



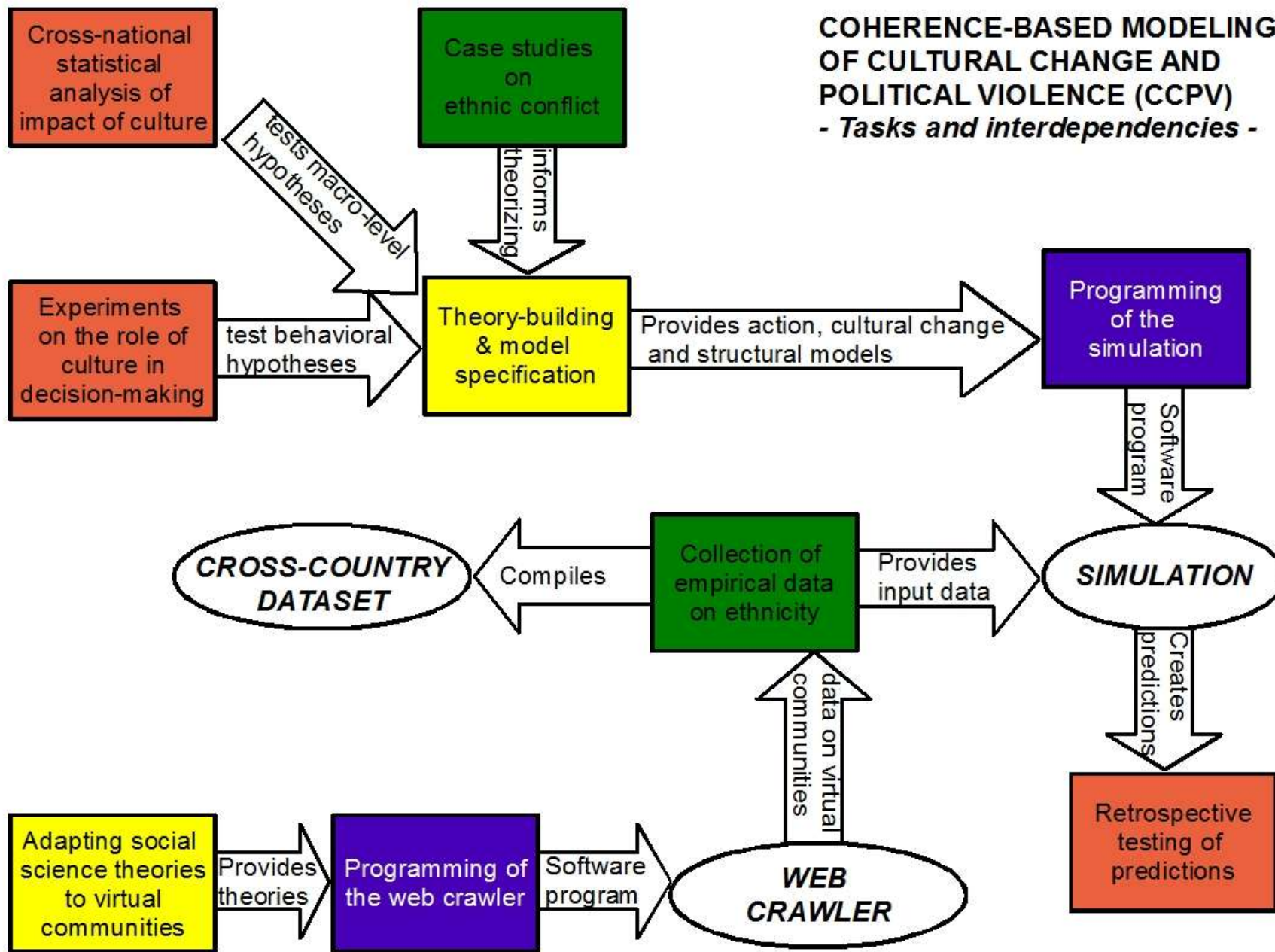
CLASSIC: Software Platform for Solving Real-World Problems by Understanding Virtual Communities

Supported by AFOSR grants FA9550-07-1-0253 (CCPV) and FA9550-09-1-0261; ONR grant N00014-09-1-0805.



Role in CCPV Project

COHERENCE-BASED MODELING OF CULTURAL CHANGE AND POLITICAL VIOLENCE (CCPV) - Tasks and interdependencies -





Past and Present Members of CCPV Team

Sun-Ki Chai	(Dept. of Sociology)
David Chin	(Dept. of Information & Computer Sciences)
Scott Robertson	(Dept. of Information & Computer Sciences)
Gurdal Arslan	(Dept. of Electrical Engineering)
Min-Sun Kim	(Dept. of Speech Communication)
Jang-Hyun Kim	(Dept. of Speech Communication)
Mooweon Rhee	(Dept. of Management & Industrial Relations)
Ruben Juarez	(Dept. of Economics)
Keisuke Nakao	(Dept. of Economics, Hilo)
Ilan Noy	(Dept. of Economics)
Katherina Sherstyuk	(Dept. of Economics)
Abdul Karim Khan	(Dept. of History, LCC)
Leonard Andaya	(Dept. of History)

Research Assistants: Aaron Herres, Dong-Wan Kang, Evan Yazawa, Suren Dorj, Ming Liu, Debbie Gundaya, Kiran Sagoo, Cecilia Noble, Quincy Edwards, Turro Wonkaren, Kar-Hai Chu, Leo DeCandia, Robert Puckett



Why the Need for General Web Social Analysis Application?

- The internet growing importance of as medium for communication, social interaction, commerce, and political activity.
- Amount of data available is unprecedented.
- Most data publicly available, accessible from anywhere in the world, and can be collected with minimal intrusion on human subjects.
- Business, government, academia, are rapidly turning to internet as a resource to analyze social phenomena.
- Despite the quantity and accessibility of data, relatively few large-n studies of web social dynamics.
- Absence of convenient technology for collecting and organizing this data is the major roadblock.
- Expansion in practical uses of search technology in last several years disappointing.



General Kinds of Things You Might Want to Find Out

- Finding which combination of content, link, “annotation”, and offline (geographical, demographic) characteristics are most significant in determining a site or member’s popularity, power, influence, etc. Subset by target “audience”.
- Determining attitudes by site/member along multiple user-specified dimensions towards a particular entity or issue. Associating with lifestyle, ideology, inferred SES, demographic characteristics.
- Measuring changes in attitudes over time and the generation and diffusion of ideas across networks.
- Using virtual community analysis to identify the key concepts and areas of conflict that define a particular virtual community. What are their ancillary interests?
- Using attitudinal and relational data, along with available structural information, to predict behavior and/or explain the processes behind behavior.



“Reality Mining” of Web: A Platform for Answering Social Questions

- Improved methods for identifying boundaries of virtual communities of interest.
 - group network measures, discourse theme coherence, homogeneity characteristics
- More versatile methods for measuring the attitudinal, ideological, and cultural of defined populations
 - plug-in dictionaries, limited syntactic parsing, mini-dictionaries for general cultural typologies, inferred homogeneity
- Multi-method validation of such methods through comparison with established social science methodologies.
 - human coder comparison, survey-based indices, computer-mediated experiments, event history analysis
- General framework for generating short and long-term behavioral predictions
 - rational choice model with coherence-based endogenous preference change



4. More Control Flow Tools

4.1. if Statements

4.2. for Statements

4.3. The range Function

4.4. break and continue Statements, and Else Clauses on Loops

4.5. pass Statement

4.6. Defining Functions

4.7. More on Functions

4.8. Intermezzo: Coding Style

previous | next | modules | index

if statements known from other languages, with some twists.

elif is short for else if, and is useful to avoid excessive indentation. An if ... elif ... elif ...

```
>>> # meaning: loop strings:
... a = ['cat', 'window', 'defenestrate']
... for x in a:
...     print x, len(x)
...
cat 3
window 6
defenestrate 12
```

It is not safe to modify the sequence being iterated over in the loop (this can only happen for mutable sequence types, such as lists). If you need to modify the list you are iterating over (for example, to duplicate selected items) you must iterate over a copy. The slice notation makes this particularly convenient:

```
>>> for x in a[:]: # make a slice copy of the entire list
...     if len(x) > 6: a.insert(0, x)
```



“Intermediate” User Initiation/Runtime/Result

File Options Analysis Tools Help

--CLAS\$IC--

Seed sites and key terms:

www.example.com, www.seed.jp, www.aSeedSite.org, sports, hiking

Rank results by:

☐ Broker Power

☒ Authority

☐ Similarity

☒ Prestige

☒ Influence

☐ Popularity

Search

Advanced Search Options

Community members

This is an Example web page title

www.exampleURL.com

Put description here. Limit: 1 line.

<more information> <reseed search> <analyze this forum>

This is an Example web page title

www.exampleURL.com

Put description here. Limit: 1 line.

<more information> <reseed search> <analyze this forum>

This is an Example web page title

www.exampleURL.com

Put description here. Limit: 1 line.

<more information> <reseed search> <analyze this forum>

This is an Example web page title

www.exampleURL.com

Put description here. Limit: 1 line.

<more information> <reseed search> <analyze this forum>

Select all

Hide details

Visualizations

Data Content Options

Website Type Options

Data Content Options

Data content types

Emotions

Content type entries

Emotions
Positive
wonderful
great
Negative
Uncertain
maybe
perhaps
perchance

Chosen content

good
excellent
could be

Add

Remove

Clear

Graph Types

Pie Graph



Line Graph



Bar Graph



Network Graph





Using Online Data to Predict Behavior: Example of Grid-Group Framework

Grid: extent to which social rules prescribe and restrict action

Group: extent to which identity is directed towards others

Closely related to classical sociological concepts of regulation and integration

- ❑ Operationalization methods straightforward and well-tested
- ❑ Works well as front-end to “thin” rational choice models of decision- making
- ❑ Fits with abstract dimensions of social organization found in social theories, e.g. regulation and integration



Using Coherence Model to Predict Preference and Belief Change

Expected Regret (single-period, individual form):

$$d = \int_s (u(s, a^*(s)) - u(s, a)) p(s) ds$$

where

$$a^*(s) = \operatorname{argmax}_{a \in A} u(s, a)$$

$$a = \operatorname{argmax}_{a \in A, s \in S} \int_s u(s, a) p(s) ds$$

s states of the environment, a actions, u utility function, and p subjective probabilities

adjustment of g, h to minimize d within “non-yogic” utility and information constraints



A Priori Validation: Grid-Group and Behavior in Voluntary Contributions Mechanism

		Shuffled			Partner			Pooled		
		Contribution Level, No Pun	Contribution, with Punish	Punishment Expenditure	Contribution Level, No Pun	Contribution, with Punish	Punishment Expenditure	Contribution Level, No Pun	Contribution, with Punish	Punishment Expenditure
Grid	Pearson correlation	-0.100	0.038	0.263**	0.261*	-0.075	0.056	0.108	-0.002	0.132
	Sig. (2-tailed)	0.392	0.748	0.022	0.052	0.582	0.683	0.217	0.983	0.131
Group	Pearson correlation	0.350***	0.096	-0.116	0.251*	0.171	-0.218	.319***	0.138	-0.155*
	Sig. (2-tailed)	0.002	0.411	0.318	0.062	0.208	0.106	0.000	0.114	0.077
N		76	76	76	56	56	56	132	132	132

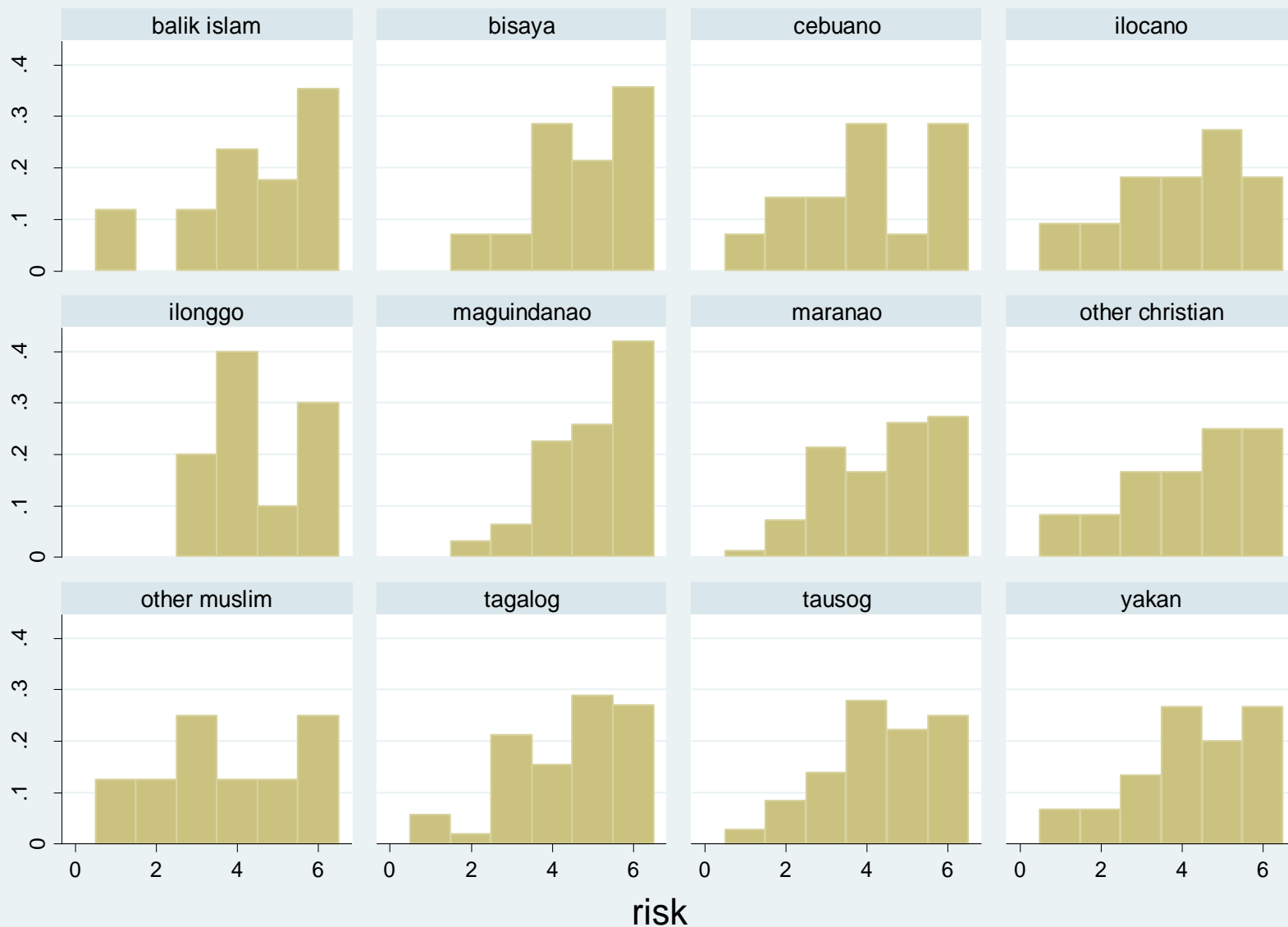
***Correlation is significant at the 0.01 level (2-tailed).

** Correlation is significant at the 0.05 level (2-tailed).

* Correlation is significant at the 0.10 level (2-tailed).



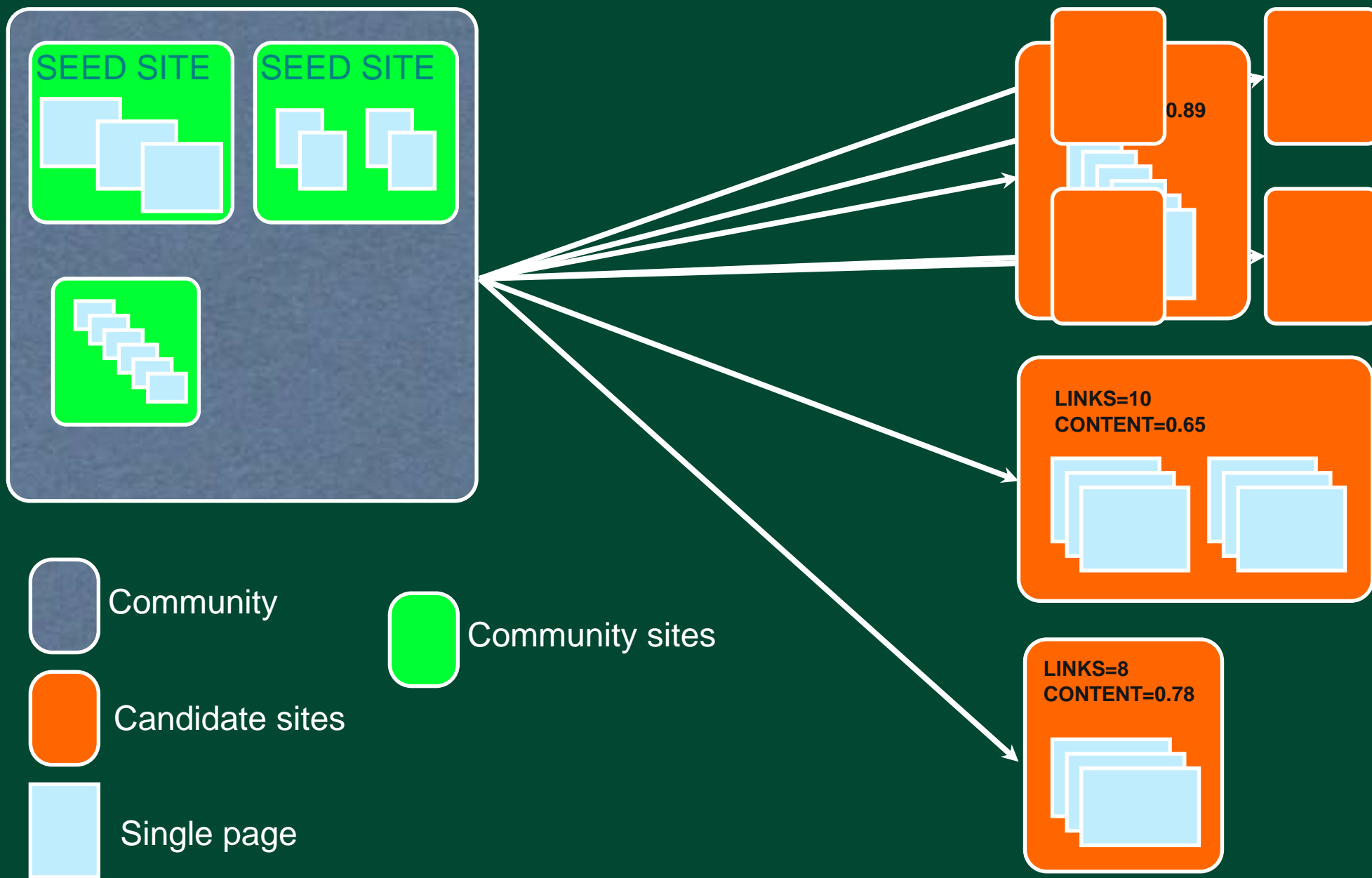
Field Study: Risk Preferences Among Moro Ethnic Groups



Graphs by eth2



How It Works





Forum Analyzer

ConfigurationResultsForum Analyzer

Forum site to crawl

http://www.v8buick.com/forumdisplay.php?f=24

Forum crawl

Start Date: 06/01/2009

End Date: 06/28/2009

Thread # to collect: 30

Start Crawl

Select a forum to view the analysis

http://www.v8buick.com/forumdisplay.php?f=24

Current: http://www.v8buick.com/forumdisplay.php?f=2

OverviewCommunity MetricsMember NetworkContent Analysis

Raw Text Message:

ese, are they?
Joe, Not sure on the switch aspect of the mirror. But I think you can get this place to re-silver the mirror. You need to send just the mirror glass.
Hey, looking for a (cheap) place to stay in Albuquerque for one night. going to be passing through on the way to Vegas from Dallas. pullin outta here (dallas) early tuesday morning. if anyone knows of a clean but cheap place to stay or something like that, let me know
Super 8 Motel on I-40, exit 166. Only US \$ 42.85
the all GM car show is at Woodley Park next sunday, June 28th. who is going? Basketball Sam says he is coming up, i should be there with a mmmmm black ummmmmm..... racecar?
Will try to make it. Just moved to Palm Springs from L.A. No Buick yet, but sounds like fun. Would like to meet B Sam too.
To anyone that's never been there: it has always been a great event. I would like to attend but might not be able to go. Some stupid little Toyota will be showing up here next weekend to get it's bad struts replace

Show Word FrequencyTerm ExtractionSave As TXT

Forum Analyzer: Community Metrics

Community Metrics

Measurements	Point
Mean Reply Depth	1.0
Non-lurker rate	3.0
Post # per member	2.0

Member Communication Network

PICKING



Attitudinal Content Tracking

Analyze in real time high-level social constructs drawn from one of multiple established content analysis dictionaries that can be plugged into system.

Constructs can be incorporated into crawling algorithm inclusion criteria
Raw n-grams also collected and can be incorporated as well.

Applet Viewer: org/ccpv/forums/ForumApplet.class

Applet

Configuration Results Forum Analyzer

Forum site to crawl

http://www.priuschat.com/forums/

Add Delete

http://www.priuschat.com/forums/ (29 available)

☒ Filter subforum crawling

☒ Choose data collection periods

Start Crawl Save Load

Show Analysis Newbie Forum

Subforum Filtering

Crawling Priority

Subforums in order

Number of Subforums

1

Apply

29 available

Data Collection Period

Start Date: End Date:

01-01-2010, 00:00 ~ 01-24-2010, 23:59

01-01-2009, 00:00 ~ 01-24-2009, 23:59

Add Delete

Overview Community Metrics Member Network Content Analysis

Category Count Probability

WordCategory	3441	0
Emotion	1455	0
affect	616	0
anger	25	0
anx	24	0
negemo	152	0

01-01-2010, 00:00 ~ 01-24-2010, 23:59

Save As TXT Show Text Dictionary Report

Show Dictionary Word Frequency

Emotion Category

pos. emo neg. emo anger anxiety sad

Applet started.



Member Network Analysis

Includes:

Connections through
thread sharing

Time prioritization

Rank of reply
prioritization

All previous filtering
types on threads

Applet Viewer: org/ccpv/forums/ForumApplet.class

Applet

Configuration Results Forum Analyzer

Forum site to crawl

☒ Filter subforum crawling

☒ Choose data collection periods

Subforum Filtering

Crawling Priority

Subforums in order

Number of Subforums

1

29 available

Data Collection Period

Start Date: End Date:

01-01-2010, 00:00 ~ 01-24-2010, 23:59

01-01-2009, 00:00 ~ 01-24-2009, 23:59

Overview Community Metrics Member Network Content Analysis

Count Author Id Number of Posts Probability (posts by author / total posts)

1	shah	11	0.04400
2	F512M	9	0.03600
3	Eisbaer	8	0.03200
4	KLear	7	0.02800
5	Tideland Prius	7	0.02800
6	usbseawolf2000	6	0.02400

01-01-2010, 00:00 ~ 01-24-2010, 23:59

Post number per Author



Contractionist Behavioral Simulation: Identity Groups Predicted, Not Assumed

Unregistered HyperCam 2. BasicTest.class



HyperCam 2 Recording

Screen Area | Hot Keys | AVI File | Sound | Other Options

Start X: 0 Start Y: 0

Width: 1512 Height: 956

When recording:

- ☒ Show rectangle around recorded area
- ☐ make this rectangle blink
- ☐ Leave HyperCam Window Opened
- ☒ Iconize HyperCam Window to the Task Bar
- ☐ Hide HyperCam Window
- ☒ Capture layered/transparent windows (may slow down)

Stop Rec. Pause Rec. Play Default

There is a new version of HyperCam at Hyperionics here to find out about version 2.16.03

Applet started.