



---

# Lexical Link Analysis of the Haiti Earthquake Relief Operation Using Open Data Sources

Dr. Ying Zhao, Research Associate Professor, Naval Postgraduate School

Dr. Douglas J. MacKinnon, Research Associate Professor, Naval Postgraduate School

Dr. Shelley P. Gallup, Research Associate Professor, Naval Postgraduate School

16th ICCRTS, International Command and Control, Research and Technology Symposium, Québec City, Canada June 21–23, 2011



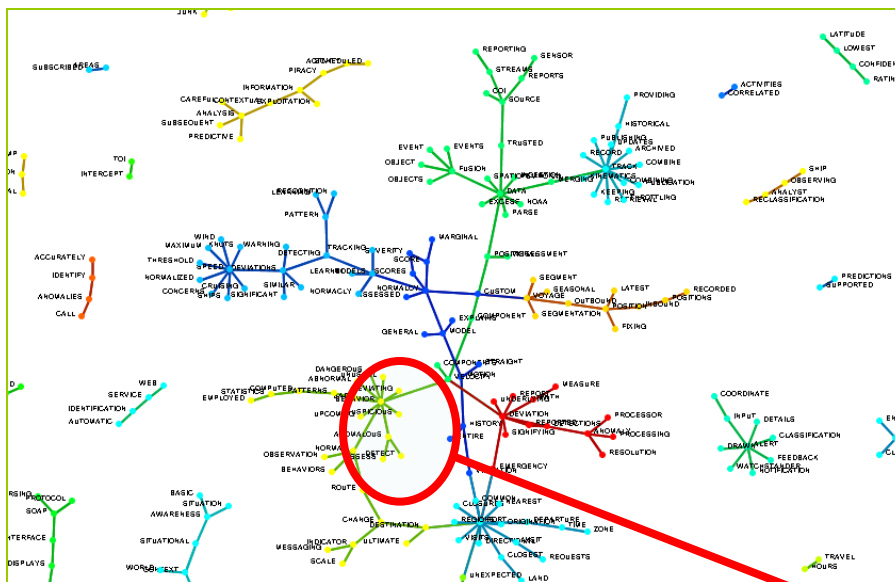
# Research Questions



- What were the roles and relationships of the organizations in the Humanitarian Assistance/Disaster Relief (*HA/DR*) operations?
- How were the operations actually conducted?
  - How can synergy, efficiency, and competency of an interagency operation be measured that involves many organizations with respect to military or civilian interests?
  - How is synergy between organizations developed over time?
  - How can we extract social networks such as people, locations, and organizations from social media or social network data?
  - How do social media, such as a discussion forums or social networking tools via Facebook and Twitter, help in an interagency operation?

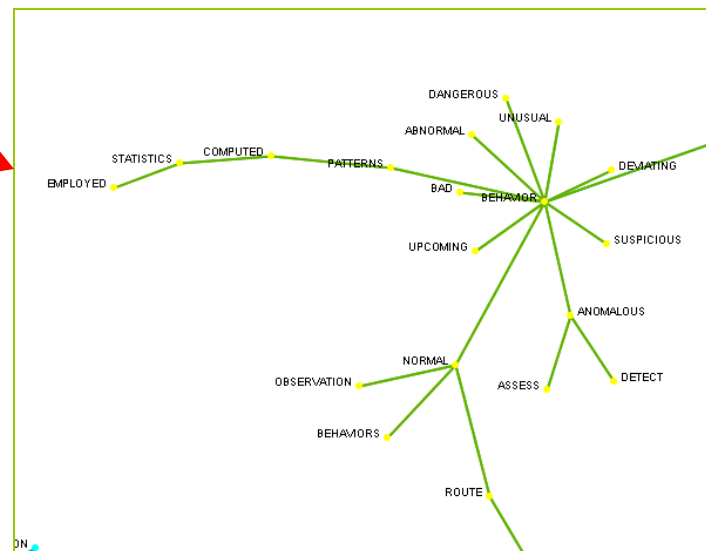


# Lexical Link Analysis



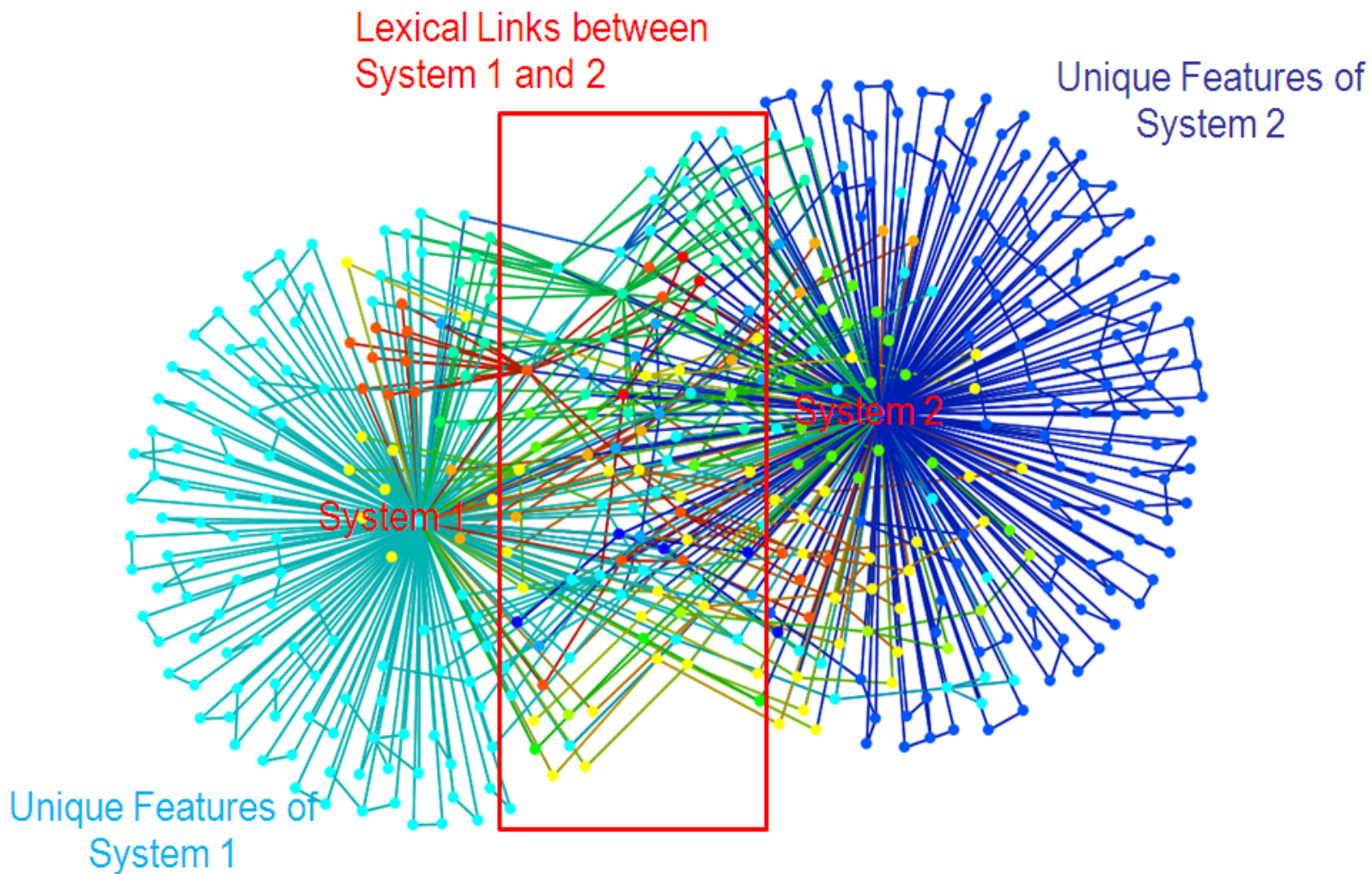
- Lexical analysis
- Link analysis
- Automatically discovers word pairs and groups them semantically

Can lexical links be used as metrics to measure the synergy, efficiency, competency of an interagency operation such HA/DR operations?

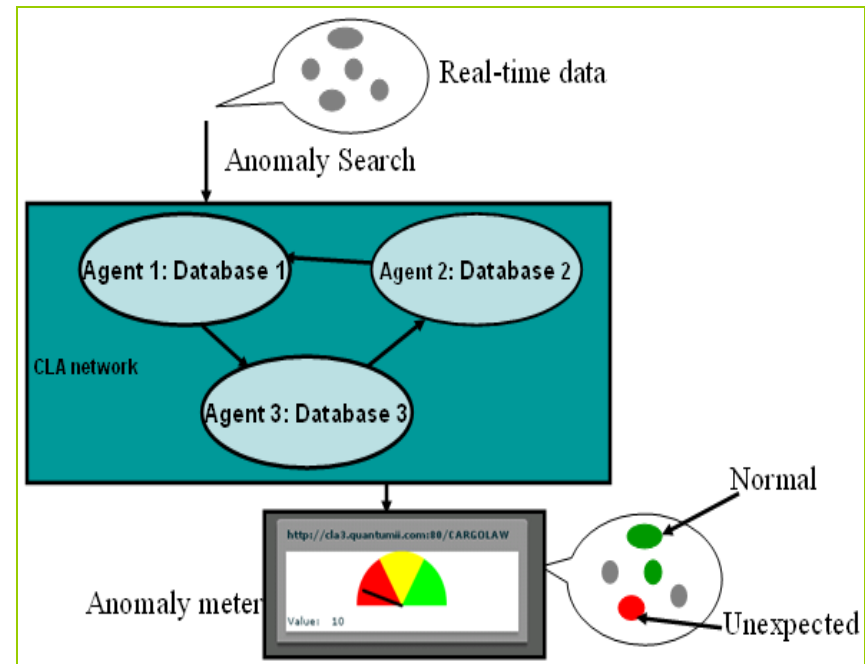
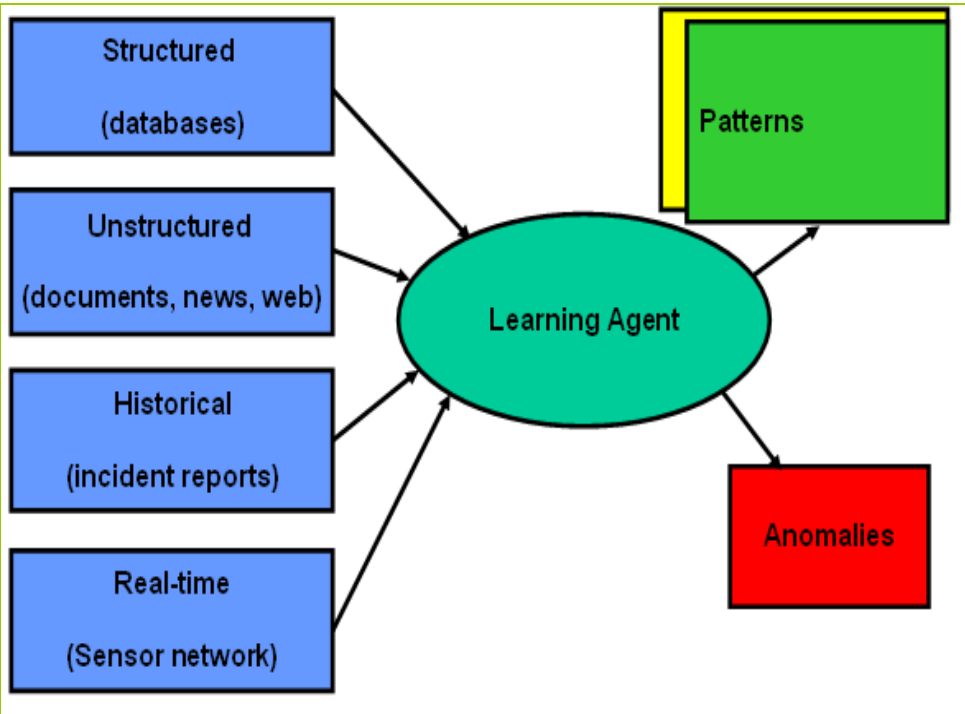




# LLA



# Agent Learning





# HAITI Project

- Data Sources
  - ~2600 open source web pages
    - Used High Performance Computer (HPC) nodes.
    - News feed from 1/13 – 2/23/2010
    - Twitter as a starting point
      - SOUTHCOM, USAID and others used Twitter to handle the situations
  - US SOUTHERN COMMAND also engaged the HAITI HA/DR Community of Interest (COI) on the All Partners Access Network (APAN) during the Haiti crisis. The APAN data was captured from a MSSQL database. The sources were:
    - Official documents and briefings in PDF: ~167 PDF file attachments related to HAITI HA/DR from 1/13/2010 to 5/26/2010
    - SITREP: ~150 Situation Report documents
    - Forum: ~1173 posts from 1/13/2010 to 6/3/2010
    - Blogs: ~3900 blog messages
- Our goal is to study organizational synergy using LLA



# Open Source Data Collection

- **Step 1:** Start with a list of web pages below, ask the crawler to go two levels deep to obtain all the links in the pages
  - <http://twitter.com/southcomwatch>
  - <http://www.southcom.mil/AppsSC/factFiles.php?id=138>
  - [http://twitter.com/USAID\\_Haiti](http://twitter.com/USAID_Haiti)
  - <http://www.inrelief.org/>
- **Step 2:** Go through all the pages collected from Step 1 and sort them according to timestamps, domains, and organizations.





# Lexical Link Metrics

- Number of nodes or word hubs/features
- Number of links between the word hubs
- Number of domains or organizations
- Number of cross-domains
  - E.g. Twitter/SOUTHCOM is a cross-domain when SOUTHCOM used Twitter to communicate
- Synergy index
  - Defined as the number of word hubs between two organizations divided by the total number of word hubs from the two organizations
    - Traditional quadratic assignment procedure correlation (QAP, Hubert & Schultz, 1976) used in social sciences and social network analysis.

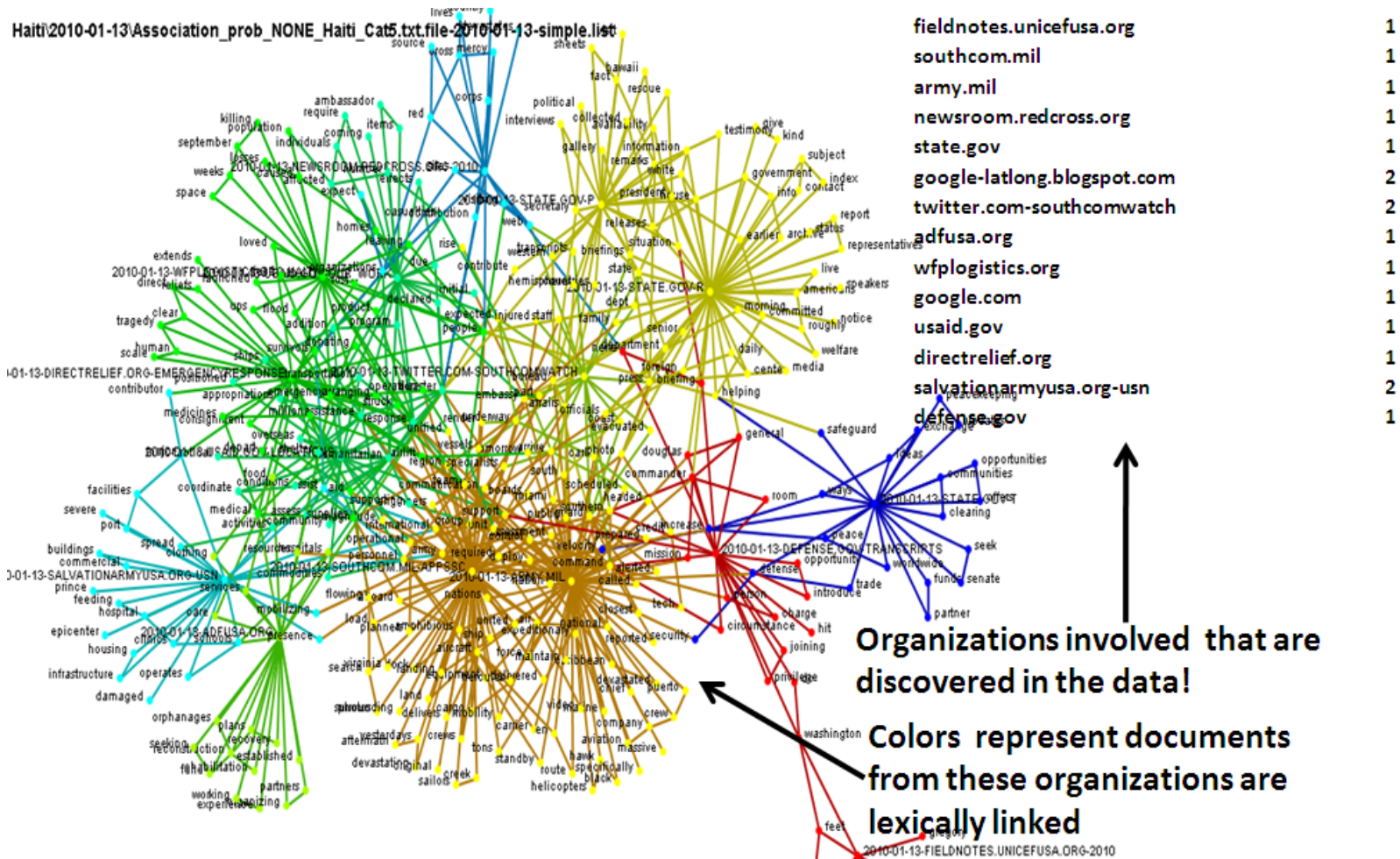




# Findings



## Low Synergy in the Beginning (2010-01-13)

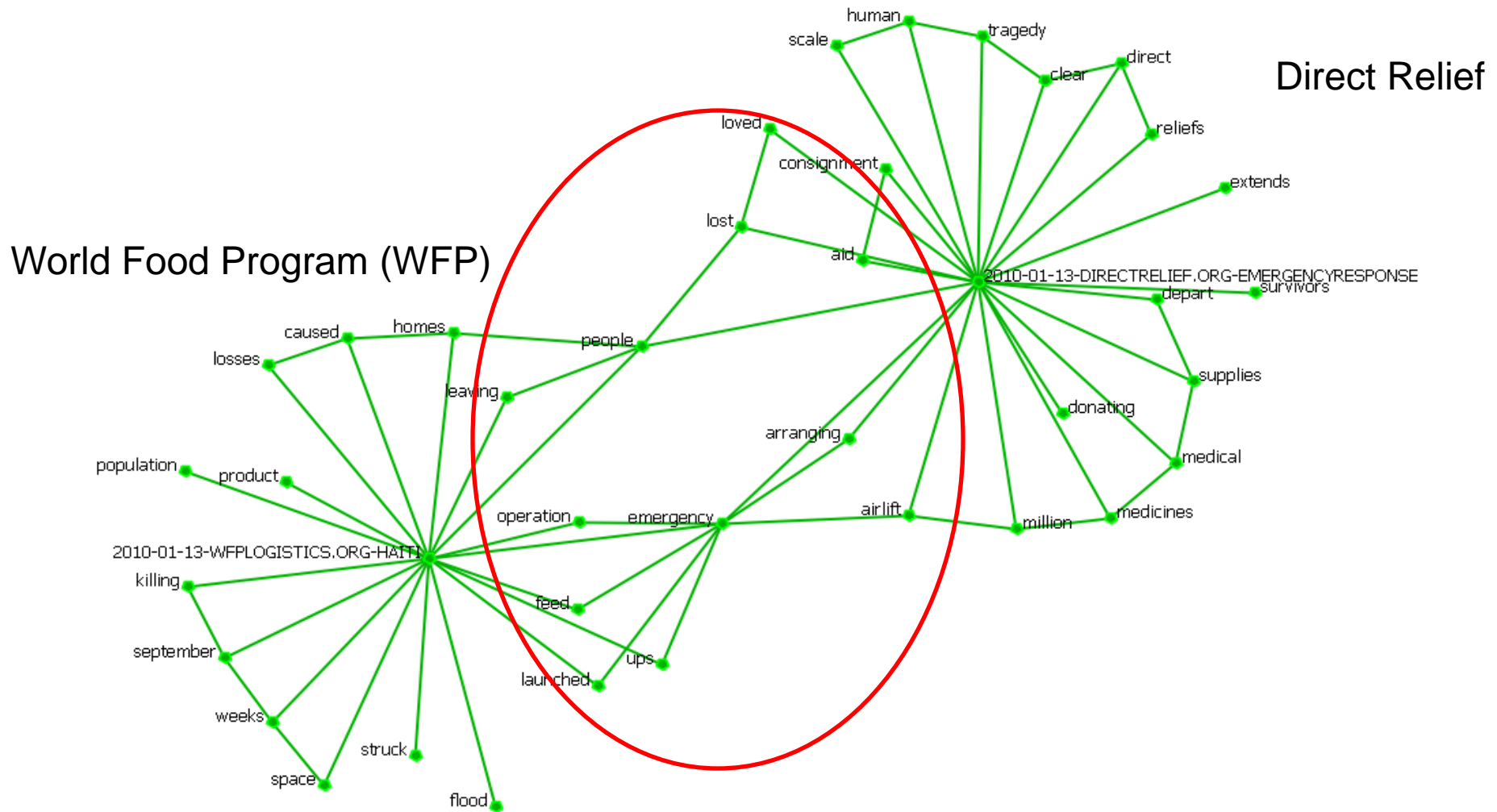






# Findings (detail)

## Synergy among NGOs (low??)

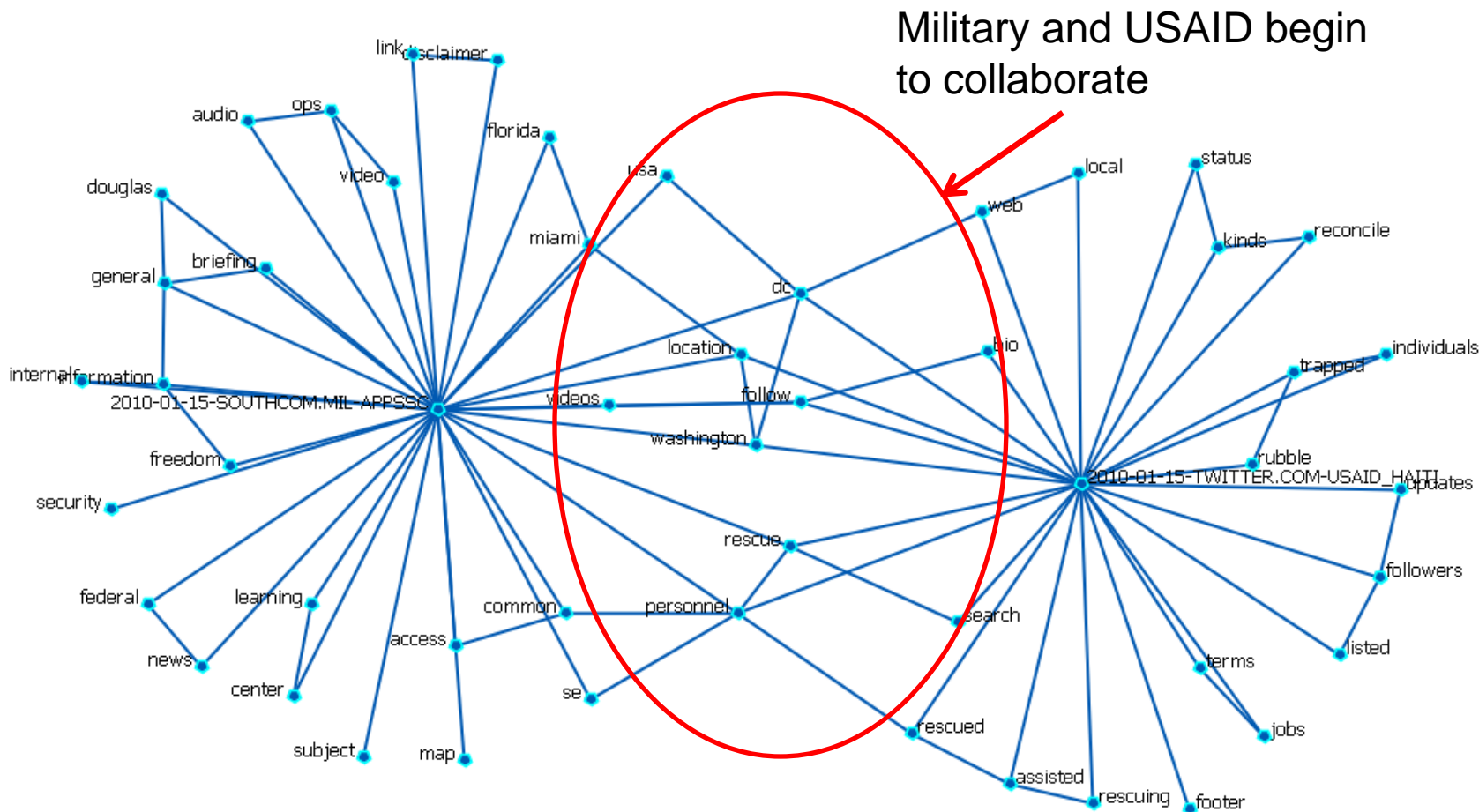




# Findings (detail)



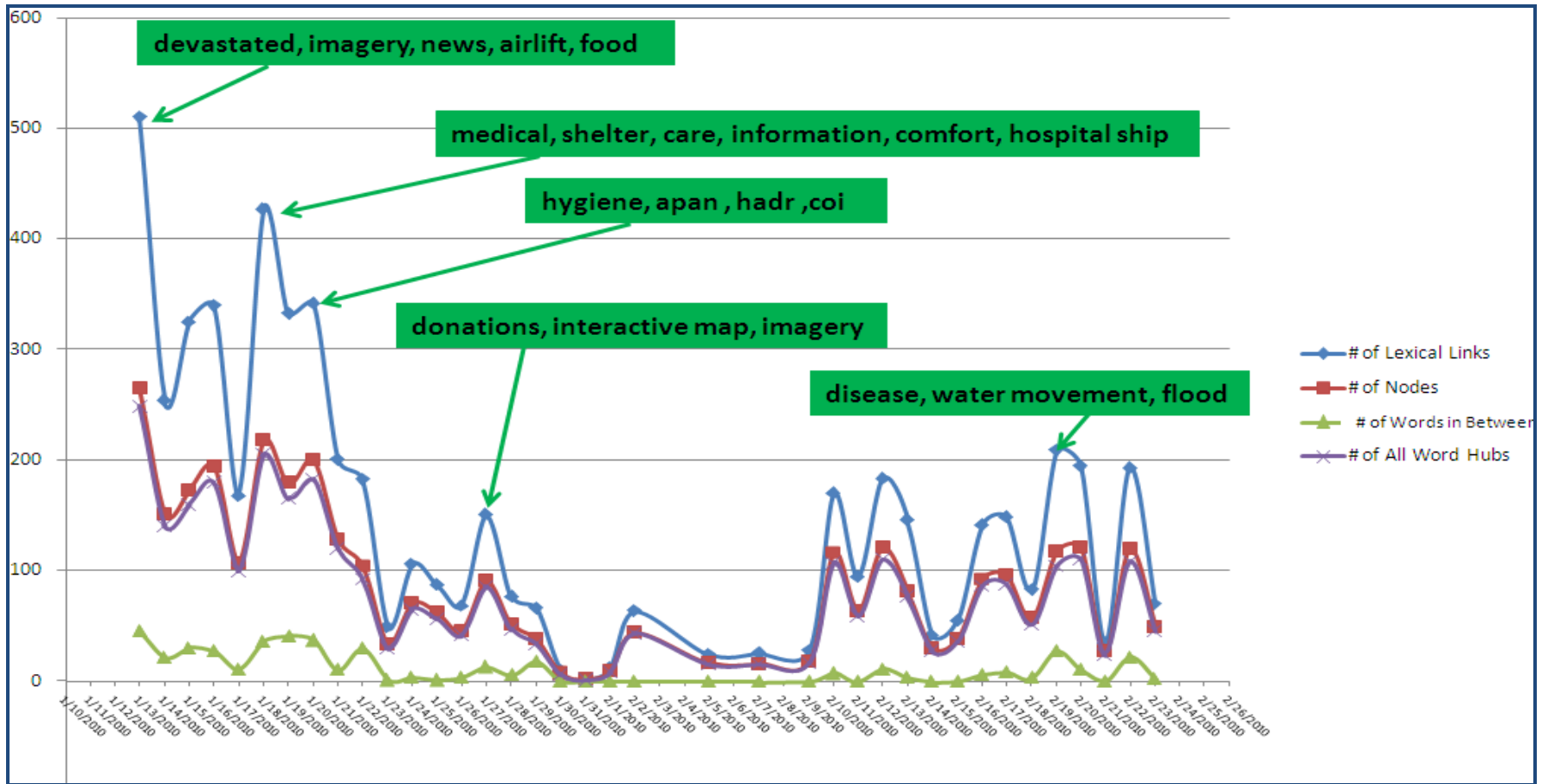
## SOUTHCOM and USAID (2010-01-15)





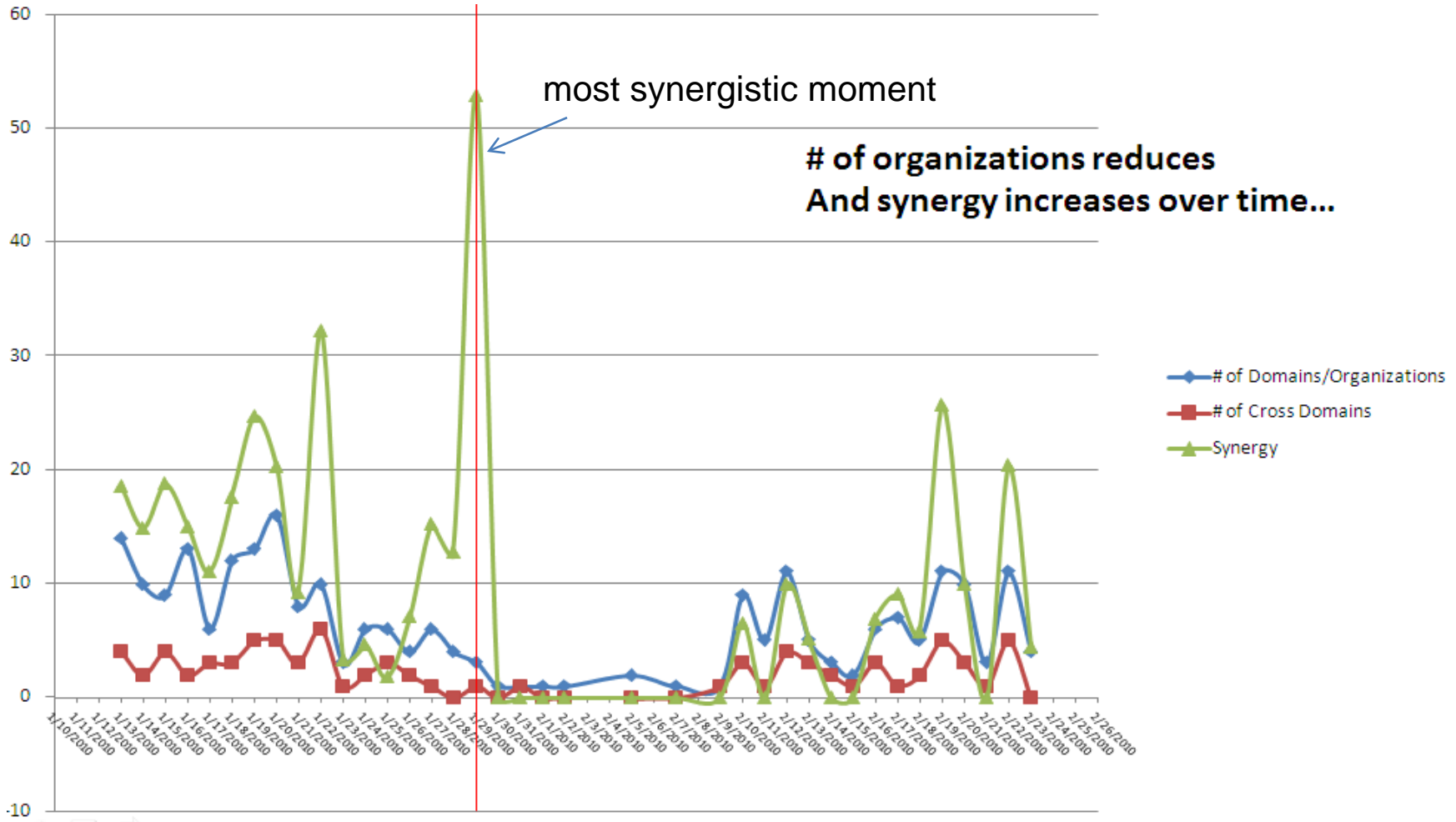


# Overall Measures and Trends





# Most Synergistic Point





# Timeline

## 6 Appendix

Table1: Organizations discovered in the open source data

	Military	Other Participating Organizations
1/13/2010		fieldnotes.umicefusa.org
	southcom.mil	
	army.mil	
		newsroom.redcross.org
	state.gov	
	twitter.com-southcomwatch	
		adfusa.org
		wfplogistics.org
		usaid.gov
		directrelief.org
		salvationarmyusa.org-usa
	defense.gov	
1/14/2010	ahp.us.army.mil	
		interaction.org
		haiti.usembassy.gov
		caritas.org
		actionaidusa.org
	dhs.gov	
1/15/2010		medicalteams.org-sf
	blogs.state.gov	

	facebook.com-chiefofnavaloperations	
		airserv.org
	jtfb.southcom.mil	
1/18/2010		
	airforcelive.dodlive.mil	
	travel.state.gov	
	coastguardallhands.blogspot.com	
1/19/2010		
	twitter.com-usnscmfort	
		heartlandalliance.org
		arcrelief.org
	navy.mil-gwf	
		inrelief.org
1/20/2010		
		cid.org
		dec.usaid.gov
		kars.ku.edu
	facebook.com-usairforce	
		facebook.com-redcross
		community.apan.org
		wfp.org
	af.mil	
1/22/2010		
		hope140.org
	whitehouse.gov	
	slideshare.net-jtfhaiti	

- SOUTHCOM.mil and Army.mil arrived on 1/13

- Inrelief.org and Community.apan.org arrived about a week after the disaster on 1/19 and 1/20.

- Cross-domain uses: twitter-southcomwatch and facebook.com-chiefofnavaloperations, began. SOUTHCOM and Chief of Naval Operations used Twitter and Facebook for the operation.





# Conclusions

---

- LLA is an effective tool to measure efficiency of interagency collaborations
- The number of overlapping lexical links can be a surrogate measure of synergy between two collaborating organizations.
- LLA examines interagency collaboration directly from the real-life communications and documents.
- Social media such as: Twitter, Facebook, and Google provide critical capabilities for military users such as SOUTHCOM and CNO.



**DISE**



# ***Backup slides***



# Lexical Link Analysis (LLA)



## Summary

- Lexical Analysis (LA wiki, 2009) is a form of text mining
  - Learns
  - Dynamically updates word and context associations with added data
- Link analysis
  - Like network analysis, explores and illustrates associations between objects
- Lexical Link Analysis (LLA)
  - Combines data mining with network analysis
  - Can dynamically identify, assess, and predict trends, patterns and features
- Data mining tools
  - Analyzes structured and unstructured data
  - Confirms previously known patterns, or discovers new patterns
  - Implements innovative visualization and navigation techniques
- Facilitates concept discovery, automated classification, and categorization of unstructured documents