

# SSnetViz: A Visualization Engine for Heterogeneous Semantic Social Networks

**Ee-Peng Lim**

**School of Information Systems**

**ICEC '09, August 12-15, 2009, Taipei, Taiwan**

# Acknowledgement

- Co-authors:
  - Maureen
  - Nelman Lubis Ibrahim
  - Aixin Sun
  - Anwitaman Datta
  - Kuiyu Chang
- Funding:
  - Agency for Science, Technology and Research (A\*Star)

# Outline

- Motivation
- Semantic Social Network
- Architecture
- Network Integration
- Network Search
- Network Export and Import
- Conclusion

# Motivation

- Analyzing people and their networks has become an increasingly important task
- Web has huge social network presence
  - Users participation in Web 2.0, e.g., Facebook, YouTube, Blogs, etc.
  - People mentioned in Web content, e.g., news sites, Wikipedia, etc.
- Databases have social network presence
  - E.g., customer data, terrorism data, etc.

# Social Network Analysis (SNA)

- Study of social networks and their elements
- Nodes can represent people, groups or organizations
- Tasks & Applications:
  - Association discovery
  - Expert finding
  - Recommendation
- Past SNA assumptions:
  - Only single type of nodes and links
- Today and Tomorrow's SNA:
  - **Multi-modality**: multiple types of nodes and links
  - **Distributed**: same person/group/organization in different networks

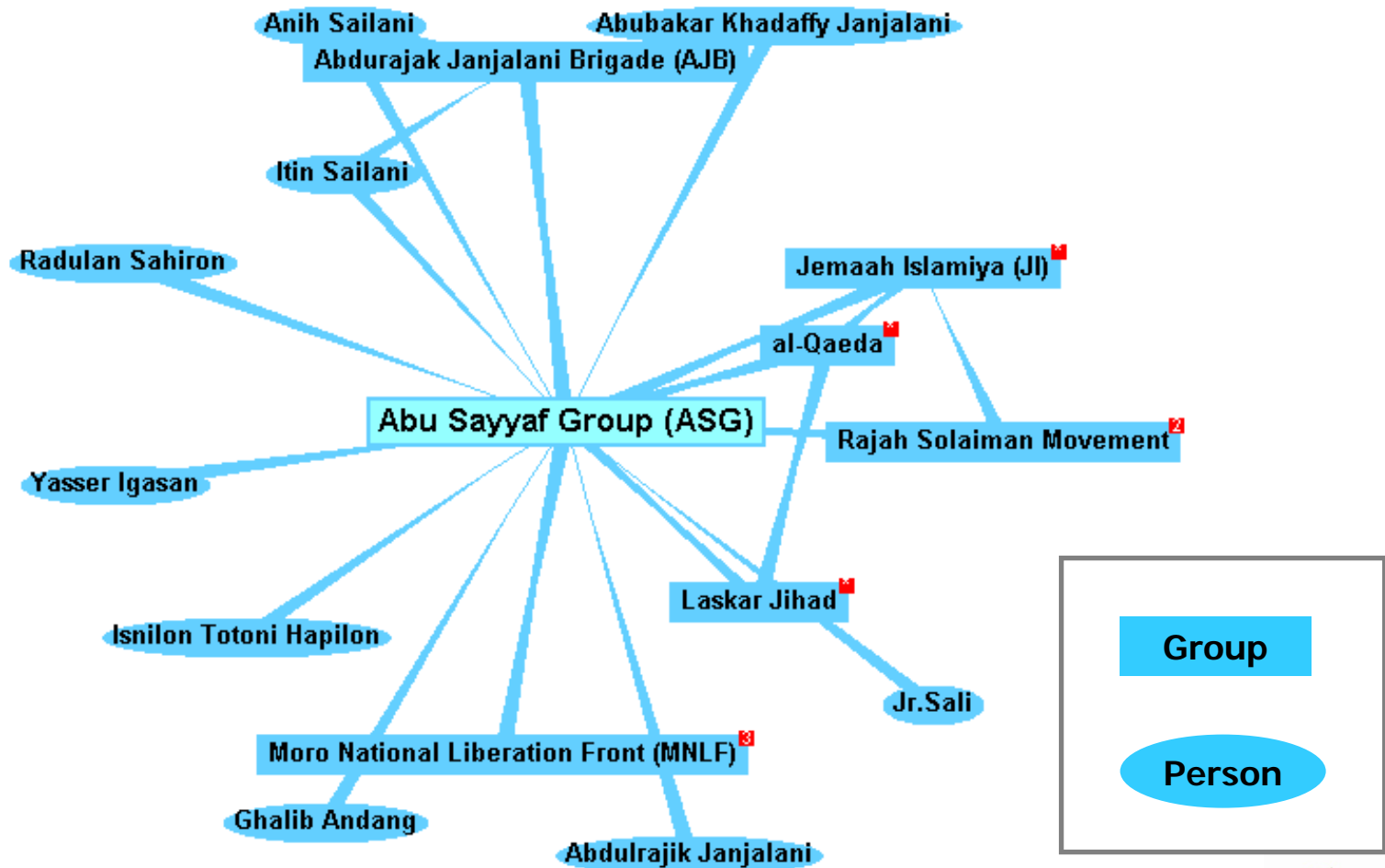
# Objectives

- Semantic social network representation
  - Multiple node types
  - Multiple link types
- Integration of multiple networks
- SSNetViz development
  - Storage
  - Integration
  - Visualization
  - Search

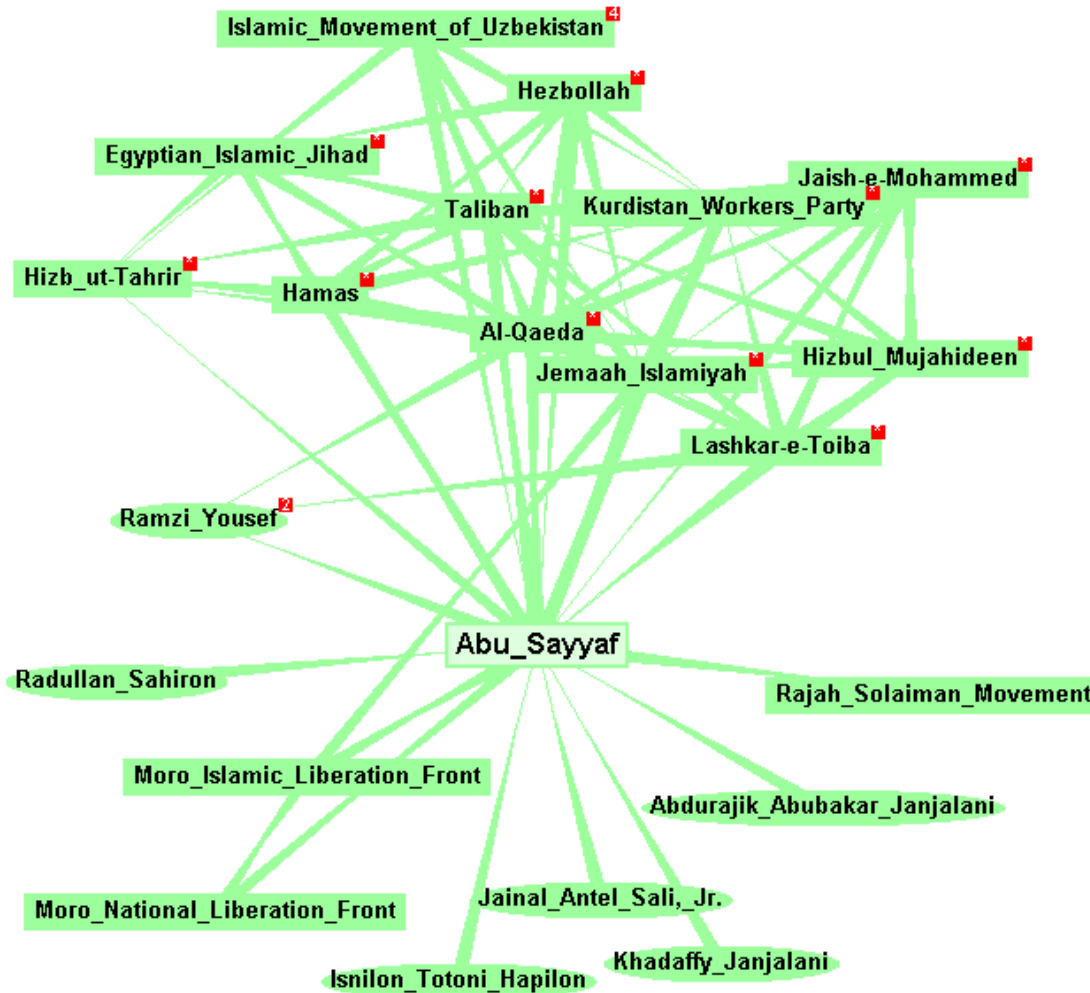
# Example: Terrorism Networks

- Domain expert collaboration:
  - Center for Political Violence and Terrorism Research (ICPVTR)
- MIPT's Terrorism Knowledge Base (TKB)
  - 858 terrorist groups + 1464 terrorists
- Wiki Terrorism Network
  - 471 terrorist groups + 542 terrorists
- Both networks
  - 277 terrorist groups + 112 terrorists

# Abu Sayyaf Group (ASG) in MIPT's Terrorism Knowledge Base (TKB)



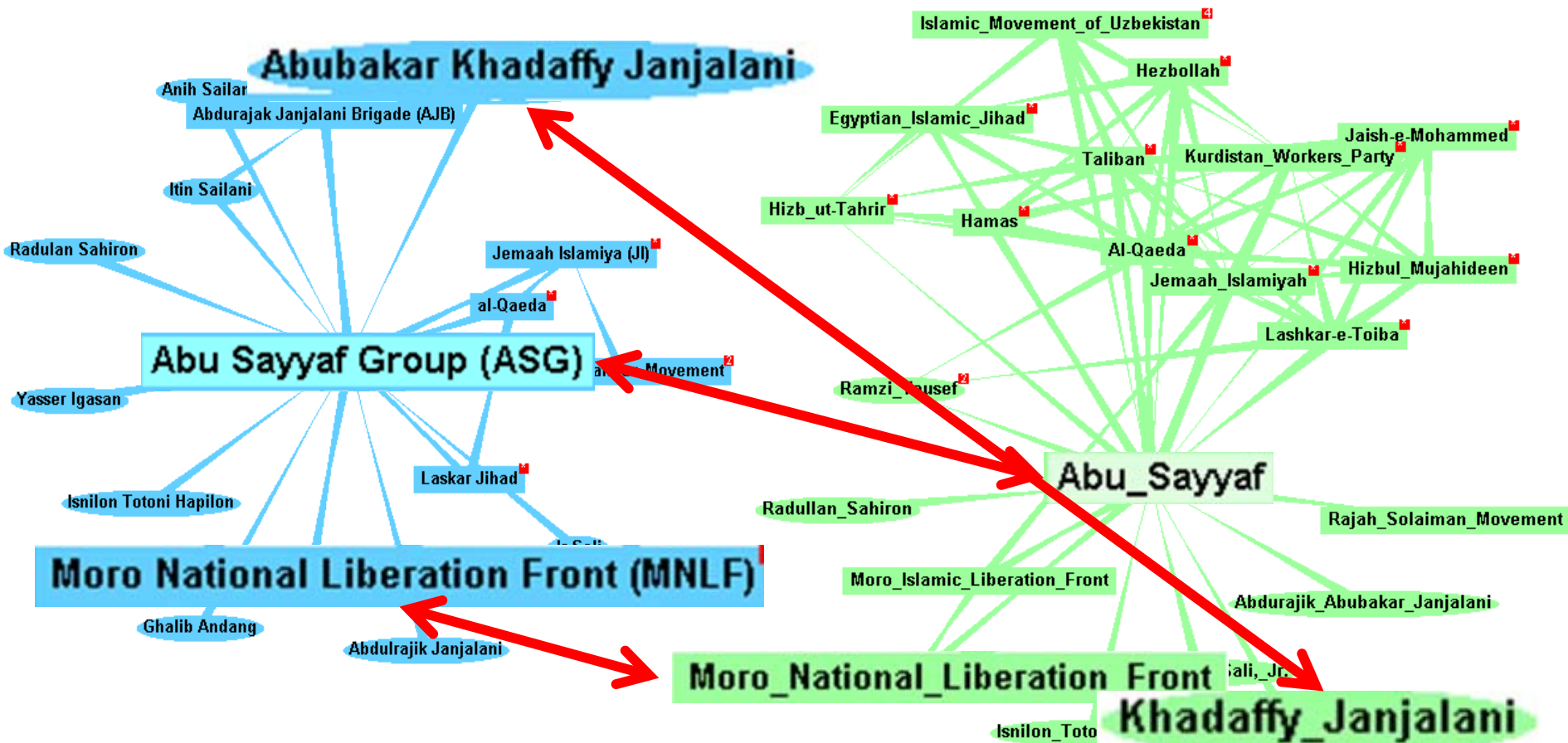
# Abu Sayyaf Group (ASG) in Wiki Terrorism Network



# Heterogeneous Attribute Sets

- TKB attribute sets
  - Group (Name, Aliases, Locations, Strength, Classification, URL)
  - Person (Name, Aliases, URL)
- Wiki Terrorism Network attribute set
  - Group(Name, Aliases, DateFormed, Description, Categories, URL, VersionURL)
  - Person (Name, Aliases, DOB, Description, Categories, URL, VersionURL)

# Matching Nodes



# Entity Resolution

- Find matched node pairs
- **Synonyms:** Same node but different names
- **Homonyms:** Different nodes but same names

Table 1: Matched Node Pairs

	TKB Network	WikiTerrorism Network
	<b>Terrorist Group</b>	<b>Terrorist Group</b>
1	Moro National Liberation Front (MNLF)	Moro_National_Liberation_Front
2	Abu Sayyaf Group (ASG)	Abu_Sayyaf
3	Jemaah Islamiya (JI)	Jemaah_Islamiyah
4	al-Qaeda	Al-Qaeda
5	Rajah Solaiman Movement	Rajah_Solaiman_Movement
	<b>Terrorist</b>	<b>Terrorist</b>
6	Isnilon Totoni Hapilon	Isnilon_Totoni_Hapilon
7	Radulan Sahiron	Radullan_Sahiron
8	Abubakar Khadaffy Janjalani	Khadaffy_Janjalani
9	Abdulrajik Janjalani	Abdurajik_Abubakar_Janjalani

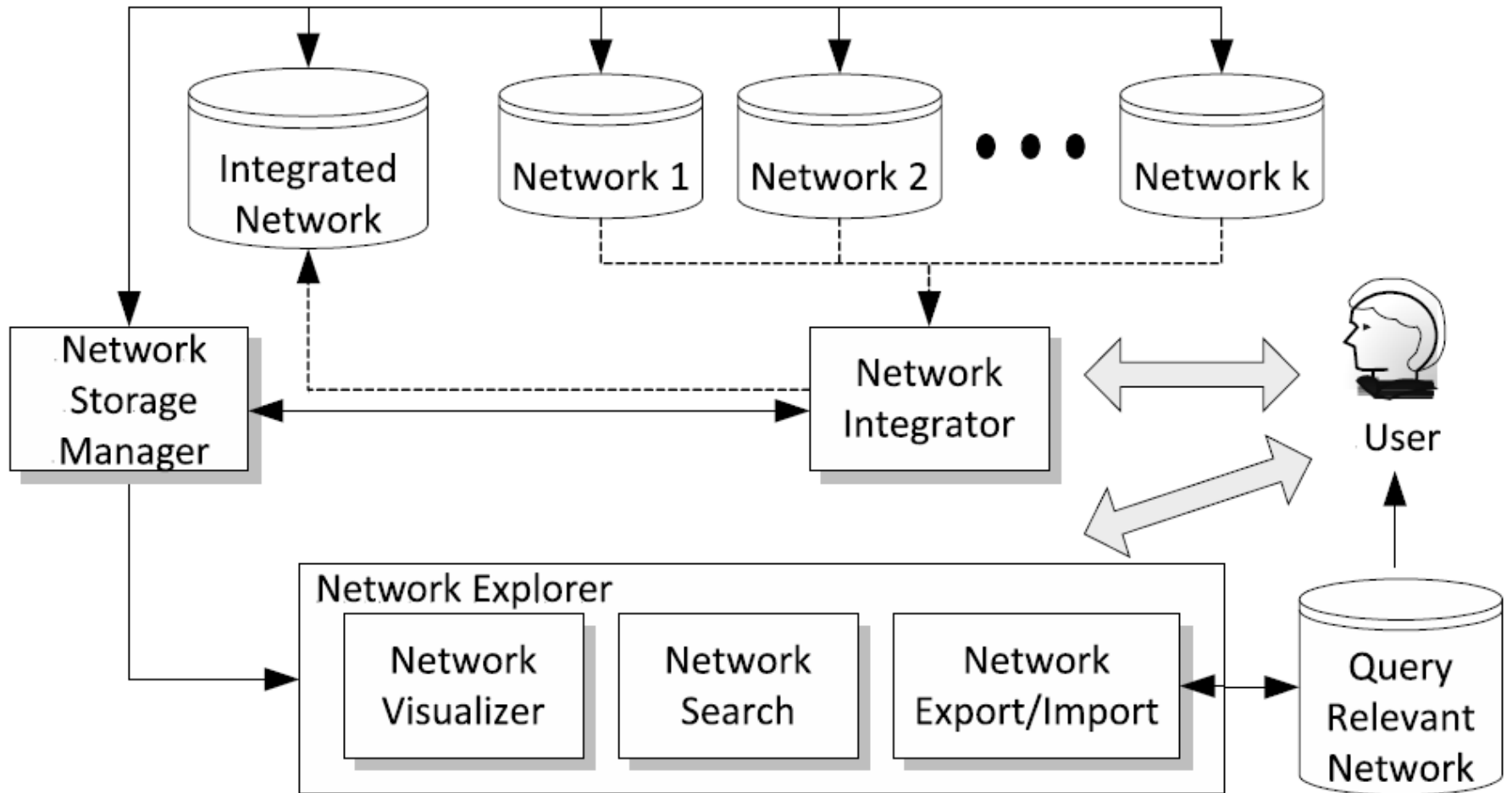
# Semantic Social Network Representation

- Each network can have different node and link types and node attributes.
- We model each network as:  
<SetofNodeTypes,SetofLinkTypes,G>
- **G** represents a graph instance with node and link instances
- Each node type **t** has a set of attributes **t.AttrSet**
- Each node instance **v** has **v.type** and **v.attr** attribute value
- Each link instance **(vi,vj)** has **(vi,vj).type**

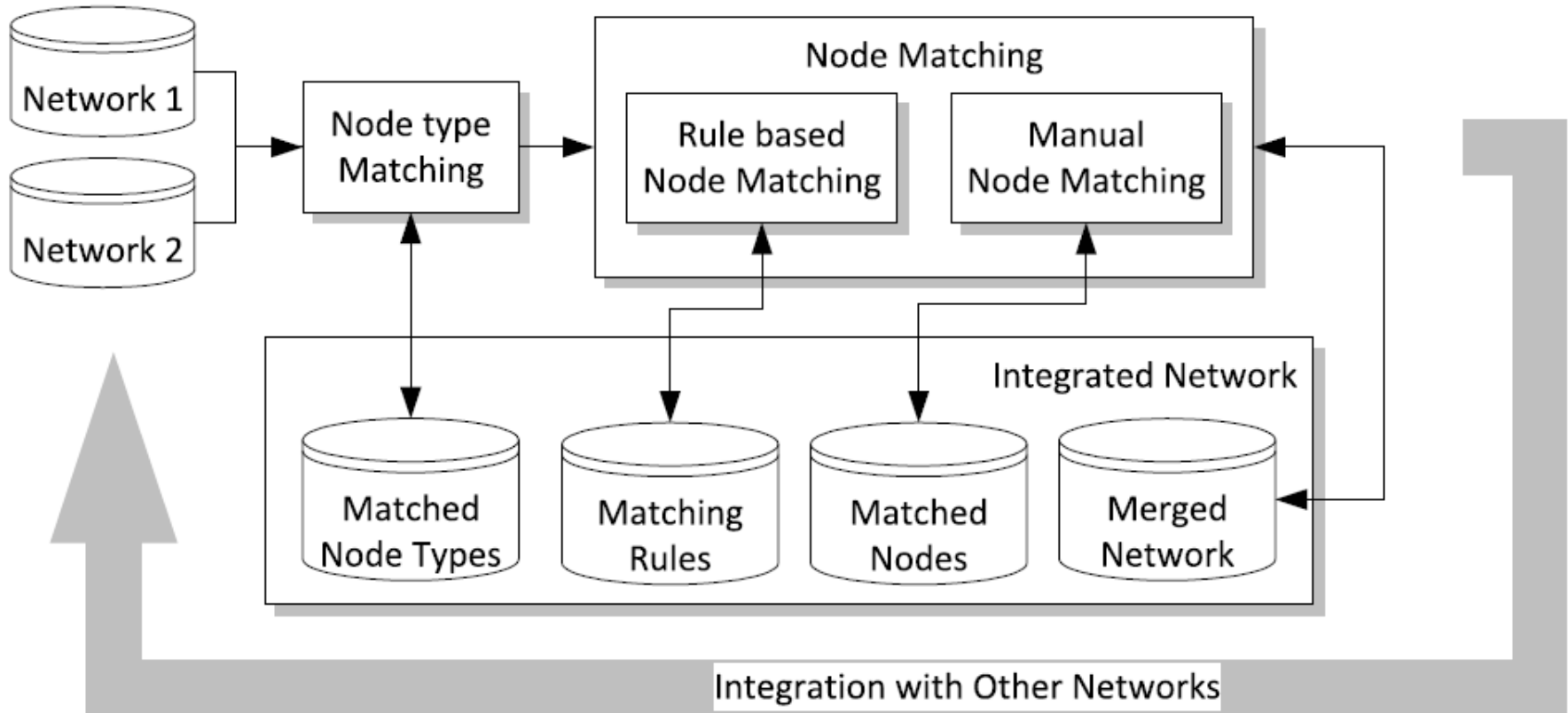
# SSNetViz Storage Scheme

- ATTRIBUTE table (for node types):  
ATTRIBUTE(sourceid, nodetypeid, attributename, attributeorder, ismultivalued, domainid)
  - stores the set of attribute names, their orders, and ids of their domains for each node type
- DOMAIN table:  
DOMAIN(domainid, domainname, datatype, minvalue, maxvalue)
  - stores the data types of attributes

# SSNetViz Architecture



# Integration Framework of SSnetViz



**Matching rule** = conjunction of conditions and each condition involves a *matching function and a pair of attributes, one from each node to be matched.*

Different matched node pairs may be derived by different matching rules.

# Matching Rule Specification

Set Rules and Match Entities for TKB and Wikipedia

**Rule Condition**

Condition

Use **Fuzzy String Match** for **Name** and **Name**

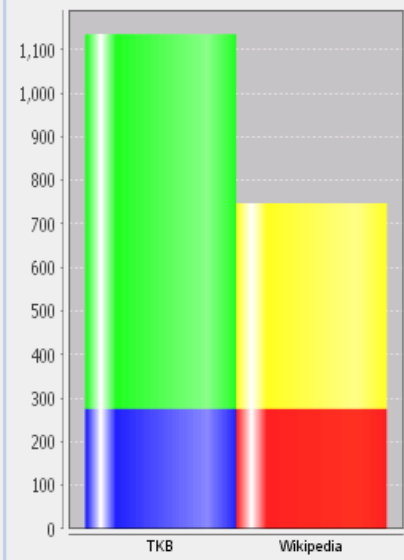
No.	Entity 1	Entity 2	Justify
1	Abu Nidal Organization	Abu Nidal Organization	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
2	Abu Sayyaf Group (ASG)	Abu Sayyaf Group (ASG)	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
3	Al Badr	Al Badr	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
4	Al Burq	Al Burq	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
5	Al Dawa	Al Dawa	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
6	Al Jihad	Al Jihad	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
7	Al Qaeda (AQ)	Al Qaeda (AQ)	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
8	Al-Aqsa Martyrs Brigades	Al-Aqsa Martyrs Brigades	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
9	Al-Barq	Al-Barq	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
10	Al-Ittihad al-Islami	Al-Ittihad al-Islami	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed
11	Al-Mansoorian	Al-Mansoorian	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unconfirmed

**Option for auto match**

Select All Yes  Select All No  Select All Unconfirmed

**Confirm judgements**

**Statistic of record pairs**



Source Name	TKB's Group Matched	TKB's Group Unmatched	Wikipedia's Group Matched	Wikipedia's Group Unmatched
TKB	270	830	0	0
Wikipedia	0	0	270	480

**Close**

# Manual Matching

**Set Rules and Match Entities for TKB and Wikipedia**

**Rule**

Name of Group from TKB (582)	Name of Group from Wikipedia (208)
1st Mechanical Kansas Militia	055_Brigade
21 June	Abu_Nidal_Organization
28 May Armenian Organization	Abu_Sayyaf_Group
313	Afghanistan_Liberation_Organization
3rd Continental Congress	African_Blood_Brotherhood
Abdurajak Janjalani Brigade (AJB)	Afrikaner_Weerstandsbeweging
Abu al-Abbas	Al-Aqsa_Martyrs_Brigades
Abu Bakr al-Siddiq Fundamentalist Brigades	Al-Assifa
Abu Mus'ab al-Zarqawi Battalion	Al-Dawa
Accolta Nazinuale Corsa	Al-Haramain_Foundation
Achik National Volunteer Council (ANVC)	Al-Ittihad_al-Islami
Actiefront Nationistisch Nederland	Al-Jihad
Action Committee of Winegrowers	Al-Ma'unah

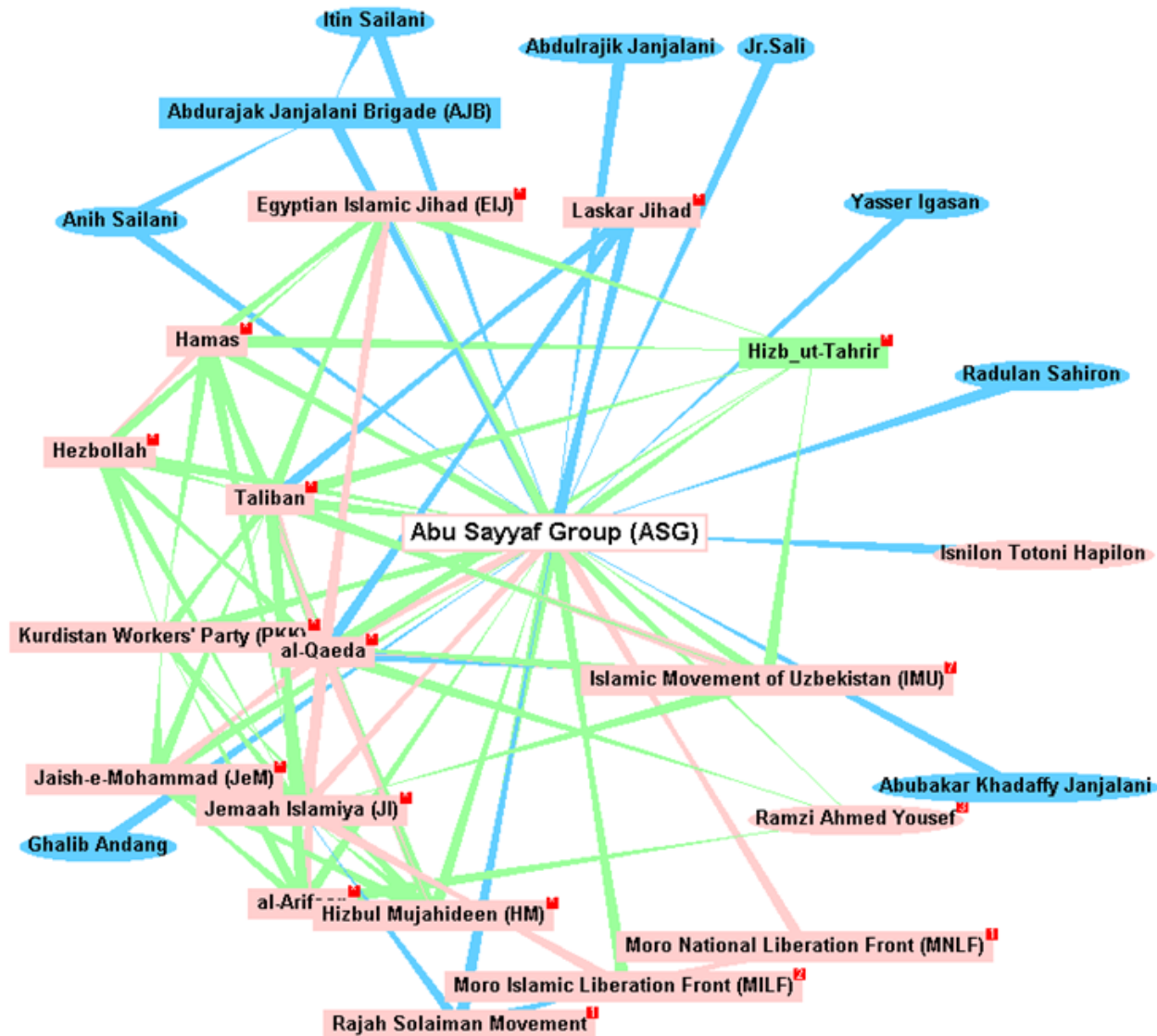
Name of Group from TKB	Name of Group from Wikipedia
Committee for Liquidation of Computers (CLODO)	CLODO
Communist Combatant Cells	Communist_Combatant_Cells
Communist Party of India-Maoist	Communist_Party_of_India_(Maoist)
Communist Party of Nepal-Maoist (CPN-M)	Communist_Party_of_Nepal_(Maoist)
Conscientious Arsonists (CA)	Conscientious_Arsonists
Continuity Irish Republican Army (CIRA)	Continuity_Irish_Republican_Army
Contras	Contras
Covenant Sword and Arm of the Lord (CSA)	The_Covenant_The_Sword_and_the_Arm_of_the_Lord
DHKP/C	Revolutionary_People's_Liberation_Party-Front
Dagestani Shari'ah Jamaat	Shariat_Jamaat
Democratic Front for the Liberation of Palestine (DFLP)	Democratic_Front_for_the_Liberation_of_Palestine
Dima Halam Daoga (DHD)	Dima_Halim_Daoga
Dukhtaran-e-Millat	Dukhtaran-e-Millat

**Statistic of record pairs**

Source Name	TKB's Group Matched	TKB's Group Unmatched	Wikipedia's Group Matched	Wikipedia's Group Unmatched
TKB	~280	~850	0	0
Wikipedia	0	0	~280	~470

■ TKB's Group Matched
 ■ TKB's Group Unmatched  
■ Wikipedia's Group Matched
 ■ Wikipedia's Group Unmatched

# Integrated Network



# Network Visualizer

SSnetViz - SSNet Visualization Tool

Graph Mode Search Help

Current Mode : Navigation Mode | Node Search:

**Your Anchor Nodes**

Abu Sayyaf Group (ASG) [Group]

Configure Anchor Node...

**Profile**

TKB Wikipedia

**Name**  
Abu Sayyaf Group (ASG)

**Aliases**  
-al-Harakat al-Islamiyah  
-Bearer of the Sword

**Locations**  
Philippines

**Strength**  
Greater than 200

**Legends**

Node	Source
<b>Grou</b>	TKB
<b>Persi</b>	TKB
<b>Grou</b>	Wikipedia

**Graph Controller**

Zoom  Rotate  Hyperbolic

**Node Controller ( Abu Sayyaf Group (ASG) )**

# Network Search

- Node search
  - Full text search on nodes
  - Rank nodes by relevance weighted by sources
- Path search
  - Returns paths from source to destination nodes
  - Rank paths by:
    - Path length
    - Rarity rank
    - Betweenness centrality
    - Closeness centrality

# Node Search

The screenshot shows a window titled "Node Search" with a search bar containing "Hamas" and a "Search" button. Below the search bar, it indicates "Search Result: 57 document(s)". There are navigation buttons labeled "1", "2", "3", and "Next".

The left sidebar shows a tree view of sources:

- All Sources (57)
  - icpvtr\_group (10)
  - icpvtr\_person (10)
  - tkb\_group (10)
  - tkb\_person (10)
  - wikipedia\_group (10)
    - Article (10)
    - Categories (2)
    - Description (4)
    - Name (2)
    - URL (2)
    - VersionURL (1)
  - wikipedia\_person (10)

The main content area displays search results for "hamas":

**hamas**  
EntityID:93 Name:**hamas** Strength:greater than 1,000 members URL:http://www.mipt.org/ Aliases...  
Locations:israel%content\_limiter%west bank/gaza Philosophy:**hamas** is a violent outgrowth of the..., **hamas'** founder and spiritual leader, sheikh ahmed yassin, was actively involved in non-violent muslim  
[tkb\\_group \(0.19\)](#) [wikipedia\\_group \(0.15\)](#) [icpvtr\\_group \(0.10\)](#) total score: 0.45

**ismail abu shanab**  
deputy to sheikh ahmed yassin, the spiritual leader of **hamas**. as yassin's deputy, he functioned as **hamas'** leader in gaza. for his involvement in **hamas**, he was imprisoned for 7 years. after his release in 1996, he served as **hamas'** observer to the plo central council (plo-cc) and served as **hamas**  
[tkb\\_person \(0.26\)](#)

**mahmud al-zahhar**  
gaza in april 1989. he then served as **hamas'** unofficial representative to the plo from january 1990  
[tkb\\_person \(0.11\)](#) [wikipedia\\_person \(0.13\)](#) total score: 0.23

**wasfi kabha**  
EntityID:3944 Name:wasfi\_kabha URL:http://en.wikipedia.org/wiki/wasfi\_kabha Categories:**hamas**... [[palestinian national authority]] and is a [[**hamas**]] member. he was detained by [[israel]]i authorities in june along with other **hamas** ministers and parliament members in a clampdown against their  
[wikipedia\\_person \(0.20\)](#)

$$combined\_score(r_i) = \sum_k w_k \times rel\_score(v_{ik})$$

# Path Search

Path Search

Source: [TKB, ICPVTR, Wikipedia]

Starting Node: tkb\_group, Jemaah Islamiya (JI)

Ending Node: tkb\_group, Abu Sayyaf Group (ASG)

Method: K\_SHORTEST\_PATH

Maximum Path Length: [ ]

Find Path Cancel

- **Path betweenness:** Sum of the betweenness of nodes in the path.  
Node betweenness: Number of shortest paths between any two nodes in the network that pass through the node.
- **Path rarity:** Average rarity of nodes along the path.  
Node rarity:  $1 - \frac{\# \text{ paths from source to target via the node}}{\# \text{ paths from source to target}}$

# Path Search

The screenshot shows a window titled "Path Search Result for Jemaah Islamiyah (JI), Abu Sayyaf Group (ASG)". The window contains a list of five paths, each with a checkbox and a list of groups. The paths are:

- Path 1: Jemaah Islamiyah (JI) (icpvtr\_group), Abu Sayyaf Group (ASG) (icpvtr\_group)
- Path 2: Jemaah Islamiyah (JI) (icpvtr\_group), Moro Islamic Liberation Front (MILF) (icpvtr\_group), Abu Sayyaf Group (ASG) (icpvtr\_group)
- Path 3: Jemaah Islamiyah (JI) (icpvtr\_group), Kurdistan Workers' Party (PKK) (tkb\_group), Abu Sayyaf Group (ASG) (icpvtr\_group)
- Path 4: Jemaah Islamiyah (JI) (icpvtr\_group), Al Qaeda (AQ) (icpvtr\_group), Abu Sayyaf Group (ASG) (icpvtr\_group)
- Path 5: Jemaah Islamiyah (JI) (icpvtr\_group), Taliban (icpvtr\_group), Abu Sayyaf Group (ASG) (icpvtr\_group)

At the bottom of the window, there are four buttons labeled 1, 2, 3, and 4, which likely correspond to the paths listed above.

# Conclusions

- SSNetViz addresses:
  - Visualization of heterogeneous social networks
  - Network integration
  - Search and exploration
  - Graph export and import
- Terrorism network datasets:
  - TKB Network
  - Wiki Terrorism Network
  - ICPVTR Network

# Future Work

- Network integration
  - Improved entity resolution
  - Attribute resolution
- Network analysis
  - Community finding
  - Trend spotting

# Thank you

**Ee-Peng Lim**

<http://aseplim.googlepages.com/socialnetworkmining>