clustering. *LECTURE NOTES IN COMPUTER SCIENCE , 2663.*

T4GM.info. About T4GM.Info. http://www.t4gm.info/about (retrieved January, 2010).

Ulicny, B., Matheus, C., Kokar, M., (2010) Metrics for modeling a Social-Political Blogosphere: A Malaysian Case Study. IEEE Internet Computing, Special Issue on Social Computing in the Blogosphere. March/April, 2010.

YouTube.com, YouTube APIs and Tools. http://code.google.com/apis/youtube/overview.html (retrieved January, 2010).