Soc. Netw. Anal. Min.
DOI 10.1007/s13278-012-0072-x

ORIGINAL ARTICLE

A system for ranking organizations using social scale analysis

Sukru Tikves · Sujogya Banerjee · Hamy Temkit · Sedat Gokalp · Hasan Davulcu · Arunaba Sen · Steven Corman · Mark Woodward · Shreejay Nair · Inayah Rohmaniyah · Ali Amin

Received: 20 December 2011 / Revised: 4 April 2012 / Accepted: 20 April 2012

Abstract In this paper, we utilize feature extraction and model-fitting techniques to process the rhetoric found in the web sites of 23 Indonesian Islamic religious organizations to profile their ideology and activity patterns along a hypothesized radical/counter-radical scale, and present an end-to-end system that is able to help researchers to visualize the data in an interactive fashion on a timeline. The subject data of this study is 37,000 articles downloaded from the web sites of these organizations dating from 2001 to 2011. We develop algorithms to rank these organizations by assigning them to probable positions on the scale. We show that the developed Rasch model fits the data using Andersen's LR-test. We create a gold standard of the ranking of these organizations through an expertise elicitation tool. We compute expert-to-expert agreements, and we present experimental results comparing the performance of three baseline methods to show that the Rasch model not only outperforms the baseline methods, but it is also the only system that performs at expert-level accuracy.

S. Tikves (✉) · S. Banerjee · H. Temkit · S. Gokalp · H. Davulcu · A. Sen · S. Corman · M. Woodward · S. Nair
Arizona State University, P.O. Box 87-8809,
Tempe, AZ 85281, USA
e-mail: sukru@asu.edu; stikves@asu.edu

S. Banerjee
e-mail: sujogya@asu.edu

H. Temkit
e-mail: mtemkit@asu.edu

S. Gokalp
e-mail: sgokalp@asu.edu

H. Davulcu
e-mail: hdavulcu@asu.edu

A. Sen
e-mail: asen@asu.edu

1 Introduction

Being able to asses information on radical and moderate actors in a geographic area is an important research topic for national security. Radicalism is the ideological conviction that it is acceptable, and in some cases, obligatory to use violence to effect profound political, cultural and religious transformations and change the existing social order fundamentally. Muslim radical movements have complex origins and depend on diverse factors that enable translation of their radical ideology into social, political and religious movements. Crelinste (2002), in his work, states that "both violence and terrorism possess a logic and grammar that must be understood if we are to prevent or control them". Therefore, analysis of Muslim radical and counter-radical movements requires attention to the global, national and local social, economic and political contexts in which they are located. Similarly, in the Islamic context, counter-radical discourse takes various different forms; discursive and narrative refutations of extremist claims, symbolic action such as ritual and other religious and

S. Corman
e-mail: scorman@asu.edu

M. Woodward
e-mail: mataram@asu.edu

S. Nair
e-mail: snair8@asu.edu

I. Rohmaniyah
Center for Religious and Cross Cultural Studies,
Gadjah Mada University, Yogyakarta, Indonesia
e-mail: rochmaniyah@yahoo.com

A. Amin
State College of Islamic Studies (STAIN),
Manado, Indonesia
e-mail: aleejtr77@yahoo.com

Published online: 30 May 2012

cultural practices, and Islamic arguments for pluralism, peaceful relations with non-Muslims, democracy, etc. The most effective counter-radicals are likely to be religiously conservative Muslims. Effective containment and defeat of radicalism depends on our ability to recognize various levels of radicalization, and detection of counter-radical voices.

In our previous work (Davulcu et al. 2010), we attempted a clustering approach to obtain "natural groupings" of a number of local non-government religious social movements and organizations in Indonesia. Social scientists on our team observed that clustering results were not fully able to separate all counter-radical or radical organizations into pure clusters. Pure radical clusters were easily identified due to high similarity among their support for violent practices. Pure counter-radical clusters were identified due to their strong reactionary opposition to violent practices through protests and rhetoric. But the rest of the groupings were mixed. We realized that binary labeling as counter-radical or radical does not capture the overlap, movement and interactivity among these organizations. In this paper, we hypothesize that both counter-radical and radical movements in Muslim societies exhibit distinct combinations of discrete states comprising various social, political, and religious beliefs, attitudes and practices, that can be mapped to a latent linear continuum or a scale. Using such a scale, an analyst can determine where exactly along the spectrum any particular group lies, and also potentially where it is heading with its rhetoric and activity.

Given the complex nature of the task, such as regional differences in local cultures, beliefs and practices, and in the absence of readily available high-accuracy parsers, highly structured religio-social ontologies, and information extraction systems; we decided to devise a multi-lingual non-linguistic text processing pipeline that relies on only statistical modeling of keyword frequency and co-occurrence information. However, we designed the system to be able to incorporate additional information extracted from the text, if available. For example, named entity recognition (NER), machine translation, and GIS-based location lookup information are part of the user-interface presentation.

We (Tikves et al. 2011) worked with social scientists on our team to come up with an orthogonal model comprising of two primary dimensions. Both dimensions, (1) radical/counter-radical and (2) violent/non-violent, are characterized as latent, partial orders of discrete beliefs and practices based on a generalization of item order in Guttman scaling (Guttman 1950) using a Rasch model (Andric 1988). A true Guttman scale is a deterministic process, i.e., if a social movement subscribes to a certain belief or practice, then it must also agree with all lower-order practices and beliefs on the scale. Of course, such perfect order is rare in the social world. The Rasch model provides a probabilistic framework for Guttman scales to accommodate for incomplete observations and measurement errors.

We have designed a web-based system to visualize this orthogonal model. The web tools provided by the system allows drilling down on specific data, and plotting the trends and trajectories of organizations on a timeline. It consists of several modules: an off-line web mining, and data-processing pipeline, two web services for application logic, and an AJAX-based presentation layer. The web-based interface built for this study can be accessed through the web site at http://www.demo.minerva-project.org. In this paper, we present several scenarios with this tool in Sect. 5.

In this paper, we present feature extraction, feature selection, and model-fitting techniques to process the rhetoric found in the web sites of 23 religious Indonesian organizations—comprising a total of 37,000 articles dating from 2001 to 2011. We aim to identify their ideology and activity patterns along a hypothesized radical/counter-radical scale, and rank them to probable positions on this scale (McPhee 1995). The automated ordering of organizations is formed by ranking the organizations according to their estimated positions on the latent scale. We used the eRm[1] package to fit the Rasch model on this data set, and identify organizations' positions based on maximum likelihood estimation (Le Cam 1990). We show that the model fits the data using the Andersen's likelihood ratio test (LR-test) (Hessen 2010). We also created a gold standard of the ranking of these organizations through an expert-opinion elicitation tool, and through the opinions of three ethnographers on our team who collectively possess 35 years of scholarly expertise on Indonesia and Islam. We computed expert-to-gold standard agreements, as well as compared the performance of three different baseline computational methods to show that the Rasch model presented here not only performs the best among the baseline methods but that it is also the only method that performs at an expert level of accuracy.

1.1 Organization of the paper

Next section provides an introduction to the theory of Guttman scaling and Rash models. Section 3 defines the problem, presents the system architecture, and the methods used to solve the problem. Section 4 describes the Indonesian corpus, expert-opinion elicitation tool, baseline computational methods, and experimental evaluations. Section 5 discusses the user-interface designed for navigating our findings. Section 6 concludes the paper.

2 Introduction of Guttman scaling and Rasch model

In social science, *scaling* is a process of measuring and ordering entities called *subjects*, based on their qualitative

[1] http://www.r-forge.r-project.org/projects/erm/.

attributes called *items*. In general, subjects are requested to respond to surveys conducted by means of structured interviews or questionnaires. Items are presented to the subjects in form of questions. Statistical analysis of the response of the subjects on the questions about items are used in scaling the subjects. Some of the widely followed scaling procedure in social science surveys are Likert scale (Likert 1932), Thurnstone scale (Thurnstone 1928), and Guttman scale (McIver 1981). In Likert scale, subjects indicate their magnitude of agreement or disagreement about an item (from strongly agree to strongly disagree) on a five- to ten-point scale. On the other hand, Thurnstone scale is a formal method of ordering the attitudes of the subjects toward the items. Guttman scaling procedure orders both the subjects and the items simultaneously with respect to some underlying cumulative continuum. In this paper, we follow the Guttman scaling process to rank the organizations based on their response on the radical and counter-radical keywords.

2.1 Guttman scaling

A Guttman scale (Guttman 1950) presents a number of items to which each subject is requested to provide a dichotomous response, e.g., agree/disagree, yes/no, or 1/0. This scaling procedure is based on the premise that the items have strict orders (i.e., the items are presented to the subjects ranked according to the level of the item's difficulty). An item "A" is said to be "more difficult" than an item "B", if any subject answering "yes" on item "A" implies that the subject will also answer "yes" on item "B". A subject who responds to an item positively is expected to respond positively to all the items of lesser difficulty. For example, to find out how extreme a subject's view is on Guttman scale, the subject is presented with the following series of items in question form. (1) Are you willing to permit immigrants to live in your country? (2) Are you willing to permit immigrants to live in your community? (3) Are you willing to permit immigrants to live in your neighborhood? (4) Are you willing to permit immigrants to live to your next door? (5) Are you willing to permit your child to marry an immigrant? If the items form a Guttman scale, any subject agreeing with any item in this series, will also agree with other items of lower rank-order in this series. Guttman scale is a deterministic process and the score of a subject depends on the number of affirmative responses he has made on the items. So, a score of 2 for a subject in the above Guttman scale not only means he has given affirmative response to two of the questions or items but also indicates that he agrees with two particular questions, namely the first and second. Scores in Guttman scale can also be interpreted as the "ability" of a subject in answering questions sorted in increasing order of "difficulty". These scores when presented on an underlying scale, give us an ordering of the subjects based on their "ability" also.

The objective of our paper is to order the Indonesian Islamic organizations based on their views on religio-social keywords which have an inherent ordering. For example, two such keywords are "Quran" and "Sharia". An organization supporting "Sharia" will also likely to "believe in Quran". So it makes sense to use Guttman scaling procedure to rank the organizations and their beliefs and practices. One drawback of Guttman scale is that it is deterministic and assumes a strict ordering of the items. In real world, it is difficult to order all the items in such a strict level of increasing difficulty, therefore, perfect scales are not often observed in practice. Furthermore, many times, the order of the items are not known since they are not straightforwardly comparable. In addition, measurement errors might lead to responses that do not strictly fit the ordering. As a result, we can no longer conclude deterministically that if a subject answers a question affirmative, whether she will be able to give affirmative answers to other questions of lower order in the same questionnaire. We use Rasch model to overcome this drawback by taking into account measurement error.

2.2 Rasch model

Rasch model (Andric 1988) provides a probabilistic framework for Guttman scales. In Rasch model, the probability of a specified binary response (e.g., a subject agreeing or disagreeing to an item) is modeled as a function of subject's and item's parameters. Specifically in the simple Rasch model, the probability of a positive response (yes) is modeled as a logistic function of the difference between the subject and item's parameters. Item parameters pertain to the difficulty of items while subject parameters pertain to the ability of subjects who are assessed. A subject of higher ability, related to the difficulty of an item, has higher probability to respond to a question affirmatively. In this paper, Rasch models are used to assess the organizations degree of being radical or counter-radical based on the religio-social keywords (items) appearing in their rhetoric.

Rasch model also maps the responses of the subjects to the items in binary or dichotomous format, i.e., 1 or 0. Let Bernoulli variable X_{vi} denotes the response of a subject v to the item i, variable θ_v denotes the parameter of "ability" of the subject v and β_i denotes the parameter of "difficulty" of an item i. According to the simple Rasch model, the probability that the subject v responds 1 for item i is given by:

$$P(X_{vi} = 1|\theta_v, \beta_i) = \frac{\exp(\theta_v - \beta_i)}{1 + \exp(\theta_v - \beta_i)}.$$

Rasch model assumes that the data under analysis have the following properties.

1. *Unidimensionality* $P(x_{vi} = 1|\theta_v, \beta_i, \alpha) = P(x_{vi} = 1|\theta_v, \beta_i)$, i.e., the response probability does not depend on other variable
2. *Sufficiency* sum of responses contains all information on ability of a subject, regardless which item it has responded
3. *Conditional independence* for a fixed subject, there is no correlation between any two items
4. *Monotonicity* response probability increases with higher values of θ, i.e., subject's ability.

Items with $s_i = \sum_v^n x_{vi}$ value of 0 or n, and subjects with $r_v = \sum_i^k x_{vi}$ value of 0 or k are removed prior to estimation, where n is the total number of subjects and k is the total number of items. Running Rasch model on the data gives us an item parameter estimate or a score for each item. In general, the estimation of β_i or score for an item i is calculated through conditional maximum likelihood (CML) estimation (Pawitan 2001). The conditional likelihood function for measuring item parameter estimate is defined as:

$$Lc = \prod_v P(x_{vi}|r_v) = \frac{\exp(-\beta_i s_i)}{\prod_r \sum_{x|r} \exp(-\beta_i x_{vi})}$$

where r represents the sum over all combinations of r items. Similarly, the maximum likelihood is used to calculate subject parameter estimation θ_v or score for each subject. Expectation-maximization algorithms (Hunter 2004) are used in implementing CML estimation in Rasch model. We can also assess whether the data fit the model by looking at goodness of fit indices, such as the Andersen's LR-test.

To evaluate the quality of these measurements, we run Anderson LR-test (Hessen 2010) on the set of data. The test gives us a goodness of fit of the data in Rasch model, i.e., it tells us whether the data follows the assumptions of Rasch model. A p value, returned by the test, indicates the goodness of fit and a p value[2] higher than 0.05 indicates no presence of lack of fit.

2.3 Implementing Rasch model in the text mining domain

In this paper, we use Guttman scaling and Rasch model to find a ranking of 23 religious organizations based on extremity of their views are on radicalism and counter-radicalism. In our application, Rasch-model *subjects* correspond to a group of religious organizations, and *items* correspond to a set of keywords for socio-cultural, political, religious *radical* and *counter-radical* beliefs, and practices. An organization responding "yes" to a feature means the organization exhibits that feature in its narrative, while an organization responding "no" to a feature indicates that the organization does not exhibit such a feature. *Difficulty* of an item translates to *strength* of the corresponding attitude in defining radical or counter-radical ideology of any organization. Similarly *ability* of a subject in this case means the *degree* of radicalism or counter-radicalism exhibited by an organization's rhetoric. Other works in text-mining domain, such as sentiment analysis, have used Rasch model in their analysis (Drehmer et al. 2000). Details of keyword extraction and selection are presented in Sect. 3.3.

3 Methods

3.1 Problem definition

The primary goal of this study is to build a semi-automated method to rank religious organizations from a certain geographical region on a scale of radicalism versus counter-radicalism using their web sites. The efficacy of the generated model is evaluated by comparing it against baseline methods and expert-level performance. In addition to accomplishing these goals, we also present an end-to-end system architecture, and a graphical user-interface design to facilitate faceted search and browsing of this corpus.

3.2 System architecture

A summary of the system architecture can be seen in Fig. 1. The system is a composition of four components: a data-gathering component, which does web crawling, and text extraction; a scale generation component, performing scaling algorithms; application services component, which consists of several web services, and finally, a web user-interface component, presenting the data to the end user.

3.2.1 Data gathering

Initially, social scientists are invited to use their domain and area expertise to identify a set of *organizations*, and hypothesize any number of unipolar or bipolar *scales* that could explain the variance among their beliefs and practices. Next, a set of web crawling scripts are created for extraction of articles from those organizations' web sites. For each organization's corpus, we extract their top-k n-grams, and a union of all these phrases are presented to experts for feature selection. Downloaded articles are then converted into XML structures, containing their original text, their set of keywords, and extracted information such as person, location and organization names using a NER tool for Indonesian language, and their machine translations into English.

[2] http://www.en.wikipedia.org/wiki/P-value.

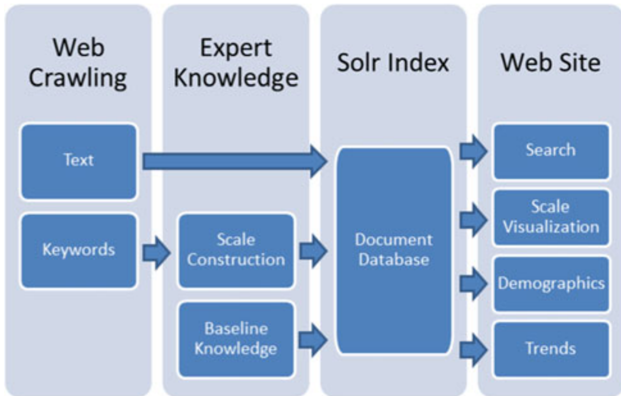

Fig. 1 An overview of the system architecture

An example document snippet is shown in Fig. 2. Here the original input (*content, source*), and a sample of the automatically extracted information corresponding to *DATE, PERSON*, and *LOCATION* can be seen. The corresponding XML versions for each input document are then stored in a document database for processing.

3.3 Keyword extraction and selection

To identify candidate keywords, one option was to translate the documents into English and apply readily available keyword-extraction methods (Michael 2010). However, it was preferable to preserve the original expression of the phrases in the original language. Hence, we utilized a non-linguistic technique that relies only on statistical occurrence, and frequency information.

Within each document, the words were separated by whitespace or punctuation marks. We considered each keyword to be an n-gram of one to three words. We treated each organization as one document and calculated the term frequency-inverse document frequency (TF-IDF) (Salton 1988) values for every single n-gram mentioned by these organizations. Top 100 n-grams with the highest TF-IDF values from each organization were used to generate a candidate list of topics that these organizations discuss most frequently. Next, we asked our team of experts to screen and manually select identify {social, political, economic, and religious} keywords corresponding to beliefs, goals and practices. During this process, our team of experts screened a total of 790 candidate keywords and they selected 29 keywords for inclusion in the radical scale, and 26 keywords for inclusion in the counter-radical scale.


```
<doc>
<field name="source">Muhammadiyah</field><field name="type">CounterRa
<field name="URL">http://www.muhammadiyah.or.id/Berita-Persyarikatan-
<field name="title">33 Rumah Sakit, dan 500 Tenaga Medis, Siap Dukung
<field name="PERSON">Demikian</field>
<field name="PERSON">Ahmad Muttaqin Alim</field>
<field name="ORGANIZATION">Seabad Muhammadiyah</field>
<field name="LOCATION">Mandala Krida</field>
<field name="LOCATIONENG">Mandala Krida</field>
<field name="date">29/06/2010</field>
<field name="event_dt">2010-06-29T00:00:00Z</field><field name="longl
<field name="content">Yogyakarta- Sebanyak 33 rumah sakit Muhammadiya

Demikian disampaikan Ahmad Muttaqin Alim, sekretaris tim kesehatan Mu
```


Fig. 2 A portion of a document represented in the system

3.4 Debates and perspective analysis

Upon inspecting the keywords selected by our team of experts, we observed that some of these keywords correspond to differing perspectives on a set of topics that are debated within these web sites. Definition of **debate** is “a formal discussion on a particular topic in a public meeting or legislative assembly, in which opposing arguments are put forward”.[3] During a debate on a particular topic, like education, both radical and counter-radical organizations discuss different perspectives such as “secular multi-cultural education” versus “Sharia based religious education”.

To design an automated perspective detection algorithm, we made the following simplifying assumptions.

1. Organizations will *mostly* discuss their own perspective in a debate.
2. Organizations will *occasionally* mention others perspectives, however, then relate them back to their own perspective.

In the following sections, we present a mathematical formulation of the perspective keyword-generation problem for a given topic, provide an NP-completeness proof, and design an exact solution through an integer linear programming (ILP)-based solver. Our future work involves finding an efficient approximation algorithm for this problem.

3.4.1 Perspective keywords-generation problem

Perspective keywords-generation problem (PKGP) is defined as follows. Given a topic (a keyword) T, and two sets of documents T_R and T_{CR} where T_R contains n documents $T_R = \{D_{R,1}, D_{R,2}, \ldots, D_{R,n}\}$ and T_{CR} contains m documents $T_{CR} = \{D_{CR,1}, D_{CR,2}, \ldots, D_{CR,m}\}$. From each document $D_{R,i} \in T_R(D_{CR,j} \in T_{CR})$, we collect a set of words $W_{R,i}$, $\forall 1 \leq i \leq n$ ($W_{CR,j}$, $\forall 1 \leq j \leq m$) which appear two words before and two words after each occurrence of the topic T in that document. Let us define W as the union of all the $W_{R,i}$, $\forall 1 \leq i \leq n$ and $W_{CR,j}$, $\forall 1 \leq j \leq m$. If the cardinality of W is p, then W can be given as $W = \{w_1, w_2, \ldots, w_p\} = \{W_{R,1} \cup W_{R,2} \cup \cdots \cup W_{R,n} \cup W_{CR,1} \cup W_{CR,2} \cup \cdots \cup W_{CR,m}\}$

Let the frequency of word w_k in document $D_{R,i}$ is given as $f_{R,i}(w_k)$ and the frequency of word w_k in document $D_{CR,j}$ as $f_{CR,j}(w_k)$.

[3] Oxford online dictionary.

Question: Are there two non-empty disjoint subsets of W, named W' and W'' and $W' \cap W'' = \emptyset$, such that for every $D_{R,i}$ $\forall 1 \leq i \leq n$,

$$\sum_{w_k \in W'} f_{R,i}(w_k) \geq \sum_{w_l \in W''} f_{R,i}(w_l) \qquad (1)$$

and for every $D_{CR,j}$ $\forall 1 \leq j \leq m$,

$$\sum_{w_k \in W'} f_{CR,j}(w_k) \leq \sum_{w_l \in W''} f_{CR,j}(w_l) \qquad (2)$$

and $|W'| + |W''| \leq K$?

In optimization version of the problem, we will try to minimize $|W'| + |W''|$.

3.4.2 Computational complexity of PKGP

Definition 1 [*Weak Partition problem* (*WPP*)] *Instance* A finite set $A = \{a_1, \ldots, a_n\}$ and a size $s(a_i) \in Z^+, \forall i, 1 \leq i \leq n$. *Question* Does the set A contain two non-empty sub-sets A_1 and A_2 that (1) $A_1 \cap A_2 = \emptyset$, (2) $A_1 \cup A_2 \subseteq A$ and (3) $\sum_{a_i \in A_1} s(a_i) = \sum_{a_j \in A_2} s(a_j)$?

WPP has been shown to be NP-complete in (van Emde Boa 1981).

Theorem 1 *PKGP is NP-complete.*

Proof It is easy to see that PKGP is in NP since a non-deterministic algorithm needs only to guess a partition of the word set W into W' and W'' and check in polynomial time if all the constraints hold for this partition and also if $|W'| + |W''| \leq K$.

WPP is a restricted version of PKGP. First we create a restricted instance of PKGP as follows: let T_R and T_{CR} contains one documents each, i.e, $T_R = \{D_{R,1}\}$ and $T_{CR} = \{D_{CR,1}\}$. Frequency of a word $w_i \in W, \forall 1 \leq i \leq n$ in document $D_{R,1}$ and $D_{CR,1}$ is taken to be equal, i.e., $f_{R,1}(w_i) = f_{CR,1}(w_i) = s(a_i)$. The parameter K is taken to be equal to $|W|$. This instance of PKGP is similar to an instance of WP in the following way: the set A contains element a_i for every word $w_i \in W$. So, $|W| = |A|$. In addition, $s(a_i) = f_{R,1}(w_i) = f_{CR,1}(w_i)$, $\forall 1 \leq i \leq n$.

If we find a weak partition of A, as sets A_1 and A_2 such that $\sum_{a_i \in A_1} s(a_i) = \sum_{a_j \in A_2} s(a_j)$, then we can find subsets of W, as sets W_1 and W_2, such that $w_i \in W_1$ if $a_i \in A_1$ and $w_j \in W_2$ if $a_j \in A_2$, respectively. In addition, $\sum_{a_i \in A_1} s(a_i) = \sum_{a_j \in A_2} s(a_j)$, implies that both the constraints $\sum_{w_i \in W'} f_{R,1}(w_i) \geq \sum_{w_j \in W''} f_{R,1}(w_j)$ and $\sum_{w_i \in W'} f_{CR,1}(w_i) \leq \sum_{w_j \in W''} f_{CR,1}(w_j)$ are true, because $s(a_i) = f_{R,1}(w_i) = f_{CR,1}(w_i)$, $\forall 1 \leq i \leq n$. Since $K = |W|$, the constraint $|W'| + |W''| \leq K$ will trivially hold. So, WPP is a restricted version of PKGP. Since WPP is known to be NP-complete, PKGP is also NP-complete.

3.4.3 Integer linear programming formulation for PKGP

We formulate an ILP to solve the PKGP optimally. For each word $w_i \in W$, we use two variables x_i and y_i. x_i is 1 if and only if the word w_i is in W_1 and y_i is 1 if and only if the word w_i is in W_2. Then constraint (3) means sets W_1 and W_2 disjoint. Constraint (4) ensures that these sets (W_1 and W_2) are also non-empty. Constraints (5) and (6) ensure the constraints 1 and 2 in problem statement. The objective minimizes the summation of cardinality of W_1 and W_2.

Variables: For each word w_i,

$$x_i = \begin{cases} 1, & \text{if word } w_i \text{ is assigned to set } \; W_1 \\ 0, & \text{otherwise.} \end{cases}$$

$$y_i = \begin{cases} 1, & \text{if word } w_i \text{ is assigned to set } \; W_2 \\ 0, & \text{otherwise.} \end{cases}$$

$$\min \quad \sum_{i=1}^{p} x_i + y_i$$

$$s.t. \quad x_i + y_i \leq 1, \quad \forall i = 1, \ldots, p \qquad (3)$$

$$\sum_{i=1}^{n} x_i \geq 1 \text{ and } \sum_{i=1}^{n} y_i \geq 1, \quad \forall i = 1, \ldots, p \qquad (4)$$

$$\sum_{w_k \in W}^{p} f_{R,i}(w_k)(x_k - y_k) \geq 0, \quad \forall i = 1, \ldots, n \qquad (5)$$

$$\sum_{w_k \in W}^{p} f_{R,i}(w_k)(x_k - y_k) \leq 0, \quad \forall i = 1, \ldots, m \qquad (6)$$

$$x_i \in \{0, 1\}, y_i \in \{0, 1\}, \quad \forall i = 1, \ldots, p \qquad (7)$$

3.4.4 Social scale generation

Social scale generation is done by building *response tables*; a pair of tables for a bipolar scale, such as radical/counter-radical (R/CR), or a single table for a unipolar scale, by thresholding the occurrence frequencies of the selected keywords in the organizations' web corpus.

The scale-generation architecture is shown in Fig. 3. Here, the flow of the processes and data can be seen as interactions between experts and automated modules. The system works as follows.

- Initially, area experts to identify a set of *organizations*, and hypothesize any number of unipolar or bipolar *scales* that could explain the variance among the beliefs and practices of the organizations.
- Next, we crawl and download the web sites of the organizations, and the system automatically *extracts the top-k candidate keywords* for consideration in the

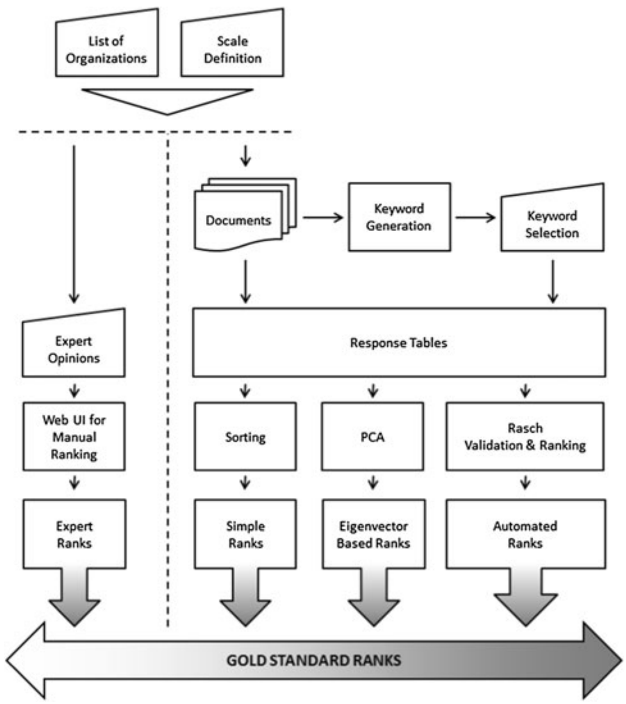

Fig. 3 A model of the system architecture

hypothesized scale. Social scientists screen the list of extracted keywords, and *select* the relevant ones for inclusion in further analysis.

– The system builds *response tables*; a pair of tables for a bipolar scale (such as radical/counter-radical R/CR), or a single table for a unipolar scale, by thresholding the occurrence frequencies of the selected keywords in the organizations' web corpus. See Figs. 4 and 5 for the response tables for the R/CR scale.
– The response tables are fed as input to the *Rasch Model building* algorithm. The algorithm produces a metric to *validate* the fitness of the model, and *rankings* of the organizations and keywords. Figures 6 and 7 show the relative positions of the organizations and keywords on the latent scales. The algorithm also produces a metric to *validate* the fitness of the model.
– Two types of other information are collected for *evaluation* purposes. First, *expert rankings* of the organizations, using a graphical drag-and-drop expert-opinion elicitation tool shown in Fig. 11. Expert rankings are merged into a consensus *gold standard* of rankings. Next, two other computational baseline methods; one based on simple sorting, and another based on principal component analysis (Jolliffe 2002), are used to generate alternative *computational rankings* shown in Fig. 12.

In addition, the data for the *violence/non-violence* are gathered using a separately developed tool, by collecting the opinion of the experts. A future work will also include automated generation of this dimension, as well.

3.5 Feature extraction

After identifying the keywords for the analysis, we needed to search the web site corpus of the organizations for the matching items. This yielded a term-document matrix.

This task was performed in a simple three-step procedure; initially, the occurrence frequencies of particular keywords were counted within each organization's corpus, then, a threshold matrix was calculated from the initial values, and finally, a binary response matrix was generated by applying these thresholds to the initial values.

The frequency metric is shown in formula 8, where k is the keyword, o is the organization, and D_o is the document set pertaining to that particular organization.

$$f_{o,k} = \frac{|\{d | k \in d, d \in D_o\}|}{|D_o|} \tag{8}$$

A threshold value for each keyword is calculated by taking the median of the values in the related column. Median was preferred over mean as a threshold, since the distribution of the values did not fit Gaussian distribution, yet median empirically proved to be a better measure.

Fig. 4 Radical subset of organizations and keywords, sorted according to aggregate row values

RADICAL	quran	Islamic teachings	Shari'a	infidel	military strength	Kufr	Caliph	prophet muhammad	prophet muhammad	Tabligh	sharia	corruption	Shirk	paganism	die shaheed	street preaching	obligation for Shari'a	Mujahideen council	the Zionist Jews	Islamic thoughts	pornography	jihad council	mujahideen	suicide bomb	communists	the idolaters	Sharia enforcement	violence	Islamic state
Org 9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	0	1	0	0
Org 7	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	1	0
Org 1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	1
Org 3	1	1	1	1	0	0	0	1	1	1	0	1	0	0	1	0	1	1	0	1	1	0	0	0	0	0	0	1	0
Org 4	1	1	1	1	0	1	0	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0
Org 8	0	0	1	0	1	1	1	0	0	1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	1
Org 2	1	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Org 5	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Org 6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Org 10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 5 Counter-radical subset of organizations and keywords, sorted according to aggregate row values

COUNTER-RADICAL	politics	equality	Hindu	pluralism	state	tolerance	gender justice	human rights	Indonesian Islam	religious freedom	intercultural dialogue	constitutional rights	democracy	election	Christian	non-Muslims	civil society groups	liberal	secularism	activists	Jewish	peace building	conflict prevention	freedom of expression	homosexual	multicultural education
Org 12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	0	0
Org 21	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	0	1	0	0	0	0	0	0
Org 15	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0
Org 11	1	1	1	0	1	1	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	1	0	0	1
Org 17	1	0	1	1	1	0	0	1	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
Org 14	1	1	0	0	1	0	1	1	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Org 16	0	1	1	1	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Org 19	1	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0
Org 18	0	1	1	1	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Org 20	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Org 13	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Org 22	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Org 23	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

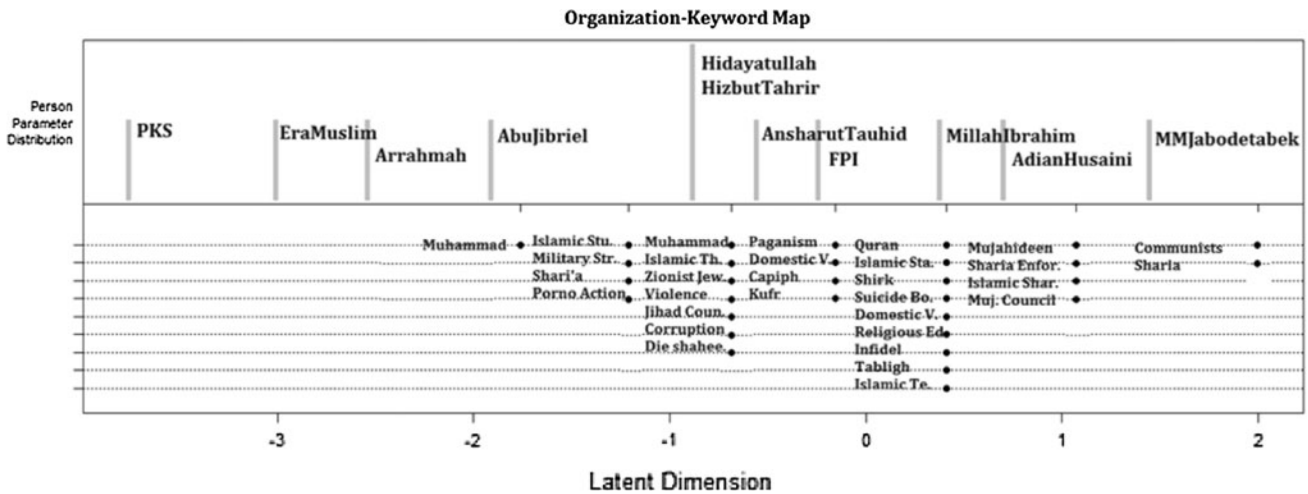

Fig. 6 Radical subset of organizations and keywords

Finally, each element was converted into a binary value by comparing it to the column's threshold. English translations of the keywords are presented for clarity in Figs. 4 and 5.

3.6 Model fitting

We fit the Rasch model on two datasets: (1) radical organizations with radical keywords and (2) counter-radical organizations with counter-radical keywords. We used the eRm package in R, an open source statistical software package,[4] to fit a Rasch model to the dataset, and obtain the organizations' scores on the latent scale, which are the the subject parameter estimates (θ_v) discussed in the previous section. The eRm package[5] fits Rasch models and provide subjects or organizations parameter estimates based on maximum likelihood estimation.

The automated scale of the organizations is formed by ranking the organizations according to their estimates on the latent scale. Not only we can provide the organization estimates but we can also assess whether the model fits the data by looking at several goodness of fit indices, such as the Andersen's LR-test.

3.7 Application services

We use two backend services in the application layer to present the data to the user interface. First, all the extracted textual information are stored in Apache Solr,[6] providing facilities like full-text search and faceting (Tunkelang 2009), using an AJAX interface. In addition, a WCF-based scaling service is used to infer scales in real time. This particular service loads the response table, and the previously generated scale data, and estimates the R/CR scale

[4] http://www.cran.r-project.org/.

[5] http://www.r-forge.r-project.org/projects/erm/.

[6] http://www.lucene.apache.org/solr/.

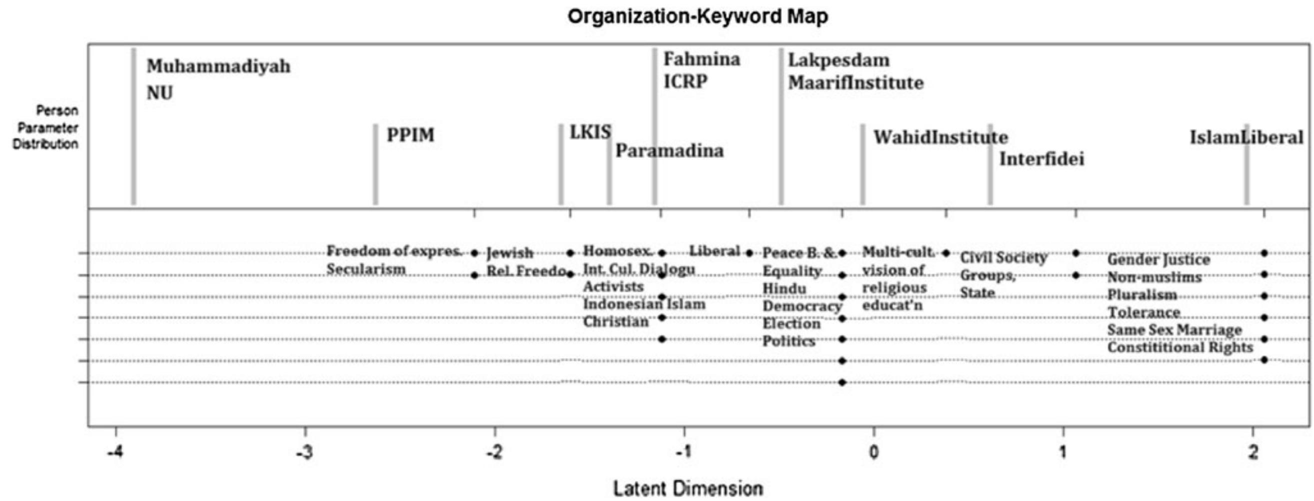

Fig. 7 Counter-radical subset of organizations and keywords

for a subset of the input. Number of positive responses are interpolated on the scale to generate the scale, and the expert opinion is used for a static violent/non-violent (V/NV) scale. While the interpolation is based on a sufficient statistics, future work on speeding up Rasch model generation for real-time use would be beneficial.

3.8 User interface

The user interface is responsible for representing our input data, and the findings to the experts in an interactive fashion. Users should be in control of the selection of the data displayed, and filtering with organization names, or a specific date range, or using other parameters such as arbitrary keywords, or geographic locations. While performing these tasks, it should provide results to the user with a minimum of delay, allowing quick drilling down to interactively model the scenarios that users have in mind.

The user interface is implemented as an interactive AJAX-based application, using ajaxsolr[7] framework. In addition to the search and navigation capabilities provided with ajaxsolr, it also adds functional widgets for visualizing the organizations on a scale, mapping the intensity of the locations, displaying demographics trends, and so on. A more detailed discussion of the user interface is provided in Sect. 5.

The presentation of the scale, however, brings the following challenges.

- It would be preferable to plot the locations on the same range as the input collection. However, the Rasch scale is on a latent range (Figs. 8, 9).
- Since this will be an interactive application, users would prefer to see almost instantaneous results. Yet, the eRm model generation is computationally expensive.

We resolve the first issue by uniformly scaling the ranges into [−10, 10], making it consistent with the inputs.

The second issue requires a more specific solution. We make use of the fact that the raw person scores pertaining to number of positive responses is a sufficient statistics for the Rasch model (G 1961) to estimate scale values on the fly. Since we know the date range, and the selected organizations currently visible in the user interface, it is possible to quickly generate a response matrix for this subset of the data, and merge it with the previously known scale information to generate interpolated scale values.

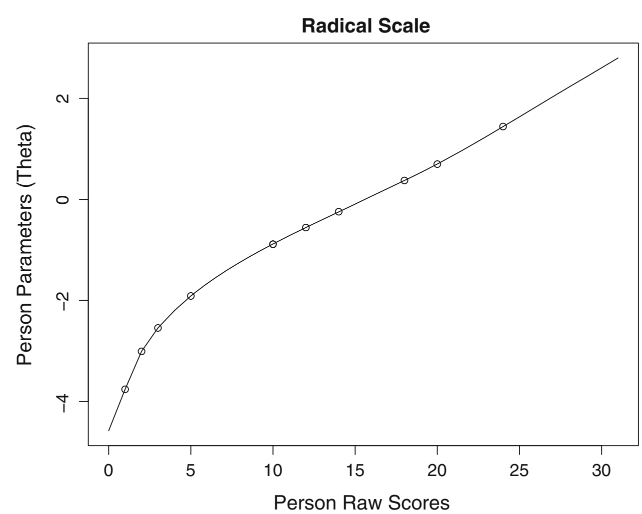

Fig. 8 Radical scale

[7] http://www.evolvingweb.github.com/ajax-solr/.

Algorithm 1 Subset scaling algorithm

1: **procedure** SUBSET-SCALE($D, O', start, end, scale$)
2: $D' \leftarrow \{d \in D \mid Date(d) \geq start \wedge Date(d) \leq end\}$
3: **if** $D' = \emptyset$ **then**
4: **return** $\emptyset$
5: **end if**

6: **for all** $o \in O$ **do** ▷ Entire set of organizations
7: $D'_o = \{d \in D' \mid Org(d) = o\}$
8: **for all** $k \in Keywords$ **do**
9: $f'_{o,k} = |\{d \in D'_o \mid k \in d\}| / |D'_o|$
10: **end for**
11: **end for**
12: **for all** $k \in Keywords$ **do**
13: $t'_k = Median(\{f'_{o,k}, o \in O'\})$
14: **end for**

15: **for all** $o \in O', k \in Keywords$ **do**
16: $r'_{o,k} = f'_{o,k} > t'_k \rightarrow (t : 1, f : 0)$
17: **end for**

18: **for all** $o \in O'$ **do**
19: $sum_o = \sum_{k \in Keywords} r'_{o,k}$
20: **end for**
21: **for all** $o \in O'$ **do**
22: $S_o = Interpolate(sum_o, scale, -10, 10)$
23: **end for**
24: **return** S
25: **end procedure**

The psuedo-code for the subset scale-generation procedure is presented in Algorithm 1. The process starts with identifying the subset of documents in the (*start*, *end*) date range (lines 2–5). Then the keyword frequencies, and thresholds are calculated for the entire set of organizations on this document subset (lines 6–14). Finally, response tables for the subset of organizations is generated (lines 15–17), and then the sums need to be interpolated (lines 18–23), to be able to generate a scale on the [−10,10] range (line 24).

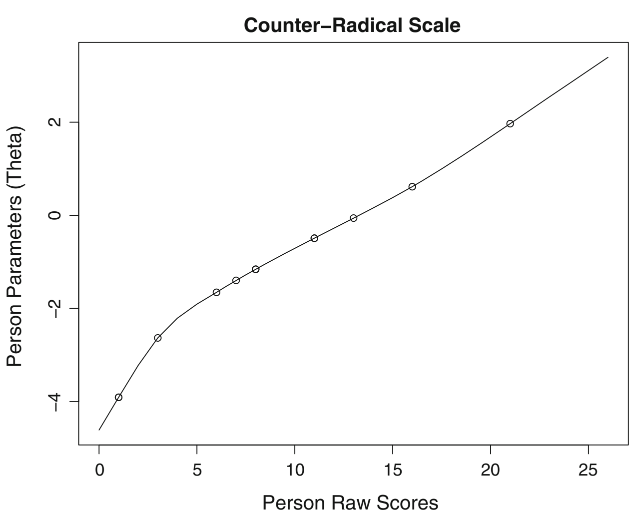

Fig. 9 Counter-radical scale

Here we have opted to include all the organizations in threshold calculations. This is because, the radical or counter-radical activity intensities are always measured relative to the other organizations participating in the same time period. However, while the scale is based on all the organizations, only the ones specifically asked will be presented to the user.

4 Experimental evaluation

4.1 Indonesian corpus

The corpus domain is the online articles published by the web sites of the 23 religious organizations identified in Indonesia, in the Indonesian language. These sources are the web sites or blogs of the identified think tanks and organizations. As discussed in the Sect. 1, each source was classified as either radical or counter-radical by the area experts. We downloaded a total of 37,000 Indonesian articles published in these 23 web sites, dating from 2001 to 2011. For each web site, a specific REGEX filter was used to strip off the headers, footers, advertising sections and to extract the plain text from the HTML code.

4.2 The quadrants model

Our project leverages the results of our previous work, which relied on social theory including Durkheim's (2004) research on collective representations, Simmel's (2008) work on conflict and social differentiation, Wallace's (1956) writings on revitalization movements, and Tilly and Bayat's studies on contemporary social movement theory (Tilly 2004; Bayat 2007). Our team has also developed, and is currently testing a theoretically based class model comprised of continuous latent scales. The first pair of scales focus on distinctions between the goals and methods of counter-radical and radical discourse, and capture the degree to which individuals, groups, and behaviors aim to influence the social order (change orientation) and the methods by which they attempt to do so (change strategies).

Quadrants model (see Fig. 10) captures multiple social trends in four quadrants (A, B, C, and D), and it makes the significant distinction between violent and not-violent dimensions of both radicalisms and counter radicalisms. Using the quadrants model, a researcher can locate organizations, individuals, and discourses in broader categories while still considering subtle differences between groups within categories. A researcher can document movement and trends from category to category, and identify points where movement is likely to happen.

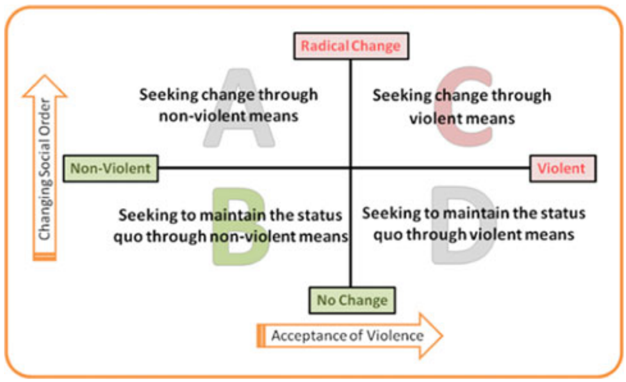

Fig. 10 The quadrants model

4.3 Expert opinion and gold standard of rankings

We collaborated with three area experts, who collectively possess 35 years of scholarly expertise on Indonesia and Islam. To build a gold standard of orderings of the organizations, we built a graphical drag-and-drop user-interface tool to collect the opinions of each of the area experts. A screenshot of the tool is shown in Fig. 11.

Each expert, separately evaluated and ranked the organizations in the dataset according to a two dimensional scale of radical/counter-radical (R/CR) and violent/non-violent (V/NV) axis. The consensus among the experts was high; since per item standard deviations among the experts' scores along the R/CR axis over a range of [−10, 10], across all organizations were 2.75. In addition, 90 % of the items have less than 22.6 % difference in their

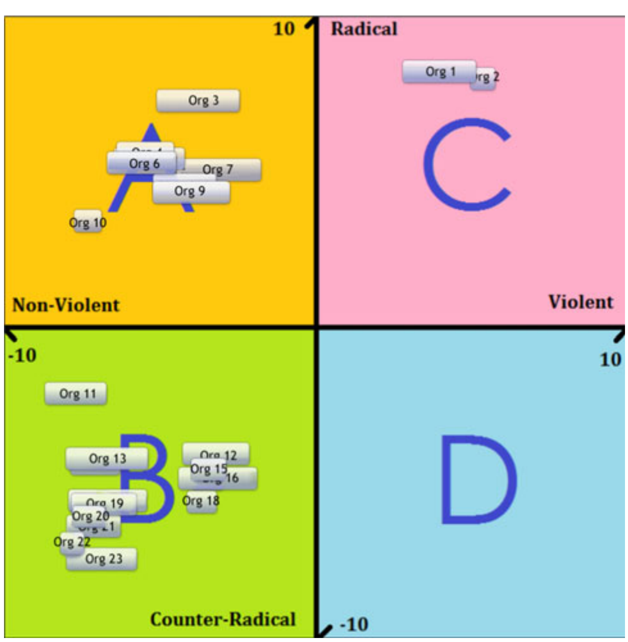

Fig. 11 The visual interface of the expert-opinion collector for manually placing the organizations on the two dimensional scale

rankings. The individual scores for each organization were combined and averaged to obtain the consensus *gold standard rankings* along the hypothesized R/CR scale.

A work is in progress for building a publicly accessible expert opinion collection toolkit. The preliminary version can be accessed at: http://www.minerva-project.org/DataCollector.

4.4 Computationally generated scale

The ranking discovered by the Rasch model fitting the corpus has been evaluated against the gold standard rankings of the organizations provided by the experts. The difference between two separate rankings have been calculated using the following misplacement error measure in Eq. (9).

$$error(G,R) = \frac{\sum_{o \in O} \frac{|G(o)-R(o)|}{|O|}}{|O|} \tag{9}$$

Here, O is the set of organizations, G and R are one to one mapping functions of rankings from set O to range $[1,|O|]$. For two exactly matching rankings, the $error(G, R)$ will be zero, whereas for two inversely sorted rankings it is expected to be 0.5 (when the size of O is even). In addition, a random ranking is expected to have a error of 0.375.

4.5 Expert-to-gold standard error

We calculated the error between each expert's ranking and their consensus gold standard of rankings. The first expert's error measure is 0.06, and the second and third expert's errors are 0.12 and 0.14 correspondingly as shown in the last row of the table in Fig. 12. The average error of our experts against their gold standard ranking is 0.11.

4.6 Baseline: sorting with aggregate score

The first baseline we used was constructed by sorting the organizations according to the number of different keywords observed in their corpus. While this provided a pattern similar to a Guttman scale, and orderings of the organizations matched to a certain degree with the gold standard as shown in Fig. 12, the *error* for this baseline was 0.19, which is higher than the average expert's performance.

4.7 Baseline: principal component analysis

A stronger baseline was built by employing principal component analysis (Jolliffe 2002), and sorting the organizations according to their projections in the first principal component of the term–document matrix. Since experts selected the R/CR scale relevant keywords only, it was expected that the first principal component would reflect the corresponding scale. PCA proved to be performing

Fig. 12 Computational and expert rankings

	Computational Rankings					Expert Rankings		
	Random	Sort	PCA	Rasch	Gold	Expert 1	Expert 2	Expert 3
Organization Rankings	21	9	1	6	1	2	7	1
	15	7	7	5	2	3	2	5
	22	1	9	2	3	1	4	8
	19	3	4	8	4	4	6	3
	4	4	12	4	5	5	3	4
	10	8	3	3	6	6	1	9
	6	2	8	1	7	8	5	2
	2	5	2	7	8	9	9	10
	17	6	17	9	9	7	8	7
	5	10	5	12	10	10	10	6
	11	23	21	21	11	14	13	12
	20	22	15	15	12	12	11	21
	13	13	11	11	13	11	19	16
	7	20	14	17	14	17	18	15
	3	18	16	14	15	15	15	19
	23	19	6	16	16	18	20	20
	1	16	20	19	17	13	22	14
	14	14	19	18	18	16	17	11
	12	17	18	20	19	22	14	18
	9	11	22	13	20	19	16	17
	16	15	13	22	21	20	23	23
	18	21	10	10	22	21	12	22
	8	12	23	23	23	23	21	13
Error	0.36	0.19	0.18	0.10		0.06	0.12	0.14

better than the aggregate score sorting, with an *error* measure of 0.18. However, this error rate is still higher than the error rate of each expert.

4.8 Performance of the Rasch model ranking system

The *p* values from the Anderson LR goodness of fit test from model (1) and model (2) (mentioned in Sect. 3.6) are 0.85 and 0.669, respectively, suggesting no evidence of lack of fit. The Rasch models allow us to get a natural order of the organizations, according to their "abilities", i.e., radicalism and counter-radicalism in this case. This system had an *error* measure of 0.10, which actually provided a higher ranking performance than the average performance of our experts'—performing better than the majority of our area experts.

4.9 Evaluations

Our experiments showed that the hypothesized compatibility of the R/CR scale for the Indonesian corpus is valid. Not only the Rasch model was statistically fitting the response matrix but also the generated ranking performance was better than the average expert performance. Among our computational baseline methods, the Rasch Model was the only method producing expert-level performance as shown in Fig. 12. This preliminary analysis with the R/CR scale shows that when experts assist the system with keyword selection, the web corpus of organizations provides rich-enough information and patterns to enable a computational method to rank them accurately.

5 Web application overview

A sample snapshot of the web application can be seen in Fig. 13. It is composed of four main widgets for visualization and navigation. The top-left section which contains the **Search and Navigation widget** (1) that allows filtering of the document subset using parametric search queries and keyword based search criteria. The top-right section is the **Quadrant widget** (2) which displays the organizations active in the currently selected time frame on a two-dimensional axis, using violence and radicalism scales. The bottom-left section consists of two **Treemap widgets** (3) which displays the demographics and the top keywords (markers) of the current selection. The bottom-right section has a **Timeline widget** (4) which provides a visualization of the keywords (markers) trends on a time line.

The navigation in the user interface starts with the Navigation widget (top-left) of the web application. Here the user is able to filter down the corpus utilizing full-text search queries, or faceting using keywords, locations, demographics, or choosing a subset of organizations.

The Quadrant widget (top-right) provides a plot of the currently selected organizations on the two dimensional scale. The radical/counter-radical (R/CR) axis is

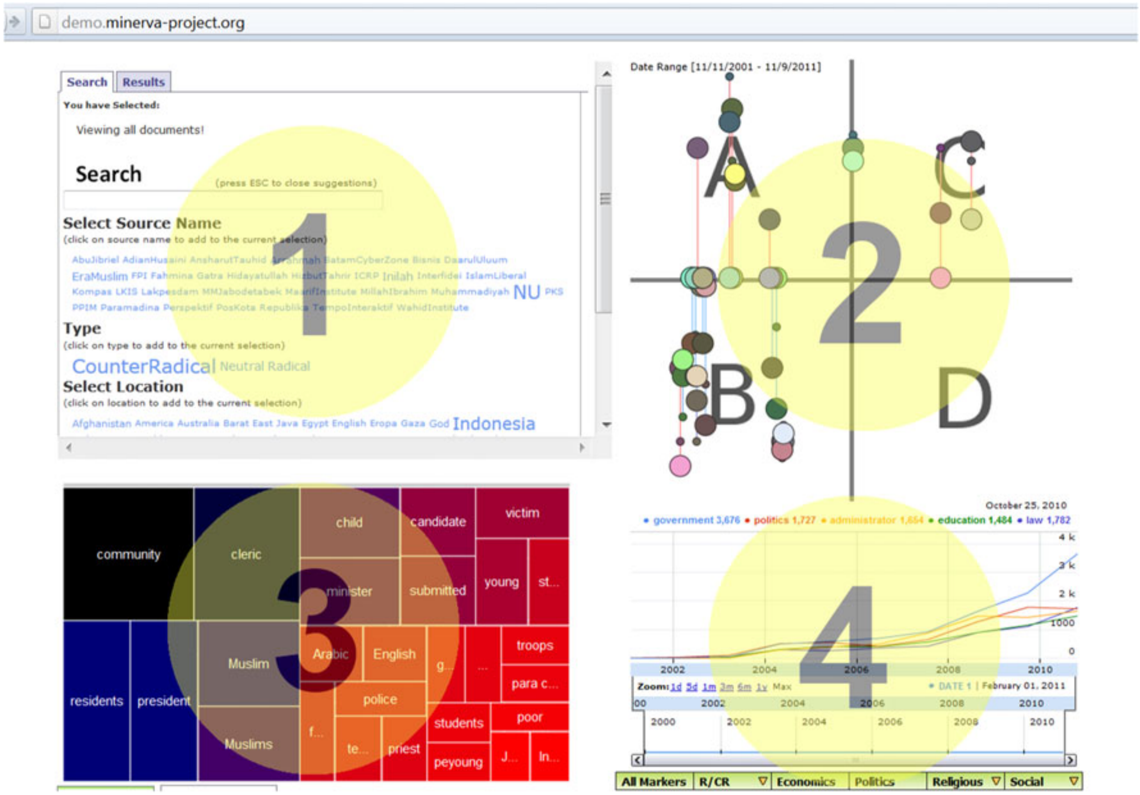

Fig. 13 A sample snapshot of the web application (color figure online)

dynamically calculated in real time, using the subset of organizations, and the time range of the current selection. The location change on the time range for each organization is shown as a color-coded path, with three markers, a light circle corresponding to the position at the beginning of the period, a dark circle corresponding to the end of the period, and a dark-small circle for the middle. A red line between the circle denotes the rise of radical activities in the organization's behavior. A blue line denotes the opposite. The smaller circle is useful to see the overall movement of an organization. For example, between the range Aug 2005 and Aug 2007, EraMuslim's activities were radical (center of A quadrant), then became almost counter radical (the smaller circle denotes this mid point in the movement), and then jumped up again. The V/NV axis is retrieved from expert opinion in the current version, and dynamic calculation of this axis is left for a future version.

The Timeline widget (bottom-right) displays the trends of the most frequent markers on a time line. Initially the subset of markers presented defaults to all available, however it is possible to restrict the selection of markers to a more limited set among radical/counter-radical, economical, political, religious, or social domains. Timeline widget can also be used for selecting a date range of interest.

The Treemap widgets (bottom-left) are used to display the relative frequencies of demographics and keywords (markers). The displayed marker category selection for this widget is synchronized with the Timeline widget.

In the following sections, we present some scenarios and findings to illustrate the capabilities of the web interface.

5.1 Scenario 1: radical organizations' trends

In this scenario, we analyze both violent and non-violent radical organizations. Our web application shows the ideologies that these organizations are propagating. We can see[8] the most prominent markers associated with these radical organizations. Markers such as "infidel", "Sharia", and "violence" show an increasing trend between 2001 and 2011. A very strict interpretation of "Sharia" is used by radical organizations to justify their actions (Widhiarto 2010; Hasan 2009). "Sharia" peaks during this period as shown in Fig. 14.

5.2 Scenario 2: C-quadrant organizations' trends

We now analyze Front Pembela Islam (FPI), an Islamic organization in Indonesia established in 1998. FPI is well known for its violent acts (Frost et al. 2010; Rondonuwu

[8] Select the filter "Radical" from the search options and then in the Markers Menu select [Religious → Radical Markers].

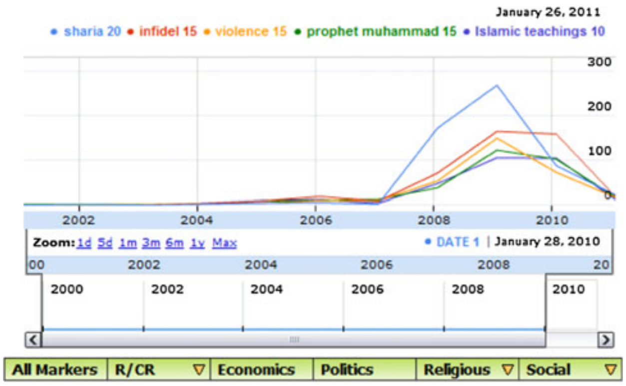

Fig. 14 Trend of radical markers

2010) justified by a strict interpretation of Sharia (for the Study of Terrorism 2011). Our documents for FPI ranges between 2000 and 2010. Using our web application's plots of the movement of FPI in the C Quadrant, we found that FPI consistently rised higher on the radical scale as shown in Fig. 15. We selected the following time ranges, 2000–2003, 2002–2006, 2006–2010 and analyzed the trends of various markers associated with FPI. There was a substantial increase in the intensity of various radical markers such as "infidel", "Mujahedin", "pornography".[9] Since 2006, we also saw a steep increase in the frequency of marker "Ahmadiyya", as shown in Fig. 16, which indicates FPI's increased opposition to this heretical sect (Rahmat and Sihaloho 2011).

5.3 Scenario 3: A-quadrant organizations' trends

We analyze Hizb ut-Tahrir also known as Hizb ut-Tahrir Indonesia (HTI), a radical organization widely believed to be non-violent (Ward 2009), which has been active in Indonesia since 1982 (Osman 2011). Between 2007 and 2009, our web application shows various radical and non-radical markers associated with this organization.

Radical	Non-Radical
"Sharia", "Infidel", "Caliph", "Violence"	"Politics", "Indonesian Islam", "Election", "Liberal", "Democracy"

During the same period, we see a steady increase in the frequency of the radical marker "Sharia". This is consistent with one of HTI's goals of implementing Sharia in Indonesia (Hasan 2009). Hizb ut-Tahrir openly propagates

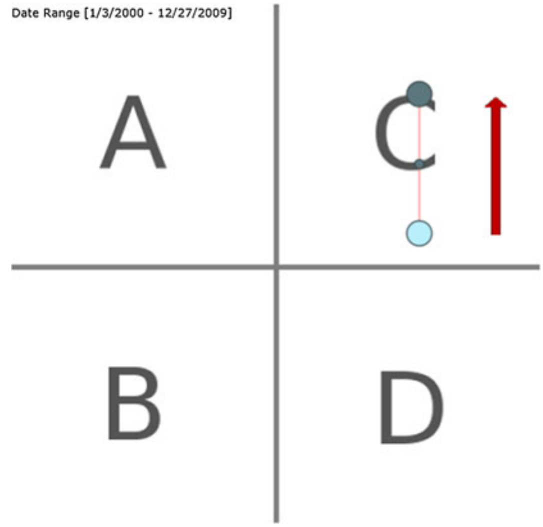

Fig. 15 Consistent rise of FPI on the radical scale

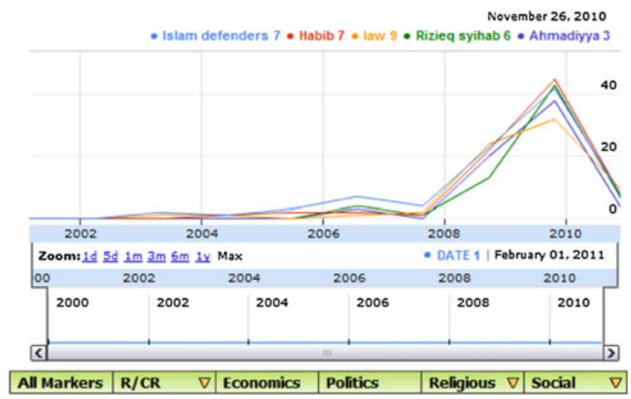

Fig. 16 "Ahmadiyya" peaking during the period 2006–2010

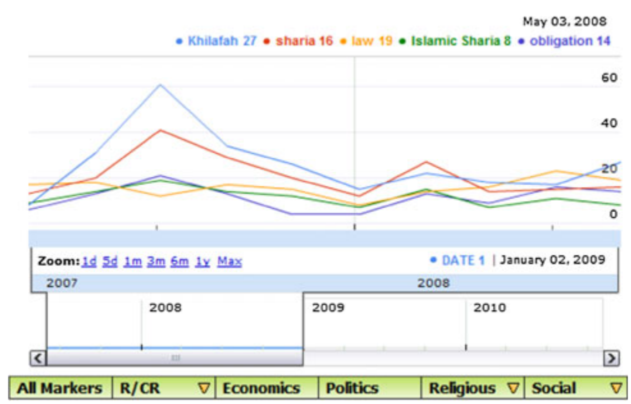

Fig. 17 "Khilafah" ideology of Hizb ut-Tahrir

the ideology of Khilafah, which believes in unification of all Muslim countries as a single Islamic State (Zakaria 2011; Mohamed Osman 2010). Figure 17 shows

[9] Select "Radical" and "FPI" from the filters, then select the time range 2002–2006 or 2006–2010, then select "radical" markers under "R/CR" menu.

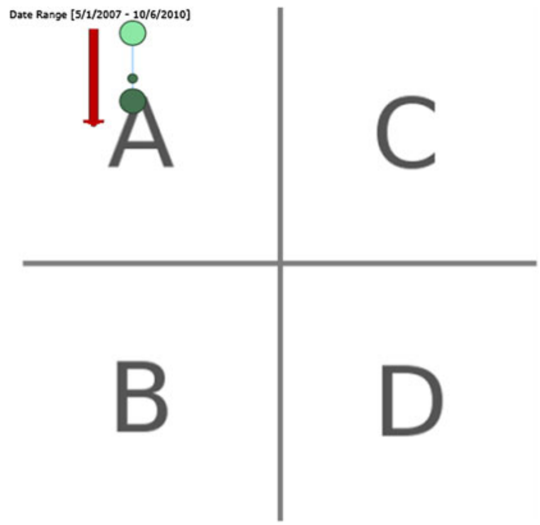

Fig. 18 Decline of the HTI in the radical scale

"Khilafah" as the most prominent marker[10] in Hizb ut-Tahrir's discourse.

By looking at the Quadrants widget (in Fig. 18), we can infer that HTI has been moderating its narrative.

5.4 Scenario 4: B-quadrant organizations' trends

In this scenario, we discuss the trends of counter radical organizations like NU and DaarulUluum. We also show an interesting scenario on the topic of "Suicide Bombing" using the keyword based Navigation widget.

The "counter radical" markers[11] associated with these organizations are: "politics", "election", "Indonesian Islam", "liberal", "human rights". These organizations support democracy and elections, which is shown by the high frequency of the markers "politics" and "election". Their narrative has local interpretation of Islam at its core, which is shown by the marker "Indonesian Islam".

On analyzing the occurrences of radical markers[12] in B-Quadrant, we find that counter radical organizations are very vocal against all of radical markers. One of the interesting radical markers is "Suicide Bombing". Most of the counter radical organizations are against suicide bombings.(Malang 2006). We will now demonstrate how combination of parametric and keyword search, and various widgets in the web application can help reveal opposition to "Suicide Bombing" by counter-radical organizations.

Searching for the text "suicide bombing", we see that one of the related markers is "ideology". Adding the keyword "ideology" to the search filter reveals a new set of markers including the "sin" keyword. Adding "sin" to our search, we obtain a set of matching documents. One of the top matches, is titled "Mengapa Saya Berubah?" (english translation: "Why I changed?")[13]. This article is by a reformed terrorist, debunking the misinterpretation of the jihad-related verses used by violent groups.

6 Conclusions and future work

In our experiments, not only did the data show fitness with the Rasch Model for the R/CR scale but also the Rasch rankings of the organizations are better than the output of the other baseline computational methods, and they are at expert-level performance when compared with the consensus gold standard rankings.

Rasch model also provided us with another output, namely the ranking of selected keywords (items) on the R/CR scale. Although preliminary observations indicates that this can be a valuable asset by itself, we plan to further investigate the quality and utility of this ranking as future work.

While the model has been demonstrated to fit on the R/CR scale, two major expansion points can be investigated in the future work, namely the violent/non-violent scale, and enhancement of feature selection. Although our experts have identified a second dimension, evaluating its correlation to R/CR axis, or existence of other significant ones could be beneficial. In addition, the features can be enhanced by experimenting with the significance of the radical keywords in the counter-radical organization corpora, and vice-versa.

A practical method to increase the automation of keyword generation has been discussed in Sect. 3.4. Future work will involve finding an efficient approximation algorithm for this model, for decreasing the necessity of expert interaction for this particular step.

Other interesting work includes making our expert opinion elicitation tool available online to a wider and more geographically distributed audience to crowdsource (Snow et al. 2008) the needed expertise for making lists of local organizations, identifying their web sources, and overcome the complex task of construction and validation of significant and fitting scales (work is currently underway to build this tool). Another interesting dimension is to look at synthesis and analysis of scales that do have a strict hierarchy of keywords, but adhere to more flexible partial order models (James and John 2002).

[10] Select "Hizb ut-Tahrir" and "radical" from filters. Select the time range 2007–2009. The markers can be seen by selecting the options of Markers Menu [Religious → Religious Markers].

[11] Select CounterRadical filter in the search option, then from the Markers Menu select [R/CR → Counter Radical].

[12] In the Markers Menu select [R/CR → Radical].

[13] http://www.islamlib.com/id/artikel/mengapa-saya-berubah/.

Acknowledgments This research was supported by US DoDs Minerva Research Initiative Grant N00014-09-1-0815, Project leader: Prof. Mark Woodward, Arizona State University, and the project title is "Finding Allies for the War of Words: Mapping the Diffusion and Influence of Counter-Radical Muslim Discourse".

References

Andrich D (1988) Rasch models for measurement. Sage, USA

Bayat A (2007) Making Islam Democratic: social movements and the post-Islamist turn. Stanford University Press, USA

Crelinsten R (2002) Analysing terrorism and counter-terrorism: a communication model. Terror Political Violence 14:77–122

Davulcu H, Ahmed ST, Gokalp S, Temkit MH, Taylor T, Woodward M, Amin A (2010) Analyzing sentiment markers describing radical and counter-radical elements in online news. In: Proceedings of the 2010 IEEE second international conference on social computing, IEEE Computer Society, SOCIALCOM'10, pp 335–340

Drehmer D, Belohlav J, Coye R (2000) An exploration of employee participation using a scaling approach. Group Org Manage 25(4):397

Durkheim E (2004) The cultural logic of collective representations: social theory the multicultural and classic readings. Wesleyan University: Westview Press

Frost F, Rann A, Chin A (2010) Terrorism in southeast asia. http://www.aph.gov.au/library/intguide/FAD/sea.html [Online accessed 21 Nov 2011]

Guttman L (1950) The basis for scalogram analysis. Meas Predict 4:60–90

Hasan N (2009) Islamic militancy, Sharia, and democratic consolidation in post-Suharto Indonesia. RSIS Working Papers. 143/07

Hessen D (2010) Likelihood ratio tests for special Rasch models. J Edu Behav Stat 35(6):611

Hunter D, Lange K (2004) A tutorial on mm algorithms. Am Stat 58(1):30–37

James AW, John LM (2002) Algebraic representations of beliefs and attitudes: partial order models for item responses. Sociol Methodol 29:113–146

Jolliffe I (2002) Principal component analysis: Springer series in statistics. Springer, Germany

Le Cam L (1990) Maximum likelihood an introduction. ISI Rev 58(2):153–171

Likert R (1932) A technique for the measurement of attitudes. Arch Psychol 140:1–55

Malang (2006) NU chairman deplores suicide bombing attempt. http://www.nu.or.id/page/en/dinamic_detil/15/28282/News/NU_chairman_deplores_suicide_bombing_attempt.html [Online accessed 22 Nov 2011]

McIver J, Carmines E (1981) Unidimensional scaling, vol 24. Sage Publications Inc, USA

McPhee RD, Corman S (1995) An activity-based theory of communication networks in organizations, applied to the case of a local church. Commun Monogr 62:1–20

Michael WB, Kogan J (2010) Text mining: applications and theory. Wiley, London

Mohamed Osman MN (2010) Reviving the Caliphate in the Nusantara: Hizbut Tahrir Indonesia's mobilization strategy and its impact in Indonesia. Terror Political Violence 22(4):601–622. doi:10.1080/09546553.2010.496317. http://www.tandfonline.com/doi/abs/10.1080/09546553.2010.496317

Osman MNM (2011) Preparing for the caliphate. Asian Stud Assoc Aust E-Bull 80:14–16. ISSN:1449-4418

Pawitan Y (2001) In all likelihood: statistical modelling and inference using likelihood. Oxford University Press, USA

Rahmat Sihaloho M (2011) FPI vows to disband ahmadiyah 'whatever it takes'. http://www.thejakartaglobe.com/home/fpi-vows-to-disband-ahmadiyah-whatever-it-takes/423477 [Online accessed 21 Nov 2011]

Rasch G (1961) On general laws and the meaning of measurement in psychology. In: Proceedings of the fourth Berkeley symposium on mathematical statistics and psychology, 4, p 332

Rondonuwu O, Creagh S (2010) Opposition grows to indonesia's hardline fpi islamists. http://www.in.reuters.com/article/2010/06/30/idINIndia-49777620100630 [Online accessed 21 Nov 2011]

Salton G, Buckley C (1988) Term-weighting approaches in automatic text retrieval. In: Information Processing and Management, vol 25, pp 513–523

Simmel G (2008) Sociological theory. McGraw-Hill, New York

Snow R, O'Connor B, Jurafsky D, Ng AY (2008) Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In: Proceedings of the conference on empirical methods in natural language processing, EMNLP '08, pp 254–263. Association for Computational Linguistics, Stroudsburg, PA, USA. http://www.portal.acm.org/citation.cfm?id=1613715.1613751

for the Study of Terrorism NC, to Terrorism R (2011) Terrorist organization profile: front for defenders of Islam. http://www.start.umd.edu/start/data_collections/tops/terrorist_organization_profile.asp?id=4026 [Online accessed 21 Nov 2011]

Thurstone LL (1928) Attitudes can be measured. Am J Sociol 33:529–554

Tikves S, Banerjee S, Temkit H, Gokalp S, Davulcu H, Sen A, Corman S, Woodward M, Rochmaniyah I, Amin A (2011) A system for ranking organizations using social scale analysis. In: EISIC. IEEE, pp 308–313. http://www.ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6059524

Tilly C (2004) Social Movements. Paradigm Publishers, USA

Tunkelang D (2009) Faceted Search: synthesis lectures on information concepts, retrieval, and services. Morgan & Claypool Publishers, UK. doi:10.2200/S00190ED1V01Y200904ICR005

van Emde Boas P (1981) Another *NP*-complete partition problem and the complexity of computing short vectors in a lattice: Tech. Rep. 81-04. Mathematisch Instituut, Amsterdam

Wallace A (1956) Revitalization movements. Am Anthropol 58:264–281

Ward K (2009) Non-violent extremists? Hizbut Tahrir Indonesia. Aust J Intl Affairs 63(2):149–164 doi:10.1080/10357710902895103.http://www.tandfonline.com/doi/abs/10.1080/10357710902895103

Widhiarto H (2010) Radical groups urge Bekasi administration to implement Sharia law. http://www.thejakartapost.com/news/2010/06/27/radical-groups-urge-bekasi-administration-implement-sharia-law.html [Online accessed 21 Nov 2011]

Zakaria Y (2011) A Global Caliphate: reality or fantasy? http://www.usa.mediamonitors.net/content/view/full/91207 [Online accessed 21 Nov 2011]