Ohio Center of Excellence on Knowledge-Enabled Computing (Kno.e.sis)

scientia potentia est

Knowledge is Power
Francis Bacon, 1597
...established and popularized deductive methodologies for scientific inquiry
Kno.e.sis’ leadership in semantic processing will contribute to basic theory about computation and cognitive systems, and address pressing practical problems associated with productive thinking in the face of an explosion of data.

Kno.e.sis: from information to meaning.
Significant Infrastructure
Exceptional Regional Collaboration

- At least 6 active projects with AFRL/WPAFB
- Human Effectiveness Directorate
- Sensors Directorate
Exceptional National Collaboration

- Univ. of Georgia, Stanford, Purdue, OSU, Ohio U., Indiana U. UC-Irvine, Michigan State U., Army
- Microsoft, IBM, HP, Yahoo
• W3C Member: WSDL-S/SAWSDL, SA-REST, SSN, ...
• U. Manchester, TU-Copenhagen, TU-Delft, DERI (Ireland), Max-Planck Institute, U. Melbourne, CSIRO, DA-IICT (India)
Current active funds of ~$10 million (supporting research of 15 faculty and 45+ funded grad students & postdocs).

NIH, NSF, AFRL, ...
MSR, IBM-R, HP Labs, Google
• Over 2,000 citations per faculty, around 1,000 refereed publications – comparable to any excellent group; granted 74 PhDs, $50 million in cumulative funding

• Prof. Sheth among the most cited Computer Science authors in the world today (h-index 65, 10th in citation in WWW area: cf Microsoft Academic Search)

• Prof. Bennett & Flach’s paper declared as one of most influential papers published in over 50 years in Journal of Human Factors; Prof. Raymer’s paper was cited in a US Supreme Court decision

• Kno.e.sis has attracted top-notch faculty

• High quality funding: NIH, NSF, AFRL.....innovation grants: Microsoft Research, Google, IBM Research, HP labs

• Entrepreneurship experience – launched several companies
• Meena Nagarajan gave a **keynote** at an international workshop—unheard of for a PhD student.
• Satya, Cory, Karthik **organized international workshops**
• Satya joining CWRU (tenure track), Meena – IBM Almaden
• Six of the senior PhD students: 80+ papers, 40+ program committees, contributed to winning NIH and NSF grants.
• Students interned at & collaborated with the very best places: Microsoft Research, Yahoo! Research, IBM Research, HP Labs, NLM, Accenture Labs, …and filed for 6 patents in 3 years
Computing for Human Experience:
Semantics *empowered* Sensors, Services, and Social Computing on ubiquitous Web

Amit Sheth
LexisNexis Ohio Eminent Scholar
Wright State University, Dayton OH
http://knoesis.org

Thanks: Meena, Cory, Kats & Kno.e.sis team
• Microsoft purchased Powerset in 2008
• Apple purchased Siri [Apr 2010]
  – “Once Again The Back Story Is About Semantic Web”
• Google buys Metaweb [June 2010]...” Google Snaps Up Metaweb in Semantic Web Play”
• Facebook OpenGraph, Twitter annotation ...”another example of semantic web going mainstream” “Google, Twitter and Facebook build the semantic web”
• RDFa adoption ....Search engines (esp Bing) are about to introduce domain models and (all) use of background knowledge/structured databases with large entity bases

Kno.e.sis is the largest US academic group in terms of # of faculty and PhD students in Semantic Web/Web 3.0 area (semantics enhanced services, cloud, social and sensor computing/Webs)
Semantic Search etc.
A Bit of History

• SYSTEM AND METHOD FOR CREATING A SEMANTIC WEB AND ITS APPLICATIONS IN BROWSING, SEARCHING, PROFILING, PERSONALIZATION AND ADVERTISING [Filed 3/2000, Granted 5/2001]

• More in this 2000 keynote: Semantic Web and Information Brokering: Opportunities, Commercialization and Challenges
Computing for Human Experience

Knowledge Discovery

Sensor Web

Social Computing

Web Services

Cloud Computing
Semantics as core enabler, enhancer @ Kno.e.sis
imagine
imagine when
meets

Farm Helper
with this

- Latitude: 38° 57’36” N
- Longitude: 95° 15’12” W
- Date: 10-9-2007
- Time: 1345h
that is sent to

Sensor Data Resource

Weather Resource

Weather Data

Location

Date/Time

Lawrence, KS

Farm Helper

Lat-Long

Services Resource

Geocoder

Structured Data Resource

Agri DB

Soil Survey

Soil Information
Pest information
and
Six billion brains
imagination today
impacts our experience tomorrow
COMPUTING FOR HUMAN EXPERIENCE
Seamless integration of technology with life*

“The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life …

Machines that fit the human environment instead of forcing humans to enter theirs …”

Mark Weiser, The Computer for the 21st Century (Ubigcomp vision)

“…technology that will allow us to combine what we can do on the Internet with what we do in the physical world.”

Ian Pearson in Big data: The next Google

* From Devices to Ambient Intelligence
Ubicomp: Mark Wisner and others

**Intelligence @ Interface**: Gruber – “the system knows about us, our information, and our physical environment. With knowledge about our context, an intelligent system can make recommendations and act on our behalf.”
What is CHE? Beyond better human interaction

• Computing for human experience will employ a suite of technologies to nondestructively and unobtrusively complement and enrich normal human activities, with minimal explicit concern or effort on the humans’ part.

• Features
  – Seamless - nondestructive and unobtrusive, with minimal explicit concern or effort on part of humans
  – anticipatory, knowledgeable, intelligent, implicit, ubiquitous
  – Encompasses: Mobile (ubiquitous) Web, Sensor (multisensory and participatory) Web, Social Web (collective intelligence and wisdom of the crowd), multimodal
## Learning from a number of exciting visions

<table>
<thead>
<tr>
<th>Name</th>
<th>Vision/Project</th>
<th>Year(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vannevar Bush</td>
<td>Trailblazing, Memex (<a href="https://books.google.com/books?id=G4mKAAAAMBAJ">As We May Think</a>)</td>
<td>1945</td>
</tr>
<tr>
<td>Mark Weisner</td>
<td>Ubicomp (<a href="https://books.google.com/books?id=gQJn9g5jA5UC">The Computer for 21st Century</a>)</td>
<td>1991</td>
</tr>
<tr>
<td>Zelkha, Epstein</td>
<td>Ambient Intelligence</td>
<td>1998-2000</td>
</tr>
<tr>
<td>Tim Berners-Lee</td>
<td>Semantic Web</td>
<td>1999-2000</td>
</tr>
<tr>
<td>Gordon Bell et al</td>
<td><a href="https://example.com/mylifebits">MyLifeBits</a></td>
<td>2002</td>
</tr>
<tr>
<td>Jonathan Rossiter</td>
<td><a href="https://example.com/humanist">Humanist Computing</a></td>
<td>2003</td>
</tr>
<tr>
<td>Ramesh Jain</td>
<td><a href="https://example.com/experiential">Experiential Computing</a></td>
<td>2003</td>
</tr>
<tr>
<td>ITU</td>
<td><a href="https://example.com/iot">The Internet of Things</a></td>
<td>2005</td>
</tr>
<tr>
<td>Ramesh Jain</td>
<td><a href="https://example.com/event">Event Web</a></td>
<td>2006-2008</td>
</tr>
<tr>
<td>A. Sheth et al</td>
<td><a href="https://example.com/relationship">Relationship Web</a> (also, Spatio-temporal-thematic analysis, Semantic Sensor Web)</td>
<td>2006-2008</td>
</tr>
<tr>
<td>Tom Gruber</td>
<td><a href="https://example.com/ii">Intelligence @Interface</a></td>
<td>2008</td>
</tr>
</tbody>
</table>
Web (and associated computing) is evolving

Web of pages
- text, manually created links
- extensive navigation

1997

Web of databases
- dynamically generated pages
- web query interfaces

2007

Web of resources
- data, service, data, mashups
- 4 billion mobile computing

Web of people, Sensor Web
- social networks, user-created casual content
- 40 billion sensors

Computing for Human Experience
Web as an oracle / assistant / partner
- "ask the Web": using semantics to leverage text + data + services
- Powerset

Enhanced Experience,
Tech assimilated in life

Situations, Events

Objects

Patterns

Keywords

Enhanced Experience,
Tech assimilated in life

Situations, Events

Objects

Patterns

Keywords

Semantic Technology

Used
Four enablers of CHE

• Bridging the Physical/Digital Divide
• Elevating Abstractions That Machines & Humans Understand: signals to observations to perception
• From Social Perception to Semantics (meaningful to other humans/observers and machine – shared, computable; crowd sourcing)
• Semantics at an Extraordinary Scale

More in Computing for Human Experience, IEEE IC, Jan-Feb 2010.
Psyleron’s Mind-Lamp (Princeton U), connections between the mind and the physical world.

Neuro Sky's mind-controlled headset to play a video game.

Sixth Sense- Gesture Computing and wearable device with a projector for deep interactions with the environment.

IoT
Emotion Sensors
Wearable Sensors
Body Area Networks
Sensors everywhere ..sensing, computing, transmitting

• 2009: 1.1 billion PCs, 4 billion mobile devices, 40+ billion mobile sensors
  (Nokia: Sensing the World with Mobile Devices)

• 6 billion intelligent sensors
  – informed observers, rich local knowledge

Christmas Bird Count
Relevant happenings today

• that all objects, events and activities in the physical world have a counterpart in the Cyberworld (IoT)

• multi-facted context of real world is captured in the cyberworld (multilevel & citizen sensors/participatory sensing)

• each object, event, and activity is represented
  – with semantic annotations (semantic sensor web)

• for a chosen context, with an ability to explicate and associate variety of relationships and events (background knowledge, Relationship Web, EventWeb)

• appropriate reasoning and human/social interaction are available and applied, insights extracted (semantic web, social semantic web, experiential computing)

• Activity anticipated/answers obtained/ decisions reached/ communicated/applied
Elevating Abstractions that Machines & Humans Understand: signals to observations to perception that lead to semantics (provide meaning and understanding)
**People Web**
(human-centric)

- **Observation**
  (senses)

- **Communication**
  (language)

- **Perception**
  (cognition)

**Sensor Web**
(machine-centric)

- **Observation**
  (sensors)

- **Communication**
  (services)

- **Perception**
  (analysis)
Enhanced Experience (humans & machines working in harmony)

Observation

Communication

Perception

Semantics to improve communication about shared spaces, events,...

Semantics for shared conceptualization and interoperability between machine and human
Semantic Sensor Web Infrastructure

Semantic annotation of SWE

```
<swc:component rdfa:about="time_1"
    rdfa:instanceof="time:Instant">
  <swc:time rdfa:property="xs:date-time">
    2008-03-08T05:00:00
  </swc:time>
</swc:component>
<swc:component>
  <swc:value name="satellite-data"
    rdfa:about="Dayton"
    rdfa:instanceof="geo:City">
    0011000111001111 ... 
  </swc:value>
```

Ontological Knowledge
space, time and theme
<om:Observation>
  <om:samplingTime><gml:TimeInstant>...</gml:TimeInstant></om:samplingTime>
</om:Observation>

<om:procedure xlink:role="http://www.w3.org/2009/Incubator/ssn/ontologies/SensorOntolgy.owl#Sensor"
xlink:href="http://www.w3.org/2009/Incubator/ssn/ontologies/SensorOntolgy.owl#sensor_xyz"/>

<om:observedProperty xlink:href="http://www.w3.org/2009/Incubator/ssn/ontologies/SensorOntolgy.owl#temperature"/>

<featureOfInterest xlink:href="http://sws.geonames.org/5758442"/>

<om:result uom="http://www.w3.org/2009/Incubator/ssn/ontologies/SensorOntolgy.owl#fahrenheit">42.0</om:result>
</om:Observation>
Semantic Sensor ML – Adding Ontological Metadata

Mike Botts, "SensorML and Sensor Web Enablement," Earth System Science Center, UAB Huntsville
Scenario of context
(Semantic Sensor Web demo)

More: Semantic Sensor Web
Active Perception (Sensing – Observation – Perception) and role of ontologies and background knowledge for Situational Awareness
To enable situation awareness on the Web, we must utilize *abstractions* capable of representing observations and perceptions generated by either people or machines.
For example, both people and machines are capable of observing qualities, such as redness.

* Formally described in a sensor/observation ontology
With the help of sophisticated inference, both people and machines are also capable of perceiving entities, such as apples.

* Formally described in a perception ontology
The ability to perceive is afforded through the use of *background knowledge*. For example, knowledge that *apples are red* helps to infer an apple from an observed quality of redness.

* Formally described in a domain ontology
The ability to perceive *efficiently* is afforded through the cyclical exchange of information between observers and perceivers.

* Traditionally called the Perception Cycle (or Active Perception)
Integrated together, we have an abstract model – capable of situation awareness – relating observers, perceivers, and background knowledge.

What is new? Relevant background knowledge/ontologies are increasingly available or possible to create. Domain independent ontologies are being developed or exist… Web, scale….
Let’s review an example execution of the perception cycle, utilizing background knowledge from the weather domain.
Background knowledge from weather ontology

Quality

- Freezing Temperature
- Not Freezing Temperature
- Snow Precipitation
- Rain Precipitation
- No Precipitation
- High Wind Speed
- Low Wind Speed

inheresIn

Entity

- Blizzard
- Flurry
- Rain Storm
- Rain Shower
- Clear
Example execution of the Perception Cycle

- **Percept:** high wind speed, snow precipitation, clear, blizzard, flurry, rain shower, rain storm

- **Focus:**
  - **Perceiver** perceives
  - **Observer** observes

- **Inheres in:**
  - sensor-observation ontology
  - perception ontology
  - perceptual theory

- **Observed qualities:**
Perceiver sends ‘wind-speed’ focus to observer

Observer

observes

Quality

Percept:

Focus:
wind-speed

Perceiver

perceives

Entity

observed qualities

inheres in

perceptual theory

clear
blizzard
flurry
rain shower
rain storm
Observer observes ‘high wind-speed’

Focus: wind-speed
Percept: high wind-speed
Entity: clear, blizzard, flurry, rain shower, rain storm

observes

inheres in

observed qualities
high wind-speed

perceptual theory
Observer sends ‘high wind-speed’ percept to perceiver

Percept: high wind-speed

Focus: wind-speed

Quality: high wind-speed

Inheres in:
- clear
- blizzard
- flurry
- rain shower
- rain storm

Perceptual theory

Observer sends 'high wind-speed' percept to perceiver
Perceiver perceives either ‘clear’, ‘blizzard’, or ‘rain storm’

Percept: high wind-speed

Focus: wind-speed

Inheres in perceptual theory:
- clear
- blizzard
- rain storm

Observer observes:
- high wind-speed

Perceiver perceives:
- clear
- blizzard
- rain storm

Entities:
- Solar Energy
- Wind Energy
- Water Energy
- Advantages
- Disadvantages
- Wind Speed
- Clear Sky
Perceiver sends ‘precipitation’ focus to observer

Percept: high wind-speed

Focus: precipitation

observes

observed qualities

high wind-speed

inheres in

perceptual theory

clear blizzard

rain storm

Perceiver

Perceiver sends 'precipitation' focus to observer

Observer

observes

Quality

observed qualities

high wind-speed

inheres in

perceptual theory

clear blizzard

rain storm

Perceiver

Percept: high wind-speed

Focus: precipitation

perceives

Entity

perceptual theory

clear blizzard

rain storm

Observer
Observer observes ‘snow precipitation’

Percept: high wind-speed

Focus: precipitation

observes

Quality

inheres in

perceptual theory

clear
blizzard

rain storm

observed qualities

high wind-speed
snow precipitation
Observer sends ‘snow precipitation’ percept to perceiver.

Observer

observes

Quality

observed qualities

high wind-speed
snow precipitation

perceptual theory

perceives

Entity

Focus:
snow precipitation

inheres in

clear
blizzard

rain storm

Percept:
snow precipitation

Perceiver
Perceiver perceives 'blizzard'

Observer \(\text{observes}\) Quality

Percept: snow precipitation

Focus: precipitation

inheres in

observed qualities

high wind-speed snow precipitation

Perceiver \(\text{perceives}\) Entity

perceptual theory

blizzard
Reference Ontologies

Sensor-Observation Ontology
• SSN-XG Sensor Ontology [1]
• Kuhn’s Ontology of Observation [2]
• etc. [3]

Perception Ontology
• Kno.e.sis’s Perception Ontology [4]

Domain Ontology: Weather
• Kno.e.sis’s Weather Ontology (derived from NOAA) [5]
Implementation of Perception Cycle

Trust

Strengthened Trust
• (What?) Trust is the psychological state comprising a willingness to be vulnerable in expectation of a valued result.
• (Why?) To act on a decision, it’s important to have confidence in the information from which it was derived.
• (How?) Through reputation, past behavior can be used to predict future behavior.

*Ontology of Trust*, Huang and Fox, 2006
Josang et al’s Decision Trust
• Trust values converge with the increasing number of observations.

![Graph showing Mean of beta vs. Time for stnID = SBE]
What next?

Trillions+ of observations (data) from human body to all devices and everything on the Globe -> What’s meaningful to human experience?

Body Sensor NW: http://www.enterprise.mtu.edu/im/projects.html
http://www.opengeospatial.org/ogc/markets-technologies/swe
mumbai, india
november 26, 2008
another chapter in the war against civilization
and
Bombay Blast by Vinu
NSG men surrounding the building. #mumbai
Mumbai, India
Nov 27, 2008 11:59 PM GMT · Reply · View Tweet

@sameer_gupta: Why is Australian government issuing advisory to not to visit Mumbai? The govt. is nuts cos it gives more confidence to ppl spreading terror
Mumbai, India
Nov 27, 2008 11:55 PM GMT · Reply · View Tweet

@sameer_gupta: CNBC is reporting that so far 7 terrorists have been killed in Taj Mahal. #mumbai
Mumbai, India
Nov 27, 2008 11:46 PM GMT · Reply · View Tweet

Mumbai: #mumbai Situation Report: mostly static. Cleanup operations at Taj, Oberoi and Nariman still in flux. Other reported attacks are rumors.
Mumbai, India
Nov 27, 2008 11:37 PM GMT · Reply · View Tweet

@sameer_gupta: If news 3 arrests is true. Its a gr8 achievement. it will help in investigation. Tough earlier news said only 1 is left in Taj. #mumbai
Mumbai, India
Nov 27, 2008 11:35 PM GMT · Reply · View Tweet

@sameer_gupta: News24 says that 3 terrorists are arrested from Taj. And they are linked with Lashkar-e-Toiba and one is Multan, Pakistan native. #mumbai
Mumbai, India
Nov 27, 2008 11:33 PM GMT · Reply · View Tweet
the world saw it

Through the eyes of the people
the world read it
Through the words of the people
PEOPLE told their stories to PEOPLE
A powerful new era in Information dissemination had taken firm ground
Making it possible for us to create a global network of citizens

Citizen Sensors – Citizens observing, processing, transmitting, reporting
SOCIAL PERCEPTIONS,
OBSERVATIONS, (SEMANTIC)
SENSE MAKING
Varied Social Perceptions lend to Varied Observations

The Health Care Reform Debate
Zooming in on Washington
RT @WestWingReport: Obama reminds the faith-based groups "we're neglecting to live up to the call" of being brother's keeper on #healthcare
Zooming in on Florida
• Social Components of content dictate how we perceive and process information
  – Textual Content
    • Latent crowd characteristics from content
  – Spatial, Temporal parameters
    • When, where the message originated
  – Poster demographics
    • Age, gender, socio-economic status..
Spatio Temporal and Thematic analysis

– What else happened “near” this event location?
– What events occurred “before” and “after” this event?
– Any message about “causes” for this event?
Spatial View….

Which tweets originated from an address near 18.916517°N 72.827682°E?

- **sameer_gupta**: At least 5-6 people held hostage in Nariman House. More than 50 NSG men surrounding the building. #mumbai
  - Mumbai, India
  - Nov 27, 2008 11:59 PM GMT · Reply · View Tweet

- **sameer_gupta**: Why is Australian government issuing advisory to not to visit Mumbai? The govt. is nuts cos it gives more confidence to ppl spreading terror
  - Mumbai, India
  - Nov 27, 2008 11:55 PM GMT · Reply · View Tweet

- **sameer_gupta**: CNBC is reporting that so far 7 terrorists have been killed in Taj Mahal. #mumbai
  - Mumbai, India
  - Nov 27, 2008 11:46 PM GMT · Reply · View Tweet

- **Mumbai**: #mumbai Situation Report: mostly static. Cleanup operations at Taj, Oberoi and Nariman still in flux. Other reported attacks are rumors.
  - Mumbai, India
  - Nov 27, 2008 11:37 PM GMT · Reply · View Tweet

- **sameer_gupta**: If news 3 arrests is true. Its a gr8 achievement. It will help in investigation. Tough earlier news said only 1 is left in Taj. #mumbai
  - Mumbai, India
  - Nov 27, 2008 11:35 PM GMT · Reply · View Tweet

- **sameer_gupta**: News24 says that 3 terrorists are arrested from Taj. And they are linked with Lashkar-e-Toiba and one is Multan, Pakistan native. #mumbai
  - Mumbai, India
  - Nov 27, 2008 11:33 PM GMT · Reply · View Tweet
Which tweets originated during Nov 27th, 2008, from 11PM to 12 PM.
Tweets originated from an address near 18.916517°N, 72.827682°E during time interval 27th Nov 2008 between 11PM to 12PM?
More meaningful spatio-temporal-thematic analysis

• Preserve social perceptions behind social data
• Extracting key phrases that describe an event
  – Separate user observations by time and space
  – Extract summaries / key phrases / n-grams
  – Weight local to global, most recent to least recent
TWITRIS : Twitter+Tetris

• Our attempt to help you keep up with citizen observations on Twitter
  – WHAT are people saying, WHEN, from WHERE

• Puts citizen reports in context for you by overlaying it with news, wikipedia articles!
Twitris demo

(Search “Twitris” on YouTube)
• Gathering and processing social observations
• Challenges with Casual text
  – Informal, Domain Dependent slangs, misspellings, non-grammatical
  – Redundancy (everyone is tweeting the same thing)
  – Variability (everyone is saying the same thing in many ways)
  – off-topic noise
• Weather conditions from tweets using key words (http://smalltalkapp.com/).
• Selecting Georgia, we get the tweets used to derive weather conditions.
• Primarily use Keywords without background knowledge.
a nickname for Hip-Hop/R&B singer Chris Brown

"Country Sunshine" is the name of a popular country song written by Dottie West in 1973
Using Domain Knowledge

Using Domain Knowledge to

• Overcome challenges with informal user-generated content
• Supplement statistical NLP / ML algorithms and techniques

Daniel Gruhl, Meenakshi Nagarajan, Jan Pieper, Christine Robson, Amit Sheth, Multimodal Social Intelligence in a Real-Time Dashboard System, to appear in a special issue of the VLDB Journal on "Data Management and Mining for Social Networks and Social Media"
Knowledge of domain helps in collecting and analyzing social observations.

Informal text – www.twitter.com

XTRABallots: Mousavi campaigner Mojtaba Gahestouni arrested in **Khuzestan** province http://is.gd/9YP5f (expand) #iranelection

**globalQZ**: #iranelection Christian Sci Monitor: 3rd anniv of disappearance of former FBI agt Robt Levinson on **Kish Island**, Iran http://tinyurl.com/yk7tgdd

PUBLIC HANGINGS IN **AHVAZ** http://j.mp/blv8Z8 #iran #iranelection #HumanRights

Another public execution advertised to happen in **Ahvaz** Saturday 13 Mar #IranElection @sayahcnn
Common Thematic Analysis Tasks on User-generated Content

- Entity identification and disambiguation
  - Context poor, Informal
  - Domain models to aid disambiguation
Geocoder (Reverse Geo-coding)

Address to location database

Vasant Vihar

Nariman House

Income Tax Office

Image Metadata
latitude: 18° 54’ 59.46″ N, longitude: 72° 49’ 39.65″ E

Structured Meta Extraction

Spatio-Temporal Analysis

Identify and extract information from tweets

Mumbai: #mumbai Situation Report: mostly static. Cleanup operations at Taj, Oberoi and Nariman still in flux. Other reported attacks are rumors.
Mumbai, India
Nov 27, 2008 11:37 PM GMT · Reply · View Tweet

sameer_gupta: Atleast 5-6 people held hostage in Nariman House. More than 50 NSG men surrounding the building. #mumbai
Mumbai, India
Nov 27, 2008 11:50 PM GMT · Reply · View Tweet
Domain Ontologies provide additional Context

Informal text, Context-poor utterances…

Supplement NL features used for NER with information from Domain models

“Ohh these sour times... rock!”
Common Thematic Analysis Tasks on User-generated Content

• Opinion Expressions
• “Your new album is wicked”
• Shallow NL Parse
• Look up: UrbanDictionary (slang dictionary, glossary and orientations)
Common Thematic Analysis Tasks on User-generated Content

• Spam / Off-topic Elimination
• Special type of spam: related to topic, not to application’s interests
  – Music Popularity applications
  – Spam: Paul McCartney’s divorce; Rihanna’s Abuse; Madge and Jesus
• Self-Promotions
  – “check out my new cool sounding tracks..”
  – Same (music) domain, similar keywords, harder to tell apart
• Standard Spam
  – “Buy cheap cellphones here..”
Spam Elimination using previous knowledge annotation cues

• Aggregate function
  – Phrases indicative of spam (regular expressions)
  – Rules over previous annotator results
    • if a spam phrase, artist/track name and a positive sentiment were spotted, the comment is probably not spam!
• Presenting social data in context is an important aspect of sense making
Example -- Social Media in Context

- SOYLENT GREEN and the HEALTH CARE REFORM
- Perceptions -> Observations -> Sense making -> Perceptions
People-Content-Network Analysis

- Sentiment
- Intention
- Cultural
- Behavioral
- Information Diffusion
- Influence
- Group formulation
Merging sensor and social data

- an image taken from a raw satellite feed
Realistic scenario

• an image taken by a camera phone with an associated label, “explosion.”
Realistic scenario

• Textual messages (such as tweets) using STT analysis
• Correlating to get
Increasingly popular social, mobile, and sensor webs exhibit these characteristics:

- spontaneous (arising suddenly)
- follow a period of rapid evolution, involving real-time or near real-time data, which requires continuous searching and analysis.
- many distributed participants with fragmented and opinionated information
- accommodate diverse viewpoints involving topical or contentious subjects.
- feature context colored by local knowledge as well as perceptions based on different observations and their sociocultural analysis.
The circle of knowledge life on the Web

- Background knowledge
  - Physical/Scientific principles
  - Historical Facts,
  - Content created by humans and through social processes

- The Web

- Information and knowledge extraction

- Discourse and analysis
Where does the Domain Knowledge come from?

For relatively static domains:
• Expert and committee based ontology creation ... works in some domains (e.g., biomedicine, health care,...)
• Community maintained knowledge-bases, dictionaries, ... (musicbrainz, IMDB, ....)

For rapidly evolving domains:
– How to create models that are “socially scalable”?
– How to organically grow and maintain this model?
KNOWLEDGE EXTRACTION
by crowd-sourcing

“Human Cognition” AND Psychology AND Neuroscience?
“Human Cognition” AND Psychology AND Neuroscience
“Human Cognition” AND Psychology AND Neuroscience

Reports that Khamenei personally ordered the banning of Shajarian’s Rabana #iranelection

Larijani: Parliament has the right to impeach Ahmadinejad, because he goes against our laws & constitution http://is.gd/eccz2 #iranelection

Int’l Campaign for Human Rights in Iran says IRIB “acts as arm of intelligence apparatus” http://is.gd/ehRxN #iranelection

From: Continuous Semantics … IEEE Internet Computing, Nov-Dec, 2010
Twitter

Query emerging events

(a)

Larijani: Parliament has the right to impeach Ahmadinejad, because he goes against our laws & constitution http://is.gd/eccz2 #iranelection

(b)

(c)

(d)

Larijani: Parliament has the right to impeach Ahmadinejad, because he goes against our laws & constitution http://is.gd/eccz2 #iranelection
Significant capabilities to come:

✧ Richer forms of relationships
  (relationships are at the heart of semantics)

✧ Richer support for events and situations

✧ Significant advances in semantics and knowledge-enriched linking of cyber-physical and social-technical systems
Suppose that we create a Web in which
- Each node is an event or object
- Each node may be connected to other nodes using
  - Referential: similar to common links that refer to other related information.
  - Spatial and temporal relationships.
  - Causal: establishing causality among relationships.
  - Relational: giving similarity or any other relationship.
  - Semantic or Domain specific:
    - Familial
    - Professional
    - Genetics,…
Karthik Gomada

Attended Google IO
Moscone Center, SFO
May 28-29, 2008

Amit Sheth

is_advised_by
Ph.D Student

Assistant Professor

Professor

Researcher

Research Paper

Journal

Conference

Location

published_in

published_in

has_location

Event

Spatio-temporal

Causal

Image Metadata

Domain Specific

Relational

Directs

kno.e.sis

Assistant Professor

Professor

Assistant Professor

Professor

Assistant Professor

Professor

Assistant Professor

Professor

Assistant Professor

Professor
Events in Iran Election, 2009

- Iran's tenth presidential election was held
- Basij paramilitary group raided Tehran University, injuring many.
- Guardian Council announced a partial recount
- More than 100,000 protesters held a candle-light vigil in Tehran
- Riot police broke up the main rally in Tehran with tear gas and live fire into the air.

- June 12: Election results announced: Ahmadinejad won
- June 13: Supporters of Mousavi started postesting
- June 14: Postest continued, Mousavi made his first post-election appearance.
- June 15: Iranian national football team wore green wristbands in support of Mousavi
- June 16: Another large protest occurred
- June 17: A young Iranian woman, Neda was shot by the Basij in Tehran
- June 18: Guardian Council certified the results of the controversial election.
- June 19: Guardian Council certified the results of the controversial election.
Online and offline worlds

• Computational abstractions to represent the physical world’s dynamic nature

• Merging online and offline activities
  – Connecting the physical world naturally with the online world

• What are natural operations on these abstractions?

• How do we detect these abstractions based on other abstractions and multimodal data sources?
Objects to Events

• If we move from this object mode to an event mode
  
  – A single user action or request or sensory observation could act as a cue for getting all (multi-modal) information associated with an event
  
  – If conditions change, systems could even modify their behavior to suit their changing view of the world

Today text is most prevalent, with increasing but disparate (non-integrated) image and video data, but human experience is event based (at higher levels of abstractions) formed based on multi-sensory, multi-perception (at lower level of abstraction) observations
On our way…

• We are already seeing efforts toward this larger goal
  – Social connections, interests, locations, alerts, comment
  – Mobile phone to social compass: LOOPT.com

Image credit - www.movilae.com
Internet of Things: “A world where inanimate objects communicate with us and one another over the network via tiny intelligent objects” - Jean Philippe Vasseur, NSSTG Systems
Direct Observation of or Participation in Events as a basis of knowledge
Entities and Events

**Entity**

- Name
- Attributes
- Processes (Services)

**Event**

- Duration
- Location
- Data-streams
- Adjacent States Related Links

Objects and Entities are static.

Events are dynamic.

EVENTS can be viewed as objects with time and space as their primary attributes.

Thanks – Ramesh Jain
Strategic Inflection Points

Documents on Web (Information)

Events on Web (Experience)

1995 2000 2005 2010

Immersive Experience
Contextual Search
Ubiquitous Devices
Semantic Search
Updates and alerts
Keyword Search

Thanks – Ramesh Jain
Challenges – Complex Events

• Formal framework to model complex situations and composite events
  – Those consisting of interrelated events of varying spatial and temporal granularity, together with their multimodal experiences

• What computational approaches will help to compute and reason with events and their associated experiences and objects?
1. Sensors observe environmental phenomena and nearby vegetation.

2. Observation analysis determines potential situation and effects.

- Through abductive reasoning, observation analysis perceives a possible storm as the best explanation hypothesis for observed phenomena.
- Through predictive deductive reasoning, observation analysis determines the effect on the crops, including the potential for the poisoning of the soil from salt carried from the ocean in the wind.
- Through query against a knowledge base of the agriculture domain, observation analysis determines that the best remedy
  - for saline soil is to “leach” the soil with excess irrigation water in order to ‘push’ the salts below the crop root zone,
  - for sodic soil is to add gypsum before leaching.
Example

1. Sensors observe environmental phenomena and nearby vegetation.

2. Observation analysis determines potential situation and effects.

3. System alerts nearby farmers of situation and possible remedy.

4. Farmer goes outside and looks at the sky and crops.

5. Farmer perceives high-winds and dark rain clouds over the ocean view and agrees with system perception.

6. Farmer calls children and neighbors to help take the necessary precautions to save the vegetables.
• Several key contributing research areas
  – Operating Systems, networks, sensors, content management and processing, multimodal data integration, event modelling, high-dimensional data visualization …. 

• Semantics and Semantic technologies can play vital role
  – In the area of processing sensor observations, the Semantic Web is already making strides
  – Use of core SW capabilities: knowledge representation, use of knowledge bases (ontologies, folkonomies, taxonomy, nomenclature), semantic metadata extraction/annotation, exploiting relationships, reasoning
THERE IS MORE HAPPENING AT KNO.E.SIS

http://knoesis.org

Also check out demos, systems at
http://knoesis.wright.edu/library/demos/
KNO.E.SIS MEMBERS – A SUBSET
Influential Works

• V. Bush, *As We May Think*, The Atlantic, July 1945. [*Memex, trail blazing*]

• Mark Weiser, *The Computer for the Twenty-First Century*, Scientific American, Sept 1991, 94-10. [*The original vision paper on ubicomp. Expansive vision albeit technical aspects focused on HCI with networked tabs, pads and boards.*]

• V. Kashyap and A. Sheth, *Semantics-based information brokering*. Third ACM Intl Conf on Information and Knowledge Management (CIKM94), Nov 29 - Dec 02, 1994. ACM, New York, NY. [*semantics based query processing (involving multiple ontologies, context, semantic proximity) across a federated information sources across the Web*]

• Abowd, Mynatt, Rodden, *The Human Experience*, Pervasive computing, 2002. [*explores Mark Wisner’s original ubicomp vision*]

• Jonathan Rossiter, *Humanist Computing: Modelling with Words, Concepts, and Behaviours*, in Modelling with Words, Springer, 2003, pp. 124-152 [*modelling with words, concepts and behaviours defines a hierarchy of methods which extends from the low level data-driven modelling with words to the high level fusion of knowledge in the context of human behaviours*]


• Amit Sheth, Sanjeev Thacker, and Shuchi Patel, Complex Relationship and Knowledge Discovery Support in the InfoQuilt System, VLDB Journal, 12 (1), May 2003, 2–27. [*complex semantic inter-domain (multi-ontology) relationships including causal relationships to enable human-assisted knowledge discovery and hypothesis testing over Web-accessible heterogeneous data*]

• Ramesh Jain, Toward EventWeb. IEEE Distributed Systems Online 8, 9, Sep. 2007. [a web of temporally related events… informational attributes such as experiential data in the form of audio, images, and video can be associated with the events]


Other Closely Related publications


• Amit Sheth and Matthew Perry, “Traveling the Semantic Web through Space, Time and Theme,” IEEE Internet Computing, 12, (no.2), February/March 2008, pp.81-86.

Interested in more background?

- Computing for Human Experience
- Continuous Semantics to Analyze Real-Time Data
- Semantic Modeling for Cloud Computing
- Citizen Sensing, Social Signals, and Enriching Human Experience
- Semantics-Empowered Social Computing
- Semantic Sensor Web
- Traveling the Semantic Web through Space, Theme and Time
- Relationship Web: Blazing Semantic Trails between Web Resources
- SA-REST: Semantically Interoperable and Easier-to-Use Services and Mashups
- Semantically Annotating a Web Service

Contact/more details: amit @ knoesis.org

Partial Funding: NSF (Semantic Discovery: IIS: 071441, Spatio Temporal Thematic: IIS-0842129), AFRL and DAGSI (Semantic Sensor Web), Microsoft Research and IBM Research (Analysis of Social Media Content), and HP Research (Knowledge Extraction from Community-Generated Content).