

CNI 2009 Fall Meeting
Washington, D.C., 12/14

Crisis, Tragedy, and Recovery Network (CTRnet)

Edward A. Fox

Virginia Tech, CS, Digital Library Research Laboratory

fox@vt.edu <http://fox.cs.vt.edu> www.dlib.vt.edu

www.ctrnet.net www.dl-vt-416.org

<http://fox.cs.vt.edu/talks/2009/20091214CNI09CTR.pdf>

The 4/16 Digital Library



A digital library for recovery, research, resources, analysis and community relating to 4/16/2007, and after, at Virginia Tech



Share

Share your stories, photos and videos with us.



Resources

Browse our multimedia and external link collections.




Analysis

Query the resources, generate visualizations, data mine for patterns, and share results.

Thanks to: NSF IIS-0736055

Broader Interest

- ▶ School Shootings and other Tragedies (>260)
 - University of Texas clock tower shooting (1966)
 - Columbine (1999), Beslan school massacre (2004)
 - Northern Illinois U. & Kauhajokki Finland school (2008)
 - Bath, MI school bombing (1927, 45 dead, 58 injured)
 - Texas A&M Bonfire disaster (1999, Project BEAM)
- ▶ Unfolding CTR Events
 - Typhoon Ketsana, ...
- ▶ **Interest:** counseling, psychology, sociology, history, oral history, technology use, ...
 - ▶ **Idea:** distributed digital library network, with world-wide collaborating community



- Request new password

• Students

- [Venkat Srinivasan](#) - Graduate Research Assistant
- [Bidisha Dewanjee](#) - Student Collaborator
- [Amine Chigani](#) - Department of Computer Science

• Local Advisory Board Members

- Jerzy Nowak - Department of Horticulture
- Erv Blythe - Information Technology
- James Hawdon - Department of Sociology
- Russell Jones - Department of Psychology
- Timothy Luke - Political Science
- Gail McMillan - Digital Library and Archives
- Christopher North - Department of Computer Science
- John Ryan - Department of Sociology

• External Advisory Board Members

- Starr Roxanne Hiltz - NJ Institute of Technology
- Murray Turoff - NJ Institute of Technology
- Kristine Hannah - Internet Archive
- Susan Metros - U of Southern California
- Ben Shneiderman - U of Maryland
- Padmini Srinivasan - U of Iowa
- Eric Van de Velde - Caltech



LATEST

- First CTRnet Local Advisory Board Meeting
- CTRnet Team Wins Award at OutreachNOW Conference
- CTRnet Team Welcomes it's New Collaborators
- CS5604 students join the project team



WHO'S

Acknowledgments

Co-PIs: Andrea Kavanaugh, Naren
Ramakrishnan, Steven Sheetz,
Donald Shoemaker

GRAs: Bidisha Dewanjee,
Venkataraghavan Srinivasan

CS5604 Team: Tristram Bethea, Yipan
Deng, Seth Fox, Min Li, Chao Peng



Departments:

Accounting Information Systems
Computer Science
Sociology



Grant:
IIS-0916733

Intro: CTR

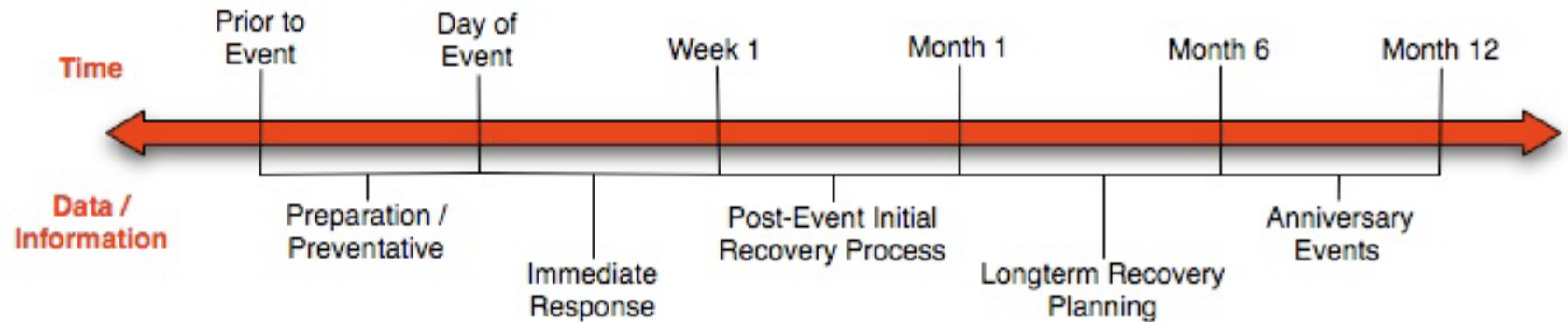
▶ **Difficulties faced:**

- information not collected or else scattered;
- stove-piped communities → survey contamination;
- few facing long-term issues/preservation;
- wide varieties of culture, content/genre, ...

▶ **Scope:** Human tragedies (loss of life) that results from

- natural or manmade events, where
- communities are effected (especially, larger communities)

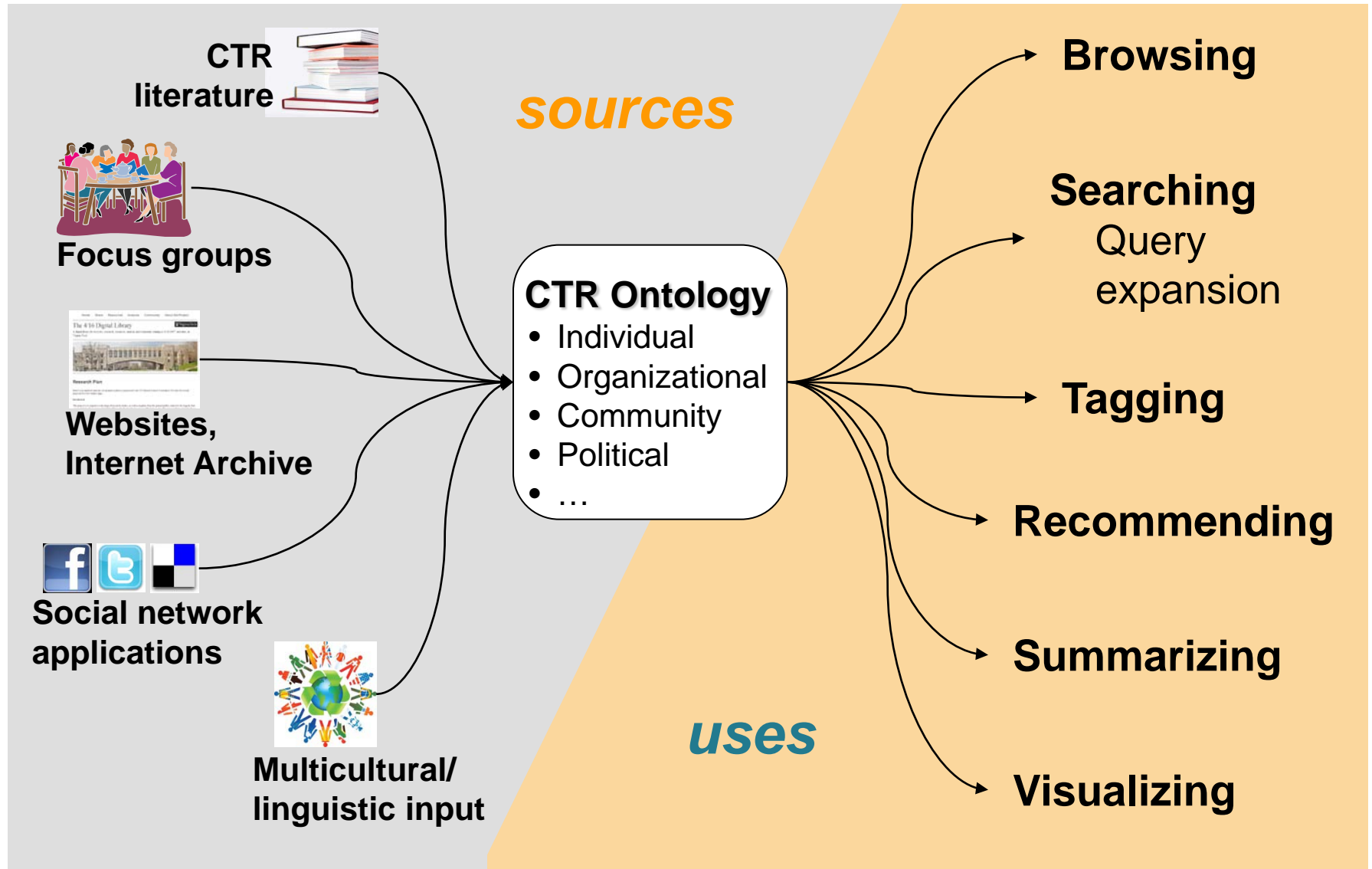
Intro: CTR Timeline



Broader Goals: Technical:

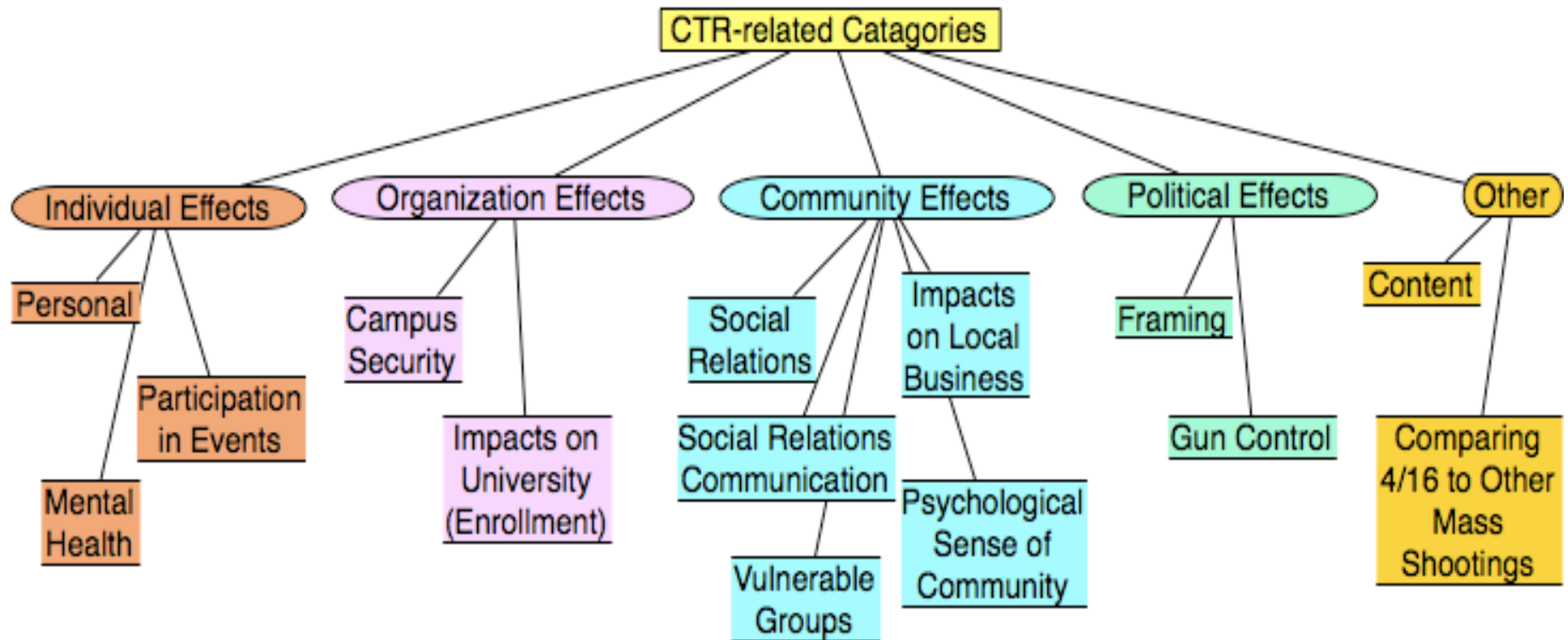
- ▶ Integrating heterogeneous information in a specific domain, making it accessible, and preserving it for long-term reuse
- ▶ Extend the scope of intelligent digital libraries with ontology
- ▶ Support information exploration with Storytelling & SSP

Goals for Ontology for CTR

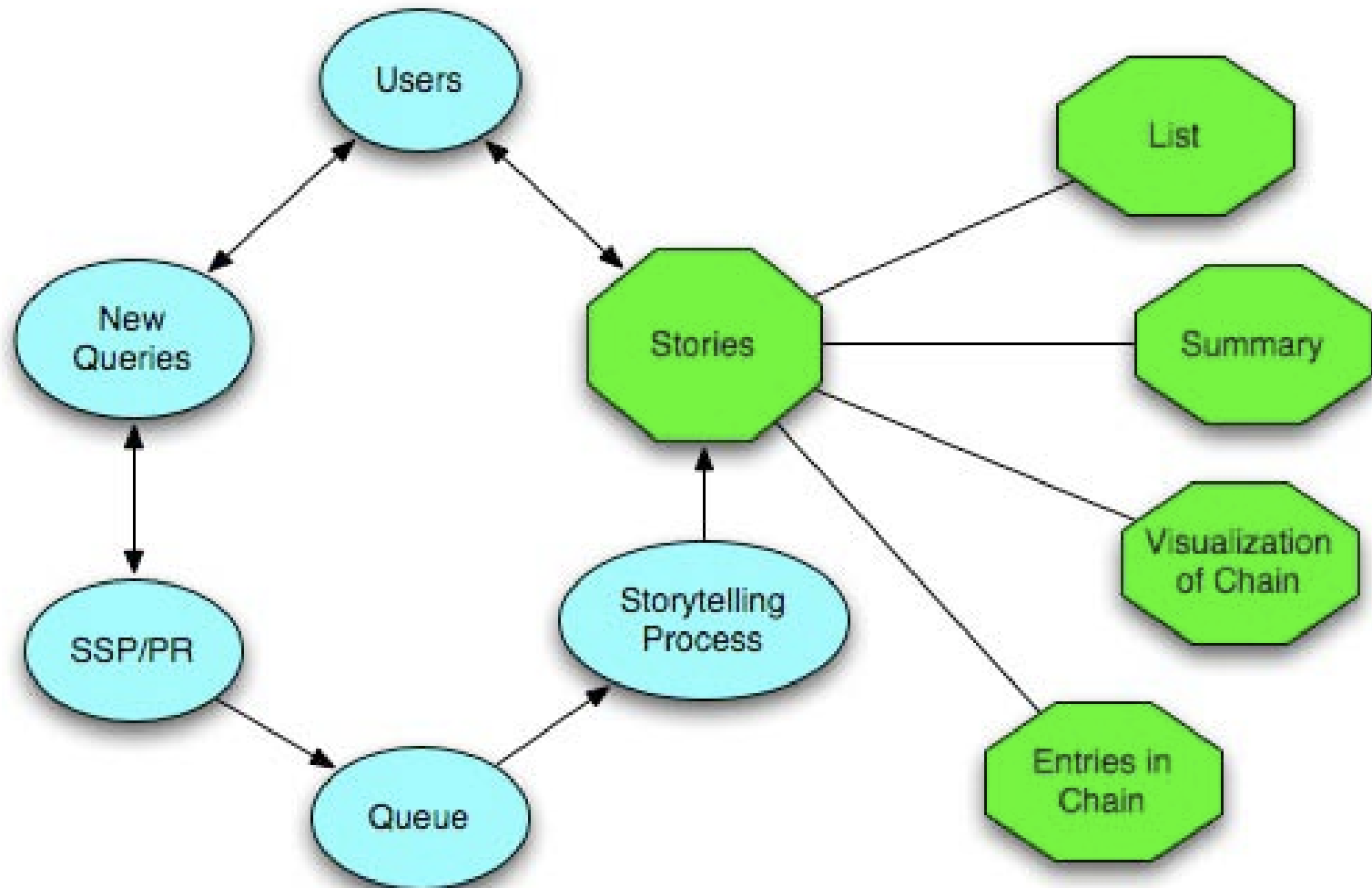


Categories from focus group study

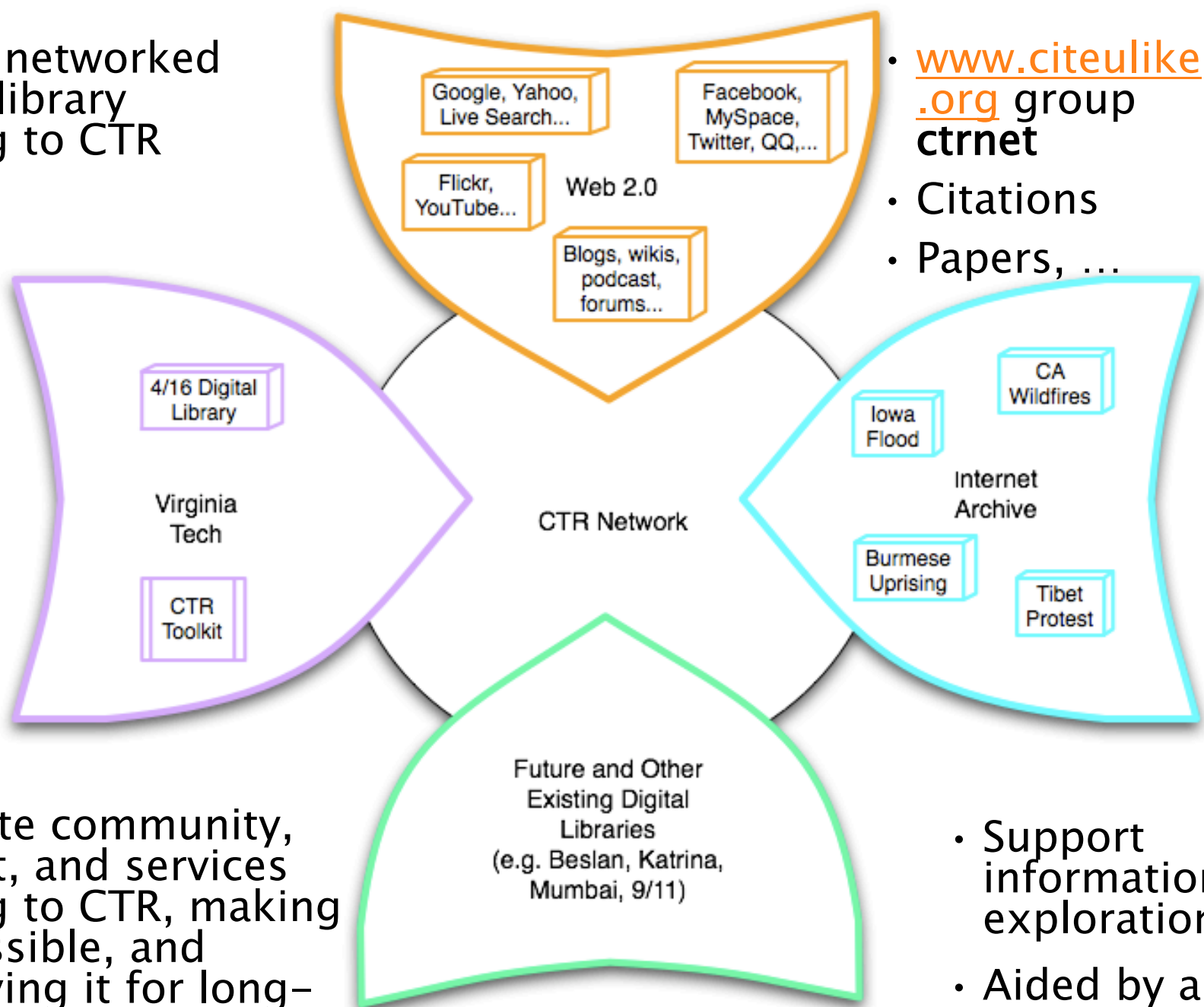
Results from focus group interviews following the April 16, 2007 tragedy at Virginia Tech



SSP¹ and Storytelling



- ▶ Build a networked digital library relating to CTR

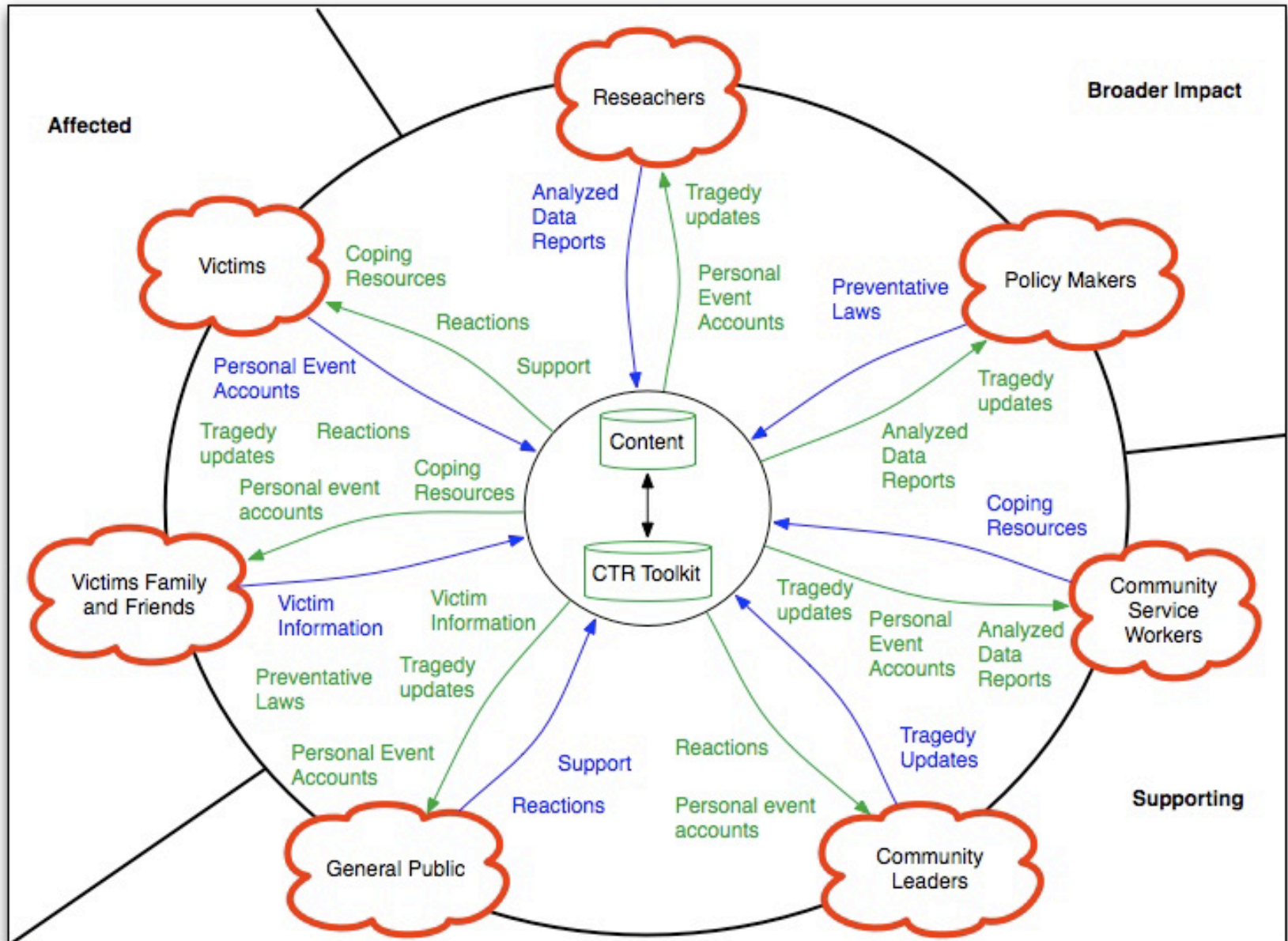


- www.citeulike.org group ctrnet
- Citations
- Papers, ...

- Support information exploration
- Aided by an ontology

- Integrate community, content, and services relating to CTR, making it accessible, and preserving it for long-term reuse

CTR stakeholders



CS5604 Team Project Objectives



Build data collections on national tragedies and disasters by crawling the web for data.



Filter and classify the crawled web data to eliminate junk within the collection.



Perform Storytelling upon the filtered data to unveil relationships among the tragedies.



Index the data collections to allow for searching and document retrieval.

Automatic Generation of Seeds

- ▶ “Pathfinder” is a new automated utility designed to find seeds relating to a certain topic.
 - Uses Google to find the top 5 related queries for a topic.
 - Retrieves information from Google’s categorical search engines.
 - Google Blogs, Google Images, etc...
 - Locates approximately 500 seeds per category and is executed on the original and related queries.



Manual vs. Generated Seeds

Manual Seeds

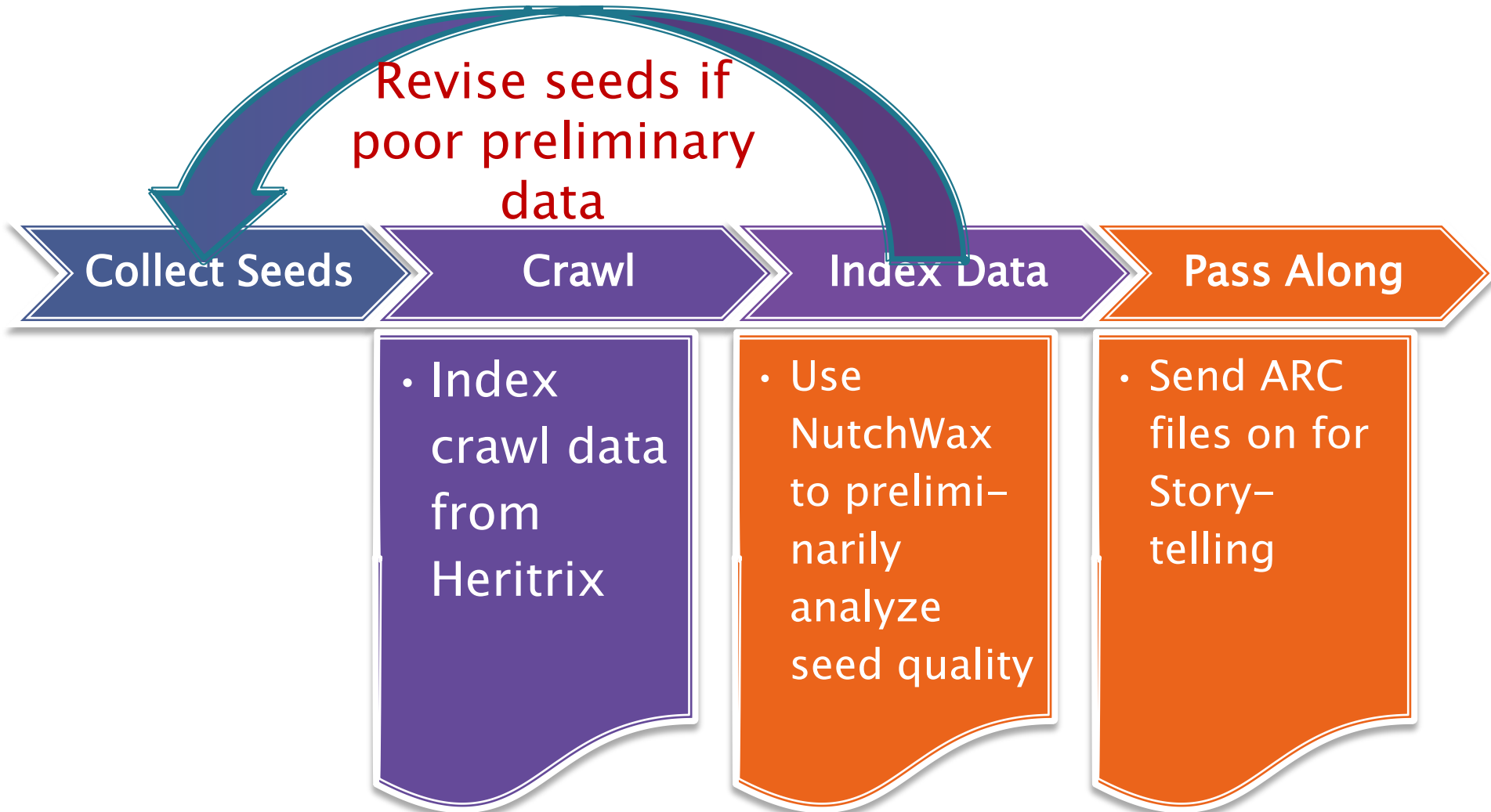
- **Accurate**
- Time Intensive to Produce
- Relatively Small in Quantity
- **Results Require Moderate Filtering**

Generated Seeds

- Cover Range of Material
- **Rapid** Production
- Large **Quantity**
- Results Require Heavy Filtering

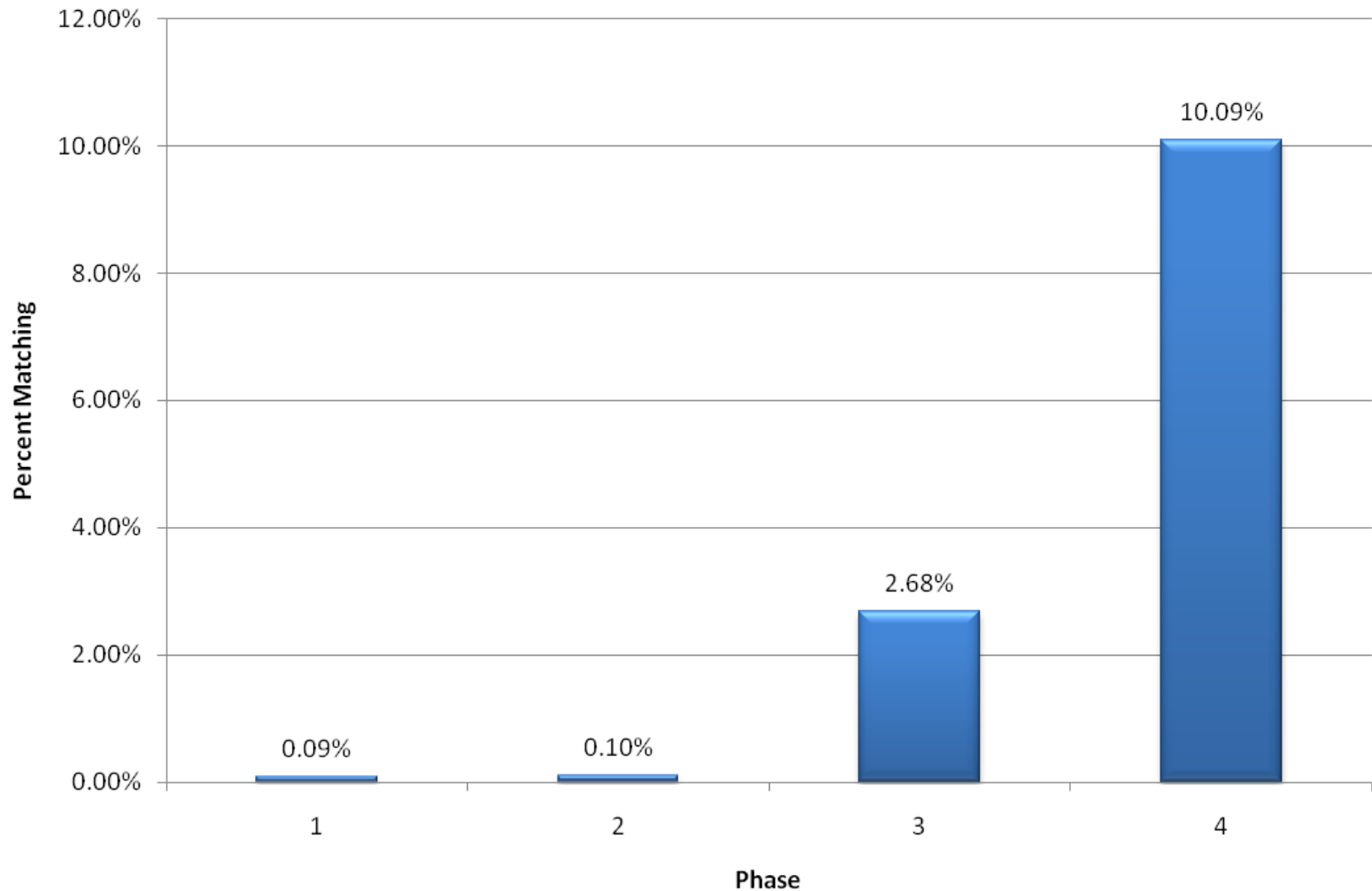


Preliminary Data Analysis



Results

Percent Matching by Phase of Development



Data Filtering and Storytelling

Crawling

Preprocessing

Classification

Storytelling

- Extracting Text
- Basic Text Cleanup

- Supervised learning methods
- Evaluation
- Classifying new data

- Generating stories
- Visualization
- Story analysis

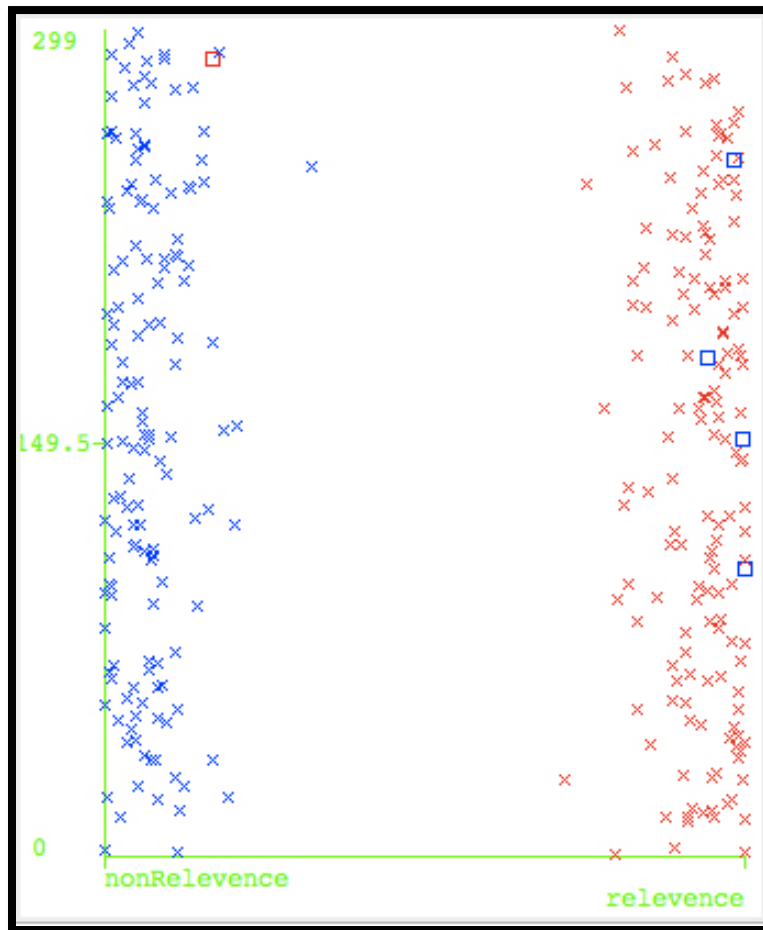
Experimental Results

	Precision				Recall				Accuracy	
	NB		SVM		NB		SVM		NB	SVM
	R	N	R	N	R	N	R	N		
VT-416	0.993	0.974	1.000	1.000	0.973	0.993	1.000	1.000	98.3%	100.0%
	0.979		1.000		0.983		1.000			
Amish	0.881	0.992	0.985	0.961	0.993	0.873	0.957	0.987	93.1%	97.3%
	0.937		0.973		0.933		0.972			
Columbine	0.810	1.000	0.961	0.980	1.000	0.767	0.980	0.960	88.3%	97.0%
	0.905		0.971		0.884		0.970			

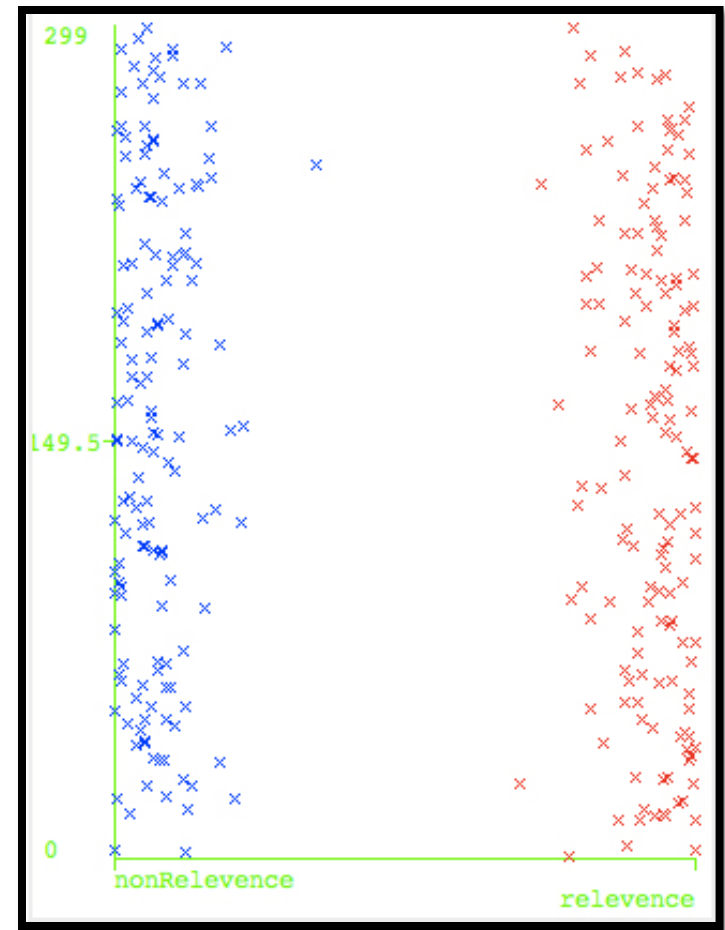
R : relevant class N : non-relevant class

“Accuracy” is computed from Confusion Matrix.

Experimental Results (contd.)

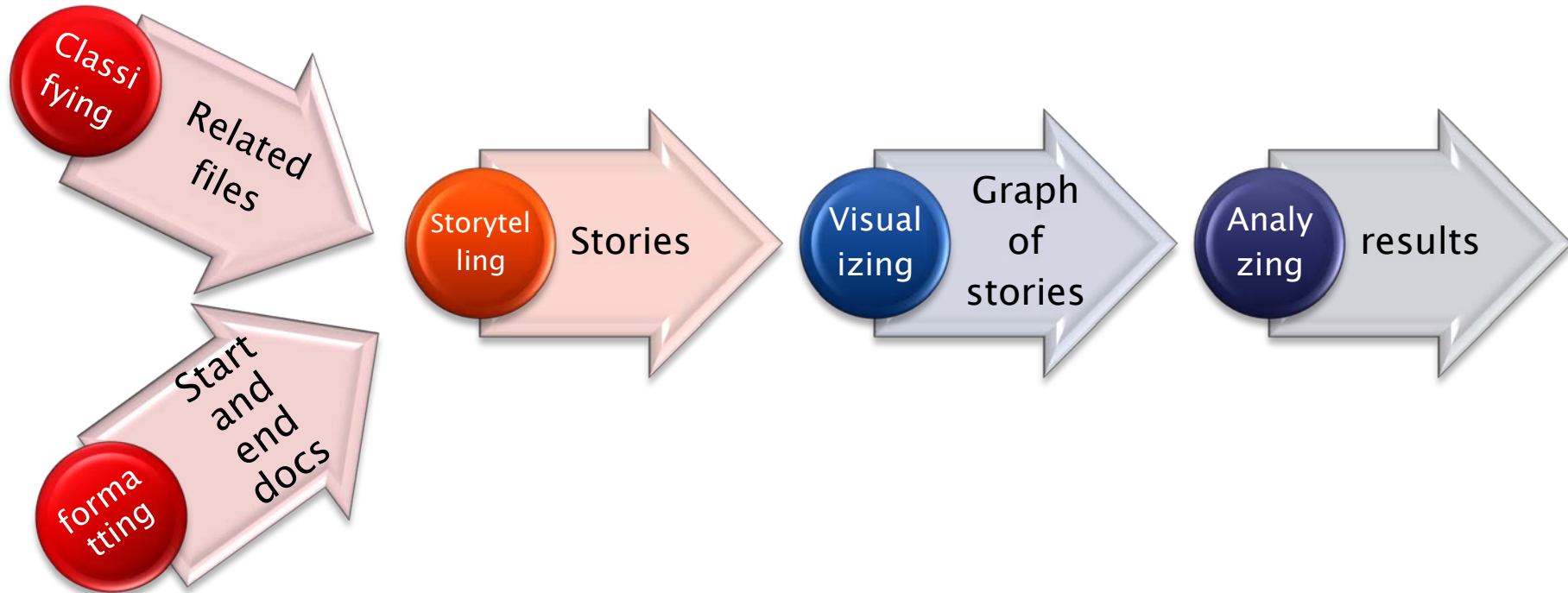


VT-416 with NB



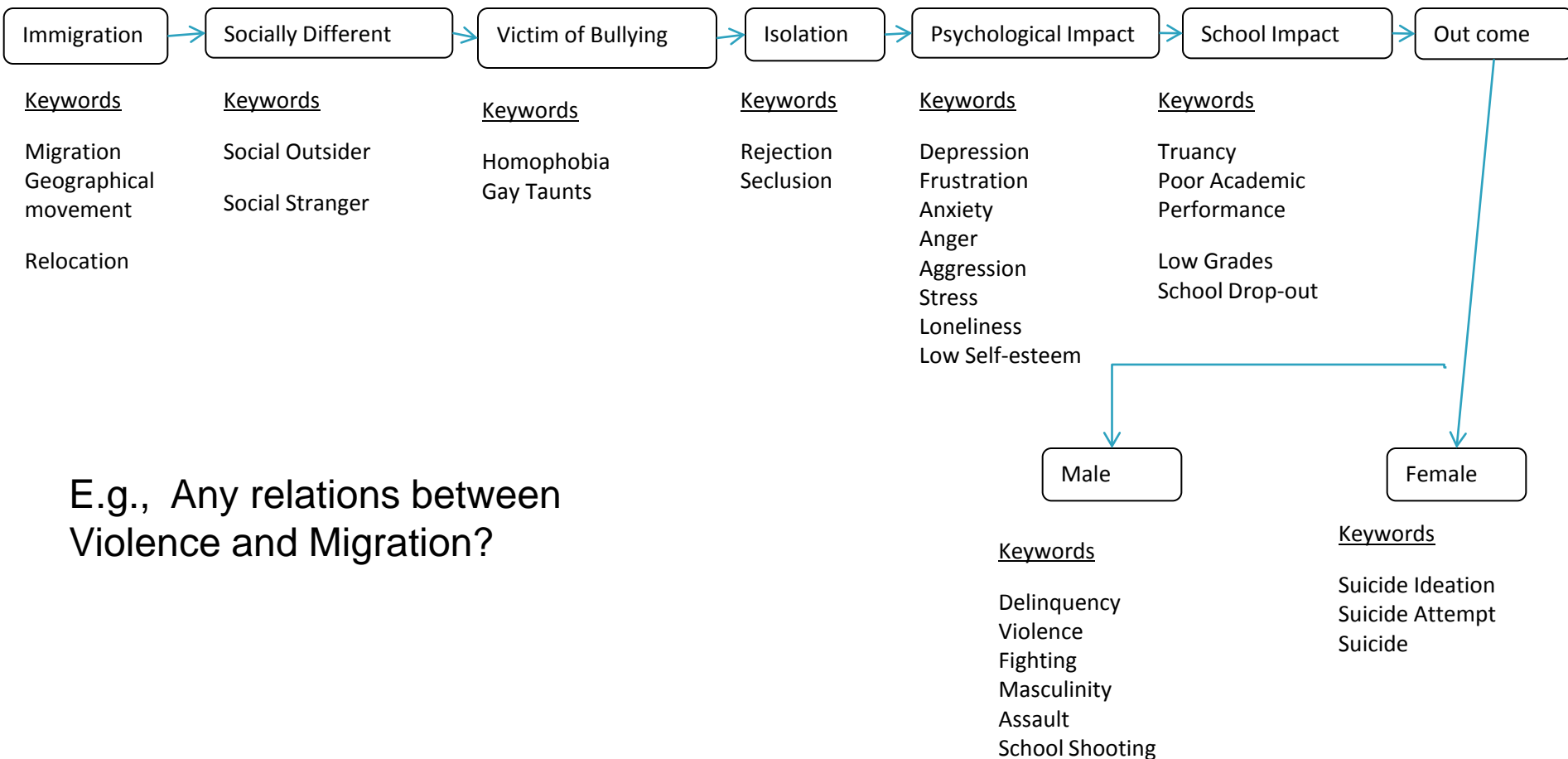
VT-416 with SVM

Storytelling

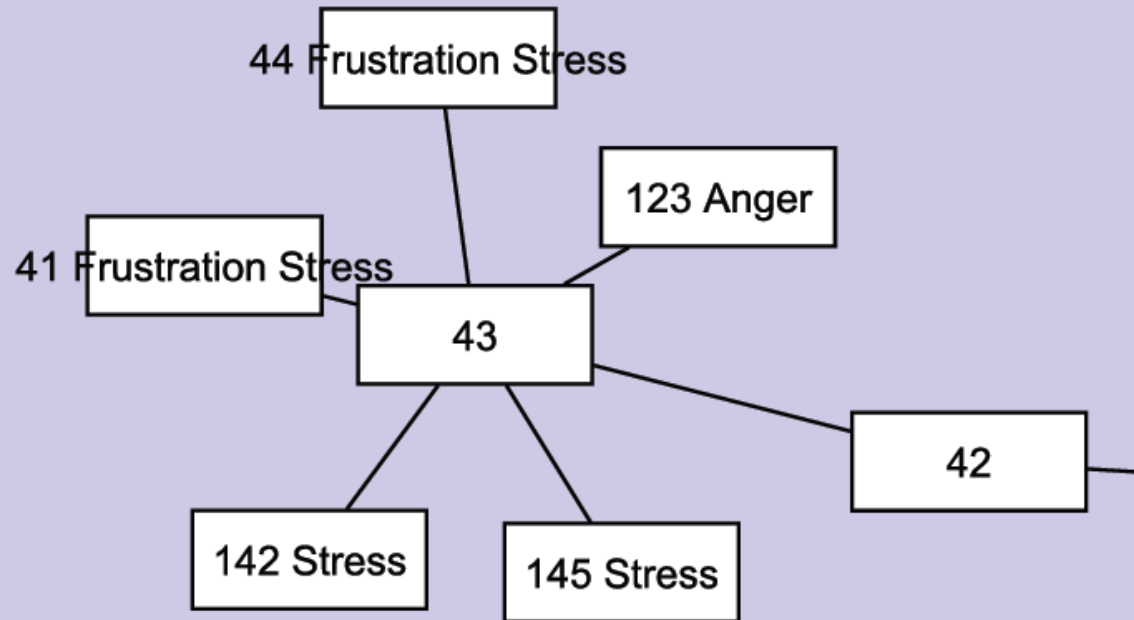


- ▶ Goal: Discover relationships between tragic events.
 - Examples: Shooting, study pressure, and mental disorders

Hypothesis of school shooting



Result



Suicide has relations with anger, stress, and frustration.

CNI for CTR – Please join!

Social Networks	Information Content	Services (Automated)
Organizations <ul style="list-style-type: none">•Schools•Gov agencies•Relief agencies	Purposes <ul style="list-style-type: none">•Assist•Plan•Recover•Understand	Information <ul style="list-style-type: none">•Analyze•Collect•Explore•Preserve
People <ul style="list-style-type: none">•Counseling•Emergency<ul style="list-style-type: none">•Assistance•Preparedness•Police•Public relations•Victims, families	Genre <ul style="list-style-type: none">•Web<ul style="list-style-type: none">•News•Public Info•Intranets<ul style="list-style-type: none">•Official reports•Scholarship<ul style="list-style-type: none">•Articles	Collaboration <ul style="list-style-type: none">•Add to CTRnet•Connect/contact<ul style="list-style-type: none">•Build lists•Research<ul style="list-style-type: none">•Ontology building•Spot events<ul style="list-style-type: none">•Seed identification

Please guide & help with CTRnet!

Thank you!

<http://www.ctrnet.net>

Ed Fox, fox@vt.edu, <http://fox.cs.vt.edu>

<http://fox.cs.vt.edu/talks/2009/20091214CNI09CTR.pdf>



VirginiaTech