

CALICO

Developed for:
USACE/CERL
Developed by:
T Hill &
Associates, Inc.



Approved for public release; distribution is unlimited.

This publication was produced for review by the US Army Corps of Engineers.

This evaluation was conducted under:

USACE Contract Number: W9132T11P00476

Table of Contents

I. Overview.....	8
A. History.....	8
B. Tasks.....	8
II. Corpus.....	9
A. Approach.....	9
B. Corpus Content	10
C. Lesson's Learned	11
III. Preprocess Efforts, Coding	12
A. Process	12
B. Descriptions of Taxonomies.....	14
1. Taxonomy of Relief Organizations in Subjects.....	15
2. Taxonomy of action verbs in Predicates	16
3. Taxonomy of Objects in Predicates.....	18
4. Taxonomy of Sociocultural Variables.....	20
IV. Text Analysis Approach	24
V. Conclusion from Analysis.....	39
A. Findings – Direct Results of Queries	39
B. Findings – Aggregate Findings and Generalizations	49
1. Who Acts?	50
2. What Do They Do?	50
3. On Behalf of Whom?.....	52
4. What Are the Objectives of the Relief Organization?.....	53
5. Sociocultural Variables in HA/DR.....	56
C. Operational Analysis	59
VI. Final Recommendations	62
A. What the Corpus Did Not Tell Us	62
B. What the Corpus Did Tell Us.....	62
C. Challenges	63
D. Hypotheses and Future Directions for Research	63
E. Operational Recommendations	64

VII. Appendix I: Evaluation of MAXQDA Qualitative Analysis Software for Coding and Taxonomy Generation	68
A. MAXQDA 10 Training	68
B. Coding Process	69
C. Analyzing Data Using MAXQDA 10.....	70
D. Comparison with Hand-Coding.....	73
E. Advantages & Disadvantages of MAXQDA 10	74
1. Advantages.....	74
2. Disadvantages	74
F. References	75
VIII. Appendix II: Taxonomy of Relief Organizations in Subjects	76
IX. Appendix III: Taxonomy of Predicate Verbs	81
X. Appendix IV: Taxonomy of Objects in Predicates.....	94
XI. Appendix V: SMA Sociocultural Taxonomy	109
A. Interrelatedness	110
B. The 10 High Level SMA Typology Sociocultural System Variables.....	111
XII. Appendix VI: Taxonomy of Socio-cultural Variables.....	114
XIII. Appendix VII: CALICO Reference Documents	121
XIV. Appendix VIII: Bibliography of Corpus Documents	122
XV. Team Contact Information	128

Table of Figures

Figure 1. SMA Sociocultural Typology.....	21
Figure 2. Architecture of the Processing Pipeline.....	24
Figure 3. User Interface for the Document Classifier.	26
Figure 4. Taxonomy Development Triad.....	27
Figure 5. Global Perspective of a Classification Run.	28
Figure 6. User Interface to Assess the Performance of a Classification Run.	29
Figure 7. Improved Taxonomy.	29
Figure 8. Sorted Unlatched Phrases.....	30
Figure 9. Internal Semio Tagger database Diagram.....	31
Figure 10. iBase Schema.	32
Figure 11. Subset of Categories and Phrases.....	33
Figure 12. Example of a Complex Query within the IBase User Interface.....	34
Figure 13. Query Graphical View.	35
Figure 14. Web User Interface for Viewing Document Classifier Results.....	35
Figure 15. Advanced Search Interface for Viewing Document Classifier Results.....	36
Figure 16. Web User Interface for Viewing Document Profile.....	37
Figure 17. Word Frequency Analysis Partial Results.....	40
Figure 18. Venn Diagram of Query Intersection Construct.....	41
Figure 19. Query Builder User Interface.	42
Figure 20. Query Results in Report Layout.	43
Figure 21. Venn Diagram of Query Intersection Construct.....	44
Figure 22. Query Builder User Interface.	45
Figure 23. Query Results in Report Layout.	46
Figure 24. Query Builder User Interface.	47

Figure 25. Query Results in Report Layout.	48
Figure 26. Identified Relief Organizations.....	50
Figure 27. Lists the First and second tier categories of relief organization activities in order of the number of documents in which their concepts were mentioned.	52
Figure 28. Beneficiary table.	53
Figure 29. Category of Concerns.....	54
Figure 30. Category of Object Actions.	55
Figure 31. Category of Material Things.....	56
Figure 32. Sociocultural variables – Economy, Technology and Other Capabilities.	58
Figure 33. Siocultural Variables - Demography.....	58
Figure 34. Sociocultural Variables – Interests.....	59
Figure 35. Recommended data types for future CALICO research to understand HA/DR indicators relevant to HA/DR responses (Source: World databank.com).	65
Figure 36. Multilayered Geospatial Analysis for visualizing research findings (source: worldmap.harvard.edu/africamap/).	66
Figure 37. Example of Geospatial Mapping of food emergencies in Africa (source isciences.com).	66
Figure 38. Example of Geospatial Mapping of Disaster Trends (source: ochaonline.un.org/MapCentre/).	67
Figure 39. MAXQDA Training Time Requirements.....	69
Figure 40. MAXQDA User Time Requirements.	69
Figure 41. Example of MAXQDA 10 Retrieved Segments pane	70
Figure 42. Example of MAXQDA 10 Retrieved Segments spreadsheet	71
Figure 43. Example of a Document Portrait.....	71
Figure 44. Example of a Codeline Visualization (taken from the MAXQDA 10 Reference Guide).....	72
Figure 45. Example of a Tag Cloud and Word Frequency List.....	72
Figure 46. User Time Requirements.	73
Figure 47. Example of Codes Applied in a Document	75

Figure 48. Single Code Highlighted in Text	75
Figure 49. Typology Hierarchy.	110
Figure 50. Basic Interrelatedness of High Level SMA Sociocultural Variables.	111

I. Overview

A. History

A team from the Engineering Research Development Center-Construction Engineering Research Laboratory (ERDC-CERL) developed a content analysis project for *Civil Affairs Language for Informing Cultural Operations* (CALICO). The major facets of the CALICO effort are aimed at increasing the understanding of indicators of cultural relevance within running text and of the connection between that relevant sociocultural information and the specific tasks described in the running text.

The CALICO contract was given to the T Hill & Associates (THA) team that includes Subject Matter Expert (SME) personnel from Millennium KI (MKI) and National Security Innovations (NSI). The team of SMEs brought expertise in OCOUNS operations of Civil Affairs, US Government Mission, Humanitarian Assistance and Disaster Relief programs, as well as part-of-speech labeling, word frequency analysis, text-coding and taxonomy development.

This Contractor team was to answer the question: "What do government and NGO reports on HA/DR activities tell us about the information needed to support effective interventions to ameliorate the human consequences of disruption and to enhance the effectiveness of indigenous governance?"

Team members first visited the ERDC-CERL office to achieve clarity on the objectives and methods expected for the project. This visit gave clarity to the team and allowed for the start of the corpus creation and execution of all assigned tasks.

B. Tasks

In order to accomplish the work under this project, it was necessary for the Contractor to complete the following tasks:

Task 1: Develop a corpus of a minimum of fifty (50) reports on actions taken by organizations engaged in HA/DR missions.

Task 2: Deliver all texts in the corpus to ERDC-CERL in standard plain-text electronic format (e.g., xml); and cleaned such that no duplications or other anomalies introduced by any conversion process are found.

Task 3: Conduct automated analysis of corpus.

Task 4: Prepare documented deliverables.

II. Corpus

A. Approach

The corpus was the set of reports used for analysis in the project. This document set was a very important building step in the process, as the study would not deliver the appropriate results without the correct set of reports. The goals in selection of the corpus were to:

- Concentrate on non-governmental organizations (NGO) efforts on the African continent
- Capture the range of NGO efforts conducted in Africa
- Cover the range of NGOs that operate on the African Continent

The approach to building the corpus was to gather a comprehensive collection of current and historical information related to a particular region, group, problem, or issue, which includes human, social, cultural, economic, religious, or other relevant information, as well as operationally-relevant issues, concerns, metrics, and/or research themes. The corpus is intended to provide a common framework and data for use by ERDC-CERL to design research and to develop solutions to address problems that are of real-world interest to the United States military.

The objective of the corpus development was to gather a minimum of fifty (50) reports on actions taken by organizations engaged in humanitarian assistance/disaster response (HA/DR) missions. Reports chosen for the corpus only consisted of publications in English from the year 2000 to the present and focused on efforts conducted outside the Continental United States in Africa.

The sources of these reports included, but were not limited to, the following organizations:

- International Red Cross and Red Crescent.
- United Nations.
- United Nations/World Food Program.
- U.S. Agency for International development.
- African Union.
- Doctors Without Borders.
- Military response to HA/DR.
- Host Nation response to HA/DR.

An explicit effort was made to avoid gathering reports on U.S. military aid activities, since this is already being researched by ERDC-CERL. When searching U.S. military aid activities the main focus was to gather information on non-governmental organizations (NGOs), international non-governmental organizations (INGOs), and intergovernmental organizations, also known as international governmental organizations (IGOs) and their roles in the HA/DR efforts.

The contractor team focused on keyword searches using Boolean logic on topics and concepts to begin the assembly of the corpus. The searches initially focused on the primary organizations listed above associated with “humanitarian assistance” or “disaster response.” Upon drilling down through the

numerous results, some disaster response operations began to emerge together with key themes such as drought and famine. These results would then lead to organizations that provide support to these types of issues.

Google was the primary search engine used to query the Internet. Other search engines such as Bing, Dogpile, and Yahoo were used to a lesser extent. Other sources of reference were the Directory of Development Organizations and the University of California Berkeley Library.

The Directory of Development Organizations provided a list of NGOs operating in Africa by country. At the time of this corpus build there are 54 independent countries and five regions in (or on the continent) of Africa. Each country has some level of humanitarian assistance or in some areas is being provided disaster response support. Due to the vastness of the continent and in order to narrow the search parameters only a few select countries from the five different regions were looked at when searching for HA/DR material.

The Berkeley Library provided an NGO custom search engine, which was connected to other sites such as the NGOs with Consultative Status with the United Nations Economic and Social Council, the University of Minnesota Human Rights Library, the Duke University Libraries' NGO Research Guide, as well as other sites. The Berkeley Library site also provided various links to International and Regional Guides & Directories along with NGOs and the United Nations.

B. Corpus Content

As per the contract, Task 1 (3), reports chosen for the corpus were selected from open-source literature that is unclassified and consisted of an average of forty (40) pages of running text, excluding summaries, tables of contents, appendices, indices, and lists of references. The entire corpus was delivered to ERDC-CERL in standard plain-text electronic format (e.g., PDF); and cleaned such that no duplications or other anomalies introduced by any conversion process. The documents in the corpus were authored by a variety of relief organizations, including NGOs, INGOs, IGOs, and other local and governmental organizations. For that reason, the more generic term, relief organization will be used to refer to the organizations studied in this effort, with more specific designations used when appropriate.

Below is a break down of the delivered corpus:

- 71 documents pertained to relief organizations and relief efforts.
- The CERL staff and Larry Kuznar reviewed the documents, and 7 documents were eliminated due to insufficient authorship, text, or military subject matter.
- The resulting corpus:
 - 64 documents
 - 3335 pages
 - Average pages per document 52.1
 - Median pages per document 25

C. Lesson's Learned

On the surface most organizations do not go into detail about their HA/DR operations. They list a lot of activities they have conducted in various places and occasionally outline the results of their efforts. Most of the first results of the search query yielded one or two page articles. While the information contained in the returned articles may have been relevant to the task initially, they did not meet the criteria to be included in the corpus. It took a lot of digging into topics and themes to find acceptable material for the corpus.

Many relief organizations operating throughout Africa are French. Some of the search results produced what appeared to in-depth documents, however they were in French and were unable to be used based on the corpus criteria.

A recommendation is to find an organization's annual report early in your research. Some organizations provide some fairly detailed reports of their activities throughout the year. This led to additional searches, search criteria, and supplementary activities and approaches of relief organizations to providing support to various areas.

Map out your search criteria first. As a team it is imperative to take a "workshop" or "word cloud" approach to developing your search criteria. Everyone makes a list of keywords they feel are important then compare notes. Eliminate the duplicates then see combinations can be used as search phrases. This saves a lot of time and brainpower of the researchers in coming up with relevant search criteria.

III. Preprocess Efforts, Coding

A. Process

An overarching requirement of this project is to allow the corpus to reveal the nature of relief organization activities in Africa, with a focus on what relief organizations do and on behalf of whom. Therefore, initial taxonomies were developed from a sample of the corpus itself, rather than beginning with taxonomies based on theoretical, or worse idiosyncratic, schemas of relief organization activities. Developing an initial taxonomy from the corpus required some hand-coding of documents in order to generate these initial taxonomies.

Not only was there a desire to allow the corpus to define the taxonomies of relief organizations, activities and beneficiaries, but that these taxonomies provide a fair representation of the range of relief organization activities in Africa. Therefore, an initial classification of documents was developed based on type of organization, authoring organization, and type of relief organization activity.

The initial collection of documents for the corpus was sufficiently broad to provide a wide range of organizations that authored the documents. These authoring organizations could be grouped into 5 basic categories:

- NGOs (43)
- US Government (9)
- Academic organizations (7)
- Foreign Government (3)
- Think Tanks (2)

NGOs were actual non-governmental entities, which included global organizations such as the United Nations and Red Cross/Red Crescent, as well as local civil society organizations (CSOs). The majority of documents (43 of 64) were authored by NGOs, the primary focus of this study. However, other types of organizations were also present and were important to the functioning of relief efforts. U.S. government, non-military, was represented by USAID documents (9 documents). Academic journals provided another 7 documents that contained materials relevant to this study. Foreign governments played a key role in relief efforts, and foreign governments primarily authored 3 of the documents in the corpus. Finally, private firms that provided analysis of relief organization activity authored 2 of the documents.

The documents in the corpus pertained to 7 types of relief activity, which include:

- Development (14)
- Security (15)
- Natural Disaster (6)
- Famine (8)
- Health Care (3)
- Education (1)
- Multiple (16)

It is important to note that no document contained information solely on any of these types of activity, reflecting the complexity of relief efforts. A document was assigned to an activity based on the primary focus of the activities described in the document.

Development referred to economic development projects that supported business and agrarian projects that in turn supported agriculture and herding. Fourteen documents were primarily focused on development activities. Documents that focused on security primarily dealt with relief efforts associated with conflict and policing efforts. The fact that 15 documents in the corpus were focused on security reflects the widespread social instability that plagues the African continent, as well as the dangers relief workers face when operating in Africa. Natural disaster and famine relief are closely related. Natural disaster refers to environmentally-caused relief efforts. These range from destructive events such as the 2004 Indian Ocean tsunami that killed approximately 230 million people around the Indian Ocean to climatic fluctuations such as drought, which undermine the agrarian economy of African societies. Of course, droughts lead directly to famine relief efforts. There were 6 documents that focused on natural disasters and 8 that focused on famine relief. Three documents focused on health care, and one on education. As noted above, no relief activity carried out any single function, but instead involved the integration of several types of relief activity. In many cases, the relief efforts described in the documents were so varied that a single focus was neither apparent, nor appropriate. In these cases, the relief organization activity was labeled multiple, and 16 of the documents in the corpus represented this kind of varied and integrated activity.

The technological goal of this study is to explore how automated methods can extract relevant information on relief organization activity, allowing the analysis of a greater number of documents and a reduction of human labor involved in the analysis. Therefore, it was desirable that a small sample of documents and pages be hand-coded to create the initial taxonomies. We decided that no more than 10% of the documents should be coded, and hopefully a lesser percentage of pages. In order to select this small sample in a representative manner, the corpus was stratified by relief organization activity, and proportionate random sample of documents was selected. The final selection of documents included:

- Humanitarian Requirements 2011, Joint Government and Humanitarian Partners Document, Addis Ababa, Ethiopia, 2011
- System of Rice Intensification, Best Practices & Innovations (BPI) Initiative Agriculture % Rural Livelihoods, AFRICARE, 2011
- Smart Development in Practice: Field report from southern Sudan, Oxfam, 2008
- Mobilizing Early response Project – Kenya, Nairobi Peace Initiative – Africa/GPPAC Eastern and Central Africa Region, 2009
- Facilitating the Observance and Protection of Child Rights in Uganda, NGO Complementary Report to the GOU First Period Report to the CRC, 2000
- East Africa: Earthquake & Tsunamis Quarterly Report, Red Cross/Red Crescent, 2006

Of these 6 documents, 2 represent development activities, 1 multiple activities, 1 famine relief, 1 natural disaster relief, and 1 security. These 6 documents represent 9.4% of the documents, but only 158 pages, or 4.7% of the pages in the corpus.

These documents were coded by hand, using different colored markers to mark and annotate the texts. Qualitative analysis software could potentially facilitate this process, and speed up the initial coding and development of taxonomies process. We conducted an experimental coding with the software MAXQDA to evaluate the costs and benefits of using such software for the development of taxonomies. The results of this experiment are presented in Appendix I.

In order to address the objective of this study, the identification of what actors perform which activities on behalf of whom, coding focused on key sentences in which a relief organization was identified as an actor. These sentences were parsed into subjects and predicates, and the predicates were further parsed into their key action verbs and the objects of those predicates. This parsing provided the basis for three separate taxonomies that, hopefully, would capture the key actors, actions and beneficiaries of relief organization activity in Africa.

Another objective of this study was to capture the key socio-cultural variables relevant to relief organization activity. As the coding proceeded, the anthropologist on the team, Dr. Lawrence Kuznar, thought that too little socio-cultural information was represented in the key sentences, and so he decided to code socio-cultural variables wherever they occurred in the texts, and to create a separate socio-cultural taxonomy. To do so, Dr. Kuznar leveraged a previously developed socio-cultural taxonomy that was developed for Strategic Multilayer Analysis (SMA) efforts in which military analysts needed to analyze socio-cultural variables pertinent to adversaries, allies and general populations.¹ The socio-cultural variables Dr. Kuznar identified were categorized according to the taxa in the SMA socio-cultural taxonomy.

B. Descriptions of Taxonomies

The following four taxonomies were created by the procedures described above:

- Taxonomy of Relief Organizations in Subjects
- Taxonomy of action verbs in Predicates
- Taxonomy of Objects in Predicates
- Taxonomy of Sociocultural Variables

These taxonomies were categorized into main categories and, where appropriate, sub-categories. The taxonomies were augmented by including acronyms, spellings of acronyms, plurals and singulars of nouns and all major English verb tenses (present first, second and third person, past, participles). The taxonomies are described in turn. The complete taxonomies are each found in separate appendices (Appendices II, III, IV, VI).

¹ 2011 Intent White Paper Anthropological Reflections on Motive vs. Intent & the SMA Typology. In *From the Mind to the Feet: Assessing the Perception to Intent Dynamic*. Edited by Allison Astorino-Courtois, Larry Kuznar, and Sarah Canna, pp. 61-72. Air University Press, Maxwell AFB, AL.

1. Taxonomy of Relief Organizations in Subjects

This taxonomy is shallow, in that it contains only one taxonomic level with 5 categories. Organizations were often not clearly relief organization or government, so organizations were classified according to their predominant influence on their operations. The categories and the number of instances within each category are listed below.

- NGOs (95 instances)
- United Nations (5 instances)
- Foreign Governments (83 instances)
- US Government, Non-military (5 instances)
- Personnel (2 instances)

Nearly all organizations were either NGO or Foreign Government, reflecting the close cooperation of NGO and Foreign governments. Another general pattern was that personnel were rarely mentioned – actors mentioned were almost entirely at the organizational level, reflecting the organizational and administrative focus of the reports in the corpus. Organizations included abstract references (“the project”), specific organizations (Somali Red Crescent), or specific projects (mobilizing early response project).

NGOs included organizations that were not clearly part of governmental structures. The NGO category contained the largest list of terms (95), reflecting the nature of the documents, and the aim of this study. Some NGOs were global in scope, such as the Red Cross/Red Crescent. Organizations affiliated with the United Nations were placed in a separate category because of their global and political nature. Other NGOs were more regionally based within the African continents, such as regional Red Cross/Red Crescent branches, IGAD (Inter-Governmental Authority on Development), and the East Africa regional Delegation. Other organizations represented coalitions, such as CRNN (Child Rights NGO Network). Yet other organizations represent local and within-country NGOs, such as Uganda Society for Disabled Children and SUMI (Sudan Microfinance Institution). Some organizations referred more to an on-going project, such as AFRICARE’s System of Rice Intensification (SRI), the Targeted Supplementary Feeding (TSF) Programme, or the Slum Aid Project. Civil Society Organizations, which can mean a vast array of localized relief and aid organizations, were often simply referred to as CSOs. Other organizations refer to institutions found in any society such as hospitals. Finally, organizations that represented types of organizations rather than specific organizations, such as “donor communities,” the committee,” “partners,” and “NGOs” were often mentioned.

Since the United Nations represents simultaneously a global and political body, we thought it would be useful to distinguish UN-based organizations. Only 5 instances were mentioned, including the World Food Program (WFP), and UNICEF.

Many host nation government organizations were mentioned in the coded documents (83), almost as many as NGOs. The high frequency of host nation governmental organizations reflects the mandate that many NGOs have, which is to assist governments in aiding their citizens. The high number of instances also indicates the importance of working with host nations in relief efforts. Terms ranged from the

abstract, such as the “state,” “relevant agencies,” and the “police,” to very specific ministries and programs. Often, these ministries and programs were referred to by acronyms. Examples include Ministry of Health, Ministry of Rural Development and Water Resources (MRDWR), Office of the Prime Minister, and the Child Labour Unit. Governmental personnel were also mentioned, including, Mobile Health and Nutrition Teams (MHNTs), and Health Extension Workers (HEW).

The United States maintains a strong presence in aid worldwide, and so non-military U.S. organizations were placed in their own category. Not many terms were used to refer to U.S. organizations (5), and nearly all were to USAID, although occasionally, the United States was mentioned more abstractly.

The final category was for relief organization Personnel. Interestingly, this was the smallest category, with only 2 reference terms, Aid Practitioner and its plural. This, in part reflects the small sample of txt that was coded from the corpus, certainly there are more personnel terms, however, it also powerfully reflects the fact that the reports that we examined are written for administrative and reporting reasons, and probably have audiences affiliated with organizations at a similar social level to the organization; their audience appears to be other organizations such as governmental agencies, other relief organizations, and themselves. The actors who actually execute field operations are largely lost in the language of the text in this small, hand-coded sample.

2. Taxonomy of action verbs in Predicates

The action verbs found in the predicates provided a richer taxonomy of terms that can be broken down into different types of relief organization activity (see below). There were 513 action verbs in all. The basic types of activity include Administrative Action (116 terms), Administrative Execution of Operations (87 terms), Supporting Actions (115), Field Actions (94 terms), and a variety of other, lesser-mentioned activities, including financial, political, education and research functions.

Taxonomy of Verbs (513 instances)

Administrative Action (116 instances)

- Initiation and Leadership (41 instances)
- Oversight (8 instances)
- Assessment (11 instances)
- Reporting (20 instances)
- Meetings (19 instances)
- Travel (13 instances)
- Personnel (4 instances)

Execution of Operations (87 instances)

Supporting Action (115 instances)

- Augmenting (29 instances)
- Assistance (62 instances)
- Advisory (16 instances)
- Replacement (8 instances)

Field Action (94 instances)

- Distribution (22 instances)
- Reconstruction (12 instances)
- Mitigation (16 instances)
- Security Rescue (8 instances)
- Preparation (16 instances)
- Operations Beginning (12 instances)
- Operations Ending (8 instances)

Financial Action (4 instances)

Political Action (25 instances)

Success (8 instances)

Observation (20 instances)

Research (24 instances)

Forecasting (8 instances)

Education (8 instances)

Translation (4 instances)

Administrative actions include administrative activities other than those involved in directly running operations. Administrative actions had the largest number of terms associated with it (116 out of 513 terms), and involved terms for initiating and leading programs, overseeing programs, assessing programs, filling reports, holding meetings, travel and dealing with personnel issues. Furthermore, these activities involved the internal functions of the relief organization; in other words, these are things the relief organization did internal to its own operation, and not necessarily in cooperation or with the participation of the organizations the organization was leading, assessing, etc.

The second largest category of relief organization actions involved activities in support of other organizations (115 of 513 terms). These other organizations were most often smaller relief organizations at the community level and foreign governments (see Taxonomy of Predicate Objects below). The action verbs most associated with supporting actions included augmentation of existing programs, assistance, advising, and replacing programs.

Another large category of relief organization actions involved the direct execution of the organization's operations (87 terms). These verbs involve proactive activities such as carrying out, managing, conducting, controlling, addressing, developing, organizing and running.

Field Action constituted the other major category of relief organization actions, represented by 94 terms. These verbs describe activities carried out in the field, more or less directly helping beneficiaries. The terms include verbs for distributing and disseminating goods and services, mitigating problems, providing security and rescue, and reconstructing damaged infrastructure.

A variety of other, much less mentioned activities were also at this high taxonomic level. These activities include financial actions, political actions such as advocacy, being successful, and educating. There were several categories related to research activities, including observation, research, and forecasting at the highest taxonomic level.

3. Taxonomy of Objects in Predicates

Part of the answer sought in this study is for whom the relief organization acts. This information should be contained in the noun clauses of the sentence predicates. The objects of relief organization action were recorded and categorized, providing a list of 492 terms. The 6 categories at the highest taxonomic level, including:

- Beneficiaries (59 instances)
- Organizations (64 instances)
- Things Done (90 instances)
- Concerns (145 instances)
- Things (54 instances)
- Places (50 instances)

Beneficiaries are those individuals, groups, populations and entities that are the relief organization intends to benefit. Beneficiaries were fairly evenly distributed over 10 sub-categories. These categories included women, children, students, farmers and herders, as well as a few professionals (journalists) and relief organization functionaries. Governments were also occasionally mentioned as beneficiaries. Local civic organizations such as communities and villages were mentioned as beneficiaries. Finally, some relief organization activities are specifically directed to the support of businesses.

More formal organizations were another category, and were referred to by 64 terms. Many of these terms were very general and non-specific, such as projects, teams, committees, and associations, and were not very revealing. Civil Society Organizations were also mentioned, but often in generally. Governmental organizations, such as specific projects, programs and ministries were mentioned frequently (33 of 64 terms). Other relief organizations and specific business sector programs were seldom mentioned.

A large number of terms (90 of 513) specified the things a relief organization does as objects of sentences. These actions reflect many of the action verbs identified in the taxonomy of predicate verbs. The strong majority of terms involve administrative functions. Supportive functions include coordinating and assisting. Relief organizations also provide oversight, planning, and some financial and political action. One of the larger categories involves the holding of meetings (11 terms). Relief organizations also provide educational and research services. Finally, more direct support is accomplished through provisioning and technical support.

The largest category of objects is comprised of the issues that concern relief organizations (145 of 513 terms). These are the *raison d'être* of the relief organization. Social concerns such as children, gender, violence and poverty were mentioned. Health issues such as famine, illness, sanitation and AIDs were the most frequently mentioned concerns. Terms exist for economic and development concerns such as farming and herding. A number of concerns for increasing government capability were mentioned, and several were mentioned for aid organizations. Interestingly, the most terms used to represent specific relief organization concerns (16) were the coordination of activities, reinforcing the relief organization's role of facilitator.

A fundamental part of any society and culture is its material culture – societies are made of things as well as people and relations. There were 54 terms for things that were mentioned as objects in sentences describing relief organization activity. The two most common categories of things mentioned as objects of relief organization activity were documents (14 terms each), representing relief organization administrative functions, reflecting one of the primary concerns that relief organizations have. Other material things mentioned include food, shelter, educational supplies, money, and agricultural goods.

Societies exist in geographic space as well, and 50 terms were coded to capture the geographical element of relief organization activity. These terms largely represent the countries of Africa and some regional designations such as region, province and district. A number of the countries were not mentioned in the small sample of documents hand-coded, but were inserted into the taxonomy under the presumption that they are likely to be mentioned in the larger corpus.

Beneficiaries (59 instances)

- Community Village (5 instances)
- Women (6 instances)
- Children (11 instances)
- Businesses (3 instances)
- General beneficiaries (10 instances)
- Students (3 instances)
- Professionals (2 instances)
- Government (3 instances)
- NGO Functionaries (9 instances)
- Farmers (7 instances)

Organizations (64 instances)

- General Organizations (21 instances)
- Civil Society Organizations (6 instances)
- Governmental Organizations (33 instances)
- NGOs (2 instances)
- Business Sector (2 instances)

Things Done (90 instances)

- Coordinating (7 instances)
- Assisting (5 instances)
- Oversight (5 instances)
- General Actions (14 instances)
- Education (4 instances)
- Provisioning (9 instances)

- Analysis (6 instances)
- Presentation (2 instances)
- Planning (3 instances)
- Prevention (2 instances)
- Political Action (2 instances)
- Meetings (11 instances)
- Medical Interventions (10 instances)
- Technical Support (5 instances)
- Financial (5 instances)
- Concerns (145 instances)
 - Children (6 instances)
 - Gender (2 instances)
 - Violence (11 instances)
 - Famine (9 instances)
 - Poverty (8 instances)
 - Aid Organization Capability (4 instances)
 - Government Capability (9 instances)
 - Coordination (16 instances)
 - Financial (4 instances)
 - Health (22 instances)
 - Farming (16 instances)
 - Herding (3 instances)
 - Early Warning Risk Reduction (18 instances)
 - Education (2 instances)
 - General Concerns (39 instances)
- Things (54 instances)
 - Documents (14 instances)
 - Medical Supplies (14 instances)
 - Food (6 instances)
 - Shelter (5 instances)
 - Educational Supplies (4 instances)
 - Money (3 instances)
 - Agricultural Things (2 instances)
 - Other Things (6 instances)
- Places (50 instances)

4. Taxonomy of Sociocultural Variables

A key goal of this study was to identify sociocultural variables involved in relief organization activity. In order to improve our ability to capture this information, a separate taxonomy of socio-cultural variables was developed. This taxonomy used a pre-existing DoD sociocultural taxonomy to provide a basic structure and placed items into that taxonomy as they occurred in the sample of documents hand-coded.

The DoD taxonomy used was the SMA Socio-cultural taxonomy, developed for the Strategic Multilayer Analysis office of the DDR&E under the Office of the Secretary of Defense. The taxonomy was originally developed to re-evaluate U.S. deterrence policy in the 21st century, in response to the variety of threats

emanating from various cultures and contexts. The original effort involved 120 scholars from U.S. universities, and government officials and operators engaged in deterrence. The months long effort aimed at capturing the collective wisdom of scholars from a variety of social science fields (political science, psychology, anthropology, criminology, sociology) and experts in deterrence to generate a comprehensive typology that could be used to characterize key variables essential to understanding it.

The typology is organized in a hierarchy that includes increasingly detailed levels. The levels are 5 Fundamental Categories (**Interests, Capabilities, Context, Decision Making Psychology, Language**), 10 High Level Variables: **Interests (Motivating Factors Religion and Ideology, Social Identity, Objectives)**, **Capabilities (Economy, Technology and Other Capabilities)**, **Context (Roles/Life Cycle, Demography, Political and Social Organization, Environmental and Historical Context and Other Actors)**, **Decision Making Psychology, Language**), and Subcategories added as necessary. A more complete description is found in Appendix V.



Figure 1. SMA Sociocultural Typology

The high-level category, Interests, captures ideological and motivating factors salient to peoples' lives. While this is a large category in the SMA taxonomy, only 7 terms pertinent to it were mentioned in the sample of hand-coded texts. Only one social identity variable, being Muslim, was mentioned in the coded texts. A few references to material interest and cultural norms were also mentioned. Considering the fact that most social scientists consider ideology and social identity to be key to understanding human behavior, the absence of terms related to these variables indicates that a major domain of human thought and behavior is ignored by relief organizations.

Economy, technology and other material capabilities of societies, however, receive much attention from relief organizations, if the number of terms (98) used to refer to these factors is an indicator. All of these factors relate to what many anthropologist consider a society's infrastructure, and to what most other social scientists attribute to economy. Some terms concern macroeconomic variables such as poverty

levels, prices and markets. Most terms deal with specific modes of subsistence such as agriculture (54 terms), animal husbandry or herding, wage labor, income, education and specific risks to these.

Contextual variables constitute a major high-level category in the SMA sociocultural taxonomy, and correspondingly are a major source of terms from the coded texts (106). There were 3 major sub-categories represented under the taxon of Context. They include environmental and historical context (including some political variables), demography, and political organization. The environmental and historical context terms (43 terms) identified in the sample of texts are related to terrain and climate and the overall physical security of people in a region from violence. Demographic variables are very important in a holistic perspective on society, but typically constitute a small minority of terms used by social scientists at large. However, our sample contained 25 terms associated with demography, including terms related to fertility, disease, sanitation, and mortality. This comparatively large proportion of contextual terms indicates the demographic focus of relief organizations. Political and Social organizational variables are essential to understanding human social life, and figure prominently in the language of the coded texts (35 terms). About half of these terms (18) dealt with governmental organizations that are part of nation states, which of course are the organizational entities through which relief organizations often must act. However, relief organizations did not ignore non-governmental forms of social organization, with 14 terms referring to social structures such as households, villages, tribes, councils of elders, religious leaders, chiefs, and communities.

Language has always been a difficult variable for social scientists; on the one hand nearly every social scientist recognizes its pervasiveness, but on the other hand, social scientists seldom systematically incorporate its effects in their analyses of society. This ignorance of language is reflected in the language used by relief organizations, with only 1 reference to this key variable; the use of the term dialect.

The final category of sociocultural variable is an addition to the SMA typology: Social Science Methods. The most comprehensive socio-cultural taxonomy used by anthropologists is the Outline of Cultural Materials (OCM) developed and maintained by the Human Relations Area Files (HRAF), affiliated with Yale University. They incorporate this variable into their description of every society in their database, since how we have come to know a society is as important as what we claim to know about it. And so it is appropriate to augment the SMA typology with Social Science Methods. Social Science Methods were seldom mentioned in relief organization reports, but one report drew out in some detail the theoretical, methodological and epistemological basis of their knowledge of society, acknowledging the use of methods such as focus groups, discourse analysis and process tracing. Given the low level, but persistent mention of terms related to research in relief organization reports, it seems wise to include terms related to methodology in any taxonomy of relief organization activity.

Sociocultural Taxonomy

(216 instances)

Interests (7 instances)

Social Identity (1 instance)

Religious Identity (1 instance)

Motivating Factors, Religion, Ideology (6 instances)

Materialism (1 instance)

Uncodified Norms (5 instances)

Economy, Technology and Other Capabilities (98 instances)

Infrastructure (98 instances)

Transportation (1 instance)

Macroeconomics (12 instances)

Domestic Economy (85 instances)

Economic Risks (4 instances)

Agriculture (54 instances)

Animal Husbandry (13 instances)

Finance (6 instances)

Income (3 instances)

Fishing (1 instance)

Labor (2 instances)

Transportation (1 instance)

Education (1 instance)

Context (106 instances)

Environmental and Historical Context and Other Actors (43 instances)

Geographic Factors (12 instances)

Settlements (1 instance)

Security Situation (30 instances)

Roles and Lifecycle (3 instances)

Demography (25 instances)

Fertility (3 instances)

Disease Morbidity (18 instances)

Mortality (2 instances)

Migration (2 instances)

Political and Social Organization (35 instances)

State (18 instances)

Nonstate (14 instances)

Social Fissures and Unrest (3 instances)

Language and Symbolic Communication (1 instance)

Social Science Methods (4 instances)

IV. Text Analysis Approach

This section covers the general technical approach used to perform the text-analytic processing and analysis. When appropriate, processing artifacts generated will be both discussed and referenced. Most of these artifacts are in human readable formats and in some cases are also in machine readable files for further analysis. The guiding purpose for this activity was to distill and describe the relevant sociocultural information associated with specific HA/DR activities within the corpus. The specific tasks performed are as follows:

- Part-of-speech labeling and word frequency analysis
- Develop a characterization system and classifier for HA/DR tasks and sociocultural information
- Label HA/DR tasks – words and phrases that contain sociocultural information
- Analysis of associations between HA/DR tasks, agents, and sociocultural information (section VI)

The following section contains a detailed discussion of the processing and the specific tools and techniques used to create the bulk of the text-analytic output. This discussion includes one small preprocessing step not previously addressed due to its technical nature. Before diving into that detailed technical discussion, it is useful to get a clear picture of the overall process steps. The **figure 2** below describes the architecture of the processing pipeline.

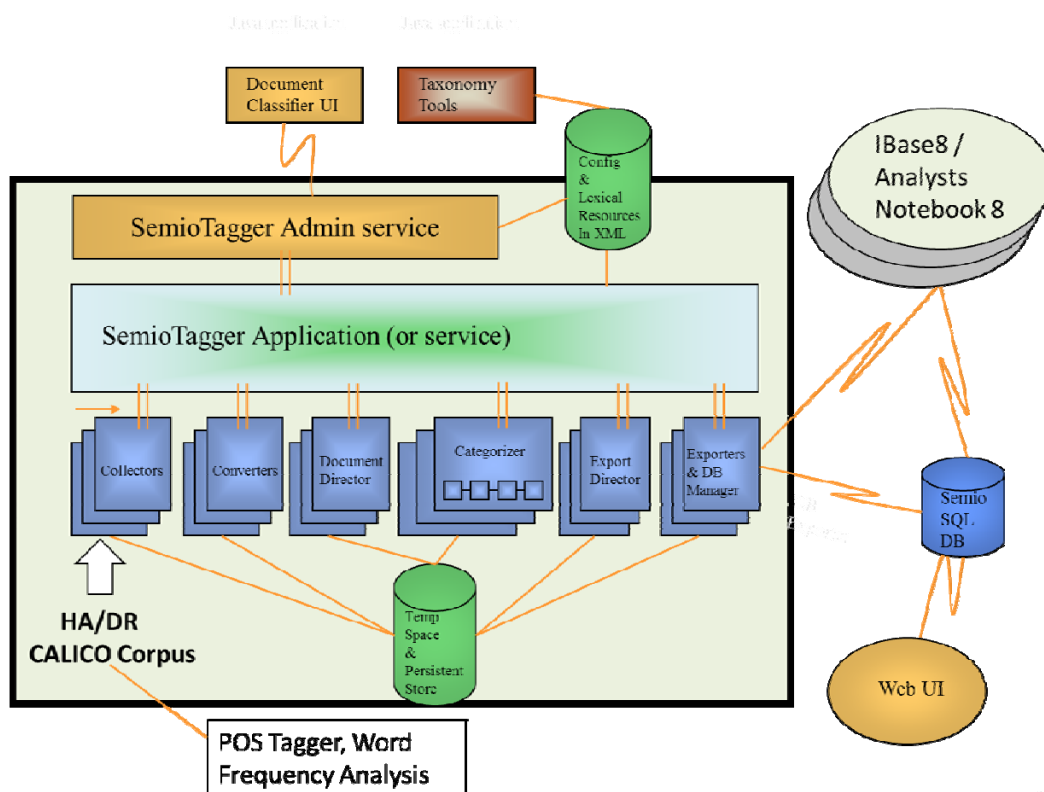


Figure 2. Architecture of the Processing Pipeline.

From a high-level view of this pipeline, a document classifier processes the corpus. This classifier is driven by a discrete human-formed taxonomy of categories and latch terms. The classifier provides a weighted mathematical value to latched terms and categories, which are described below in more detail. The output of the classifier is ingested into a relational database for complex query building. The results of these queries are then able to describe documents, categories, and phrases that answer specific questions. For example: a list of all HA/DR tasks for responding organizations. A few non-pipeline activities were also performed. These are part-of-speech tagging and word frequency analysis. Additionally, a word frequency analysis was generated just using the words contained within the phrases extracted by the document classifier for a more context specific view of word frequency. The tools used to perform the text analysis are fairly diverse. Each tool will be covered in more detail along with its role within the overall processing pipeline.

PDF Reader (Nitro PDF Reader v2.0.0.29)

The PDF Reader was used in two distinct modes. First, the entire corpus had to be converted into a text-only equivalent. This tool made this possible but could not operate in batch mode. This conversion into a text version (i.e., .txt files) was required for analysis by the part-of-speech tagger and word frequency analysis tools. The other use of this tool was to perform the same operation on a few PDF files within the corpus that were larger than 10Mbytes. This is a document classifier size constraint on any file type that it ingests. It presently supports all of the typical formats such as PDF, DOC, RTF, HTM given the 10MByte constraint.

OCR Tool (Brother Control Center v3.6.8.14)

There were some corpus member PDF files that had no extractable text. In one case it was limited by the author and in at least one other case; the text was simply inextricable due to font issues. The work-around was to print these PDF files to hardcopy and then perform an OCR scan. This process worked very well and resulted in clean text file equivalents with impressive accuracy.

Document Classifier (Semio Tagger 5.0a)

This application forms the heart of the processing pipeline. The tagger or classification engine will ingest nearly any type of content in the native formats – with size limits as described above. As ‘document classifiers’ are categorized themselves, Semio Tagger is considered a discrete taxonomically driven example. That is, no training corpus, and very high degree of knowledge engineer control using discrete lexical resources. This classification engine type relies upon human crafted taxonomies, replacement lists, exception lists, and lexicons. It was selected based upon its availability.²

The development of these taxonomies is usually performed by a knowledge engineer that will continually refine them on a periodic basis. In this case, the refinement was performed after a hand-off of the taxonomies from the social scientist to the computer scientist running the classifier. The taxonomies were mildly altered to improve latch quality and accuracy. The use of replacement lists,

² Semio Tagger software itself is no longer sold or supported commercially – a victim of the Dot.com bust.

exception lists, and lexicons was not part of this experimental framework due to the high degree of human oversight required to create and monitor their efficacy. **Figure 3** below shows the user interface for the document classifier.

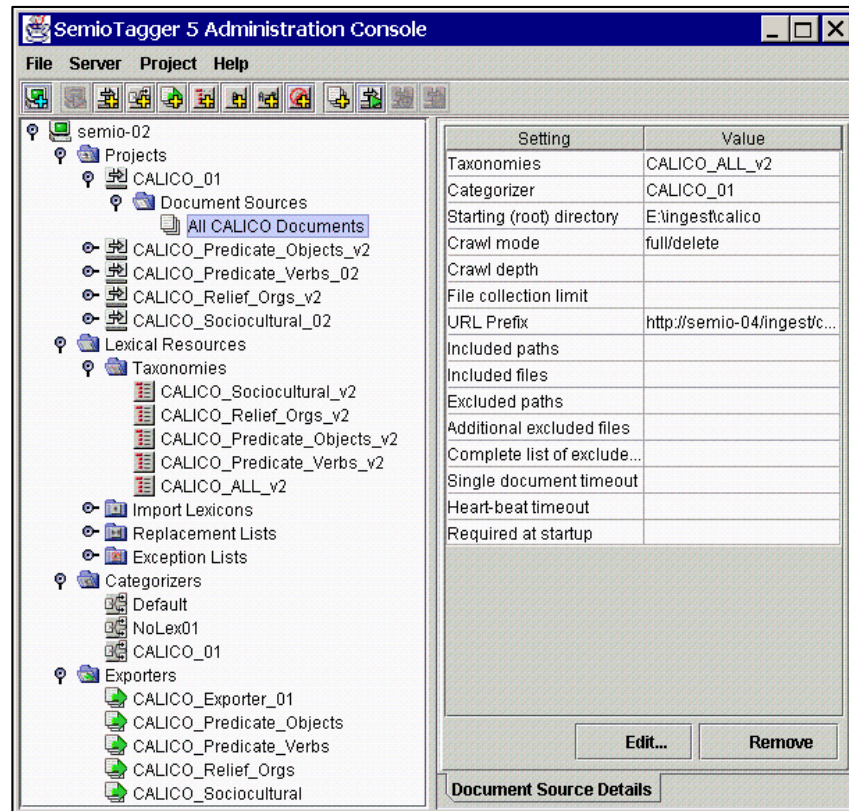


Figure 3. User Interface for the Document Classifier.

In **figure 3** above, the classification projects can be seen. There were four individual projects with one master project that represents the aggregate of all the classification projects. The aggregated project is the one used for analysis as complex queries may be constructed that are intersection, union, or even both. The best way to understand the output of the classifier for later discussion is that it is composed of: categories, phrases, documents, and weighted links. **Figure 4** below describes this object relational construct.

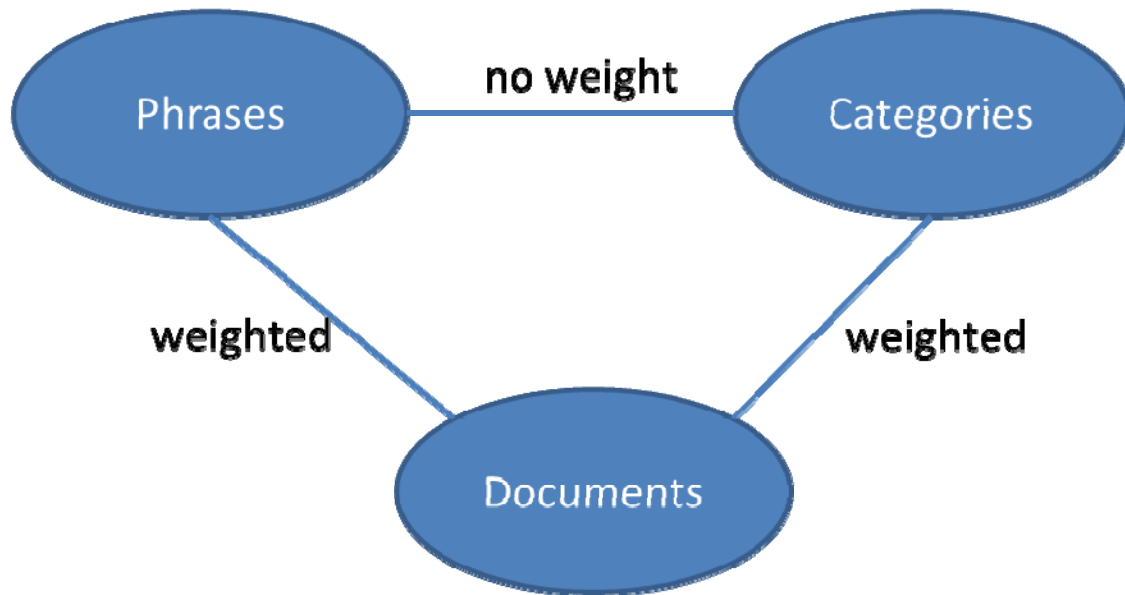
Taxonomy: defines categories & latch terms that hook phrases

Figure 4. Taxonomy Development Triad.

Figure 4 above should be interpreted as follows: 1) documents get classified into categories with some degree of fitness; 2) documents contain phrases that are weighted by importance within the scope of the document; 3) phrases are either within a category or not (i.e., a binary or non-attributed relationship).

The computer scientist driving the final taxonomy improvements used some of the KE or knowledge engineering workbench tools provided. These tools are part of the Semio Tagger application environment. As an example below, **figure 5** describes the global perspective of a classification run. In this example, 100% of the documents have been latched into at least one category. Not all categories actually latched a document. Some discussion on this follows in the next section.

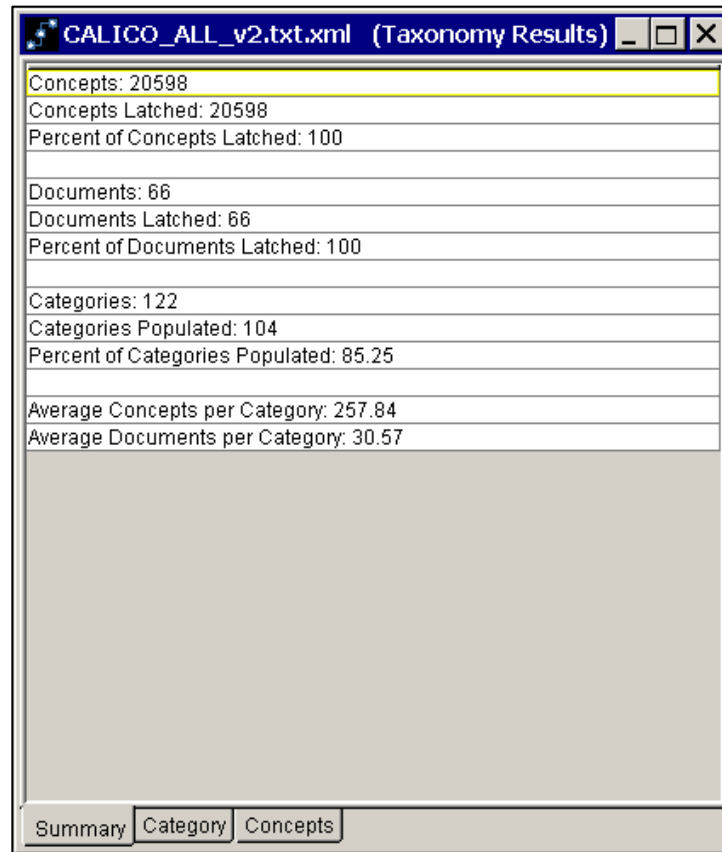


Figure 5. Global Perspective of a Classification Run.

Assessing taxonomy performance is an iterative process. These steps are: 1) make the run, 2) examine the results, and; 3) make adjustments to the taxonomy and repeat until satisfied. **Figure 6** below shows the user interface to assess the performance of a classification run. The left-hand portion of the window shows the taxonomy categories and their latching performance. Example, the category “Education” latched 341 phrases or concepts. The right-hand portion of the tool displays these phrases along with their weights and document counts (i.e., how many documents within the corpus contains that phrase). While this assessment is revealing, what a knowledge engineer really wants to know is what phrases went unlatched.

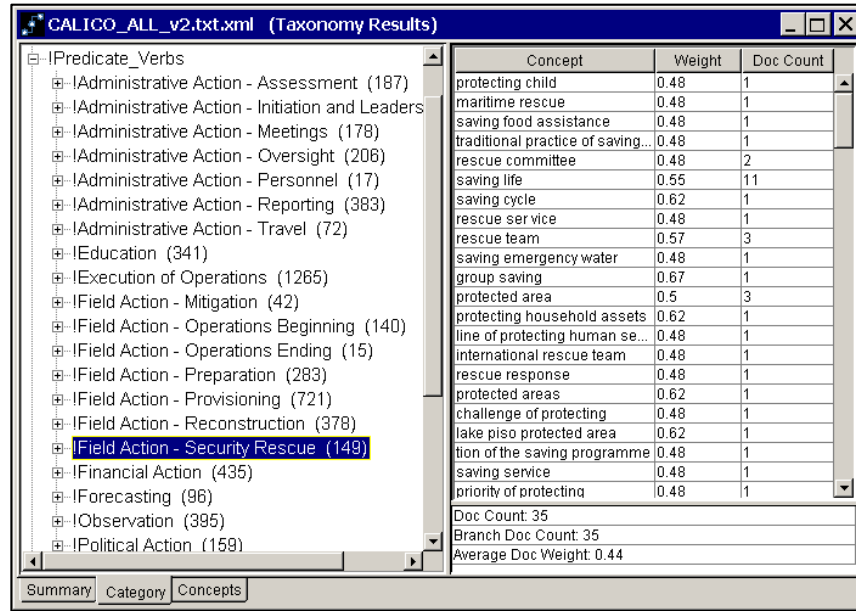


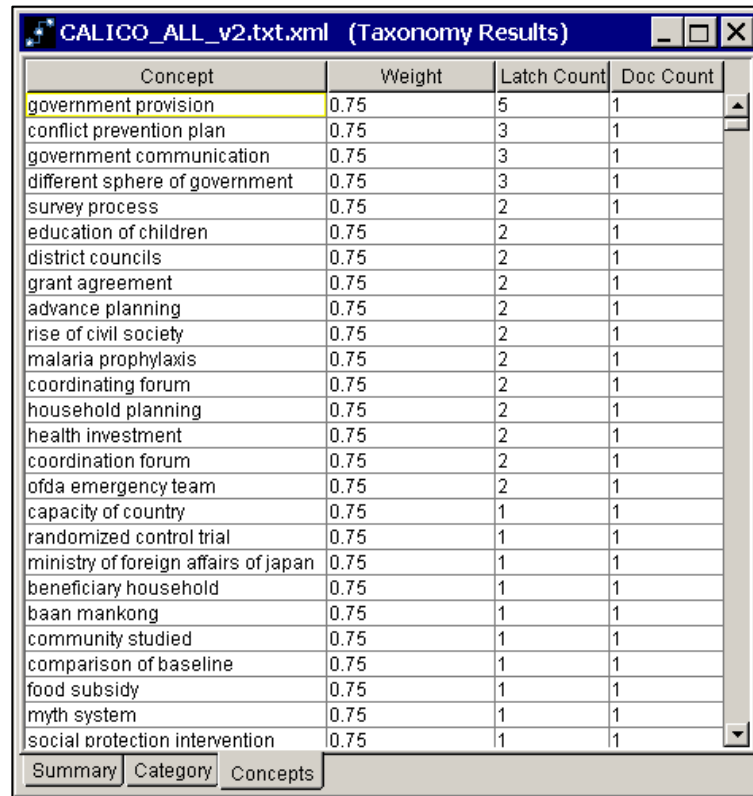
Figure 6. User Interface to Assess the Performance of a Classification Run.

The task of improving the taxonomy is essentially assessing the latch or hook rule performance. This is either an over-latching or under-latching problem. To view phrases that are unlatched, the tool as shown in the **figure 7** below can be sorted to show all unlatched phrases. Once the knowledge engineer can view these unlatched phrases, then he or she can add them to the taxonomy if appropriate.

During the processing phase of this effort, this iterative technique was used several times for three of the four taxonomies – estimated 7-10 runs each. The result was to improve each taxonomy prior to combining them. Once all individual taxonomies were performing better they were combined.

Taxonomy Filename	Original Latch Count	Latch Count after Improvement	% change	Difference
CALICO_Predicate_Objects_v2.txt.xml	543	836	53.96%	293
CALICO_Predicate_Verbs_v2.txt.xml	530	632	19.25%	102
CALICO_Relief_Orgs_v2.tax.xml	179	206	15.08%	27
CALICO_Sociocultural_v2.txt.xml	245	245	0.00%	0

Figure 7. Improved Taxonomy.



Concept	Weight	Latch Count	Doc Count
government provision	0.75	5	1
conflict prevention plan	0.75	3	1
government communication	0.75	3	1
different sphere of government	0.75	3	1
survey process	0.75	2	1
education of children	0.75	2	1
district councils	0.75	2	1
grant agreement	0.75	2	1
advance planning	0.75	2	1
rise of civil society	0.75	2	1
malaria prophylaxis	0.75	2	1
coordinating forum	0.75	2	1
household planning	0.75	2	1
health investment	0.75	2	1
coordination forum	0.75	2	1
ofda emergency team	0.75	2	1
capacity of country	0.75	1	1
randomized control trial	0.75	1	1
ministry of foreign affairs of japan	0.75	1	1
beneficiary household	0.75	1	1
baan mankong	0.75	1	1
community studied	0.75	1	1
comparison of baseline	0.75	1	1
food subsidy	0.75	1	1
myth system	0.75	1	1
social protection intervention	0.75	1	1

Summary Category Concepts

Figure 8. Sorted Unlatched Phrases.

One of the primary persistence mechanisms of the document classifier is a relational database. This database contains all matters related to the corpus and by default will ingest all categories defined within taxonomies. The database tables are representative to some degree of **figure 4** [the triad] above describing the relationship between documents, categories, and phrases/concepts. The database also contains text equivalents of each document as well as an XML tagged version that may be used by external applications. The tool discussed next will describe how this database is ingested into another relational data store to support complex query construction. The following **figure 9** shows the internal Semio Tagger database diagram.

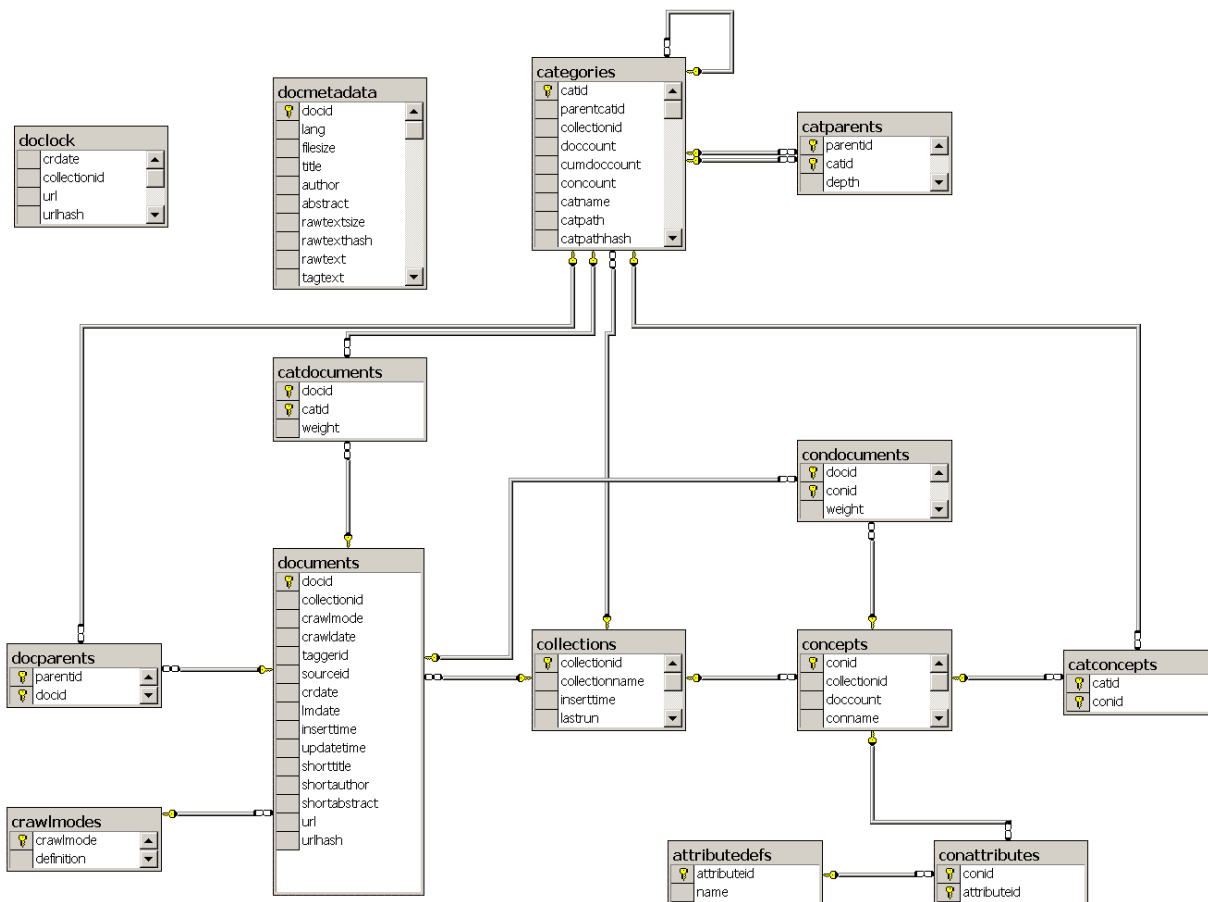


Figure 9. Internal Semio Tagger database Diagram.

Node Link Analysis Tools (IBase v8.5.7 / Analyst's Notebook v8.7.3)

The purpose of these two related tools is to support the complex query construction for the final analysis. The details of these queries will be addressed in the next section. As mentioned above, the content from the document classifier is persisted into a relational database – Microsoft SQL Server. This database is designed for access and exploitation by other tools. In this case, the tools covered here then take that database content into yet another specialized relational database known as IBase. The IBase database is also a standard SQL Server database and feeds directly from the Semio Tagger database using scripted extract, transform, and load procedures. The IBase schema is designed to support network and link analysis. In the case of this analysis, it is used to build intersecting queries based upon the taxonomy categories. An example query would be to show the intersection of HA/DR tasks and sociocultural information. IBase has a well-documented object-relation style of schema design and has been a popular tool in both law enforcement and the intelligence community for well over a decade.

The design of the IBase schema can be seen in the **figure 10** below -- note that the entity and link types with the zero record counts are not used in the CALICO effort. These are placeholders for specific geographic latitude and longitude information that has been identified by a tool such as MetaCarta's

GeoTagger. The link type describes the three possible links and shows the number of links contained within the database. The “Category2Category” link type is simply to enable a hierarchical view of the categories graphically.

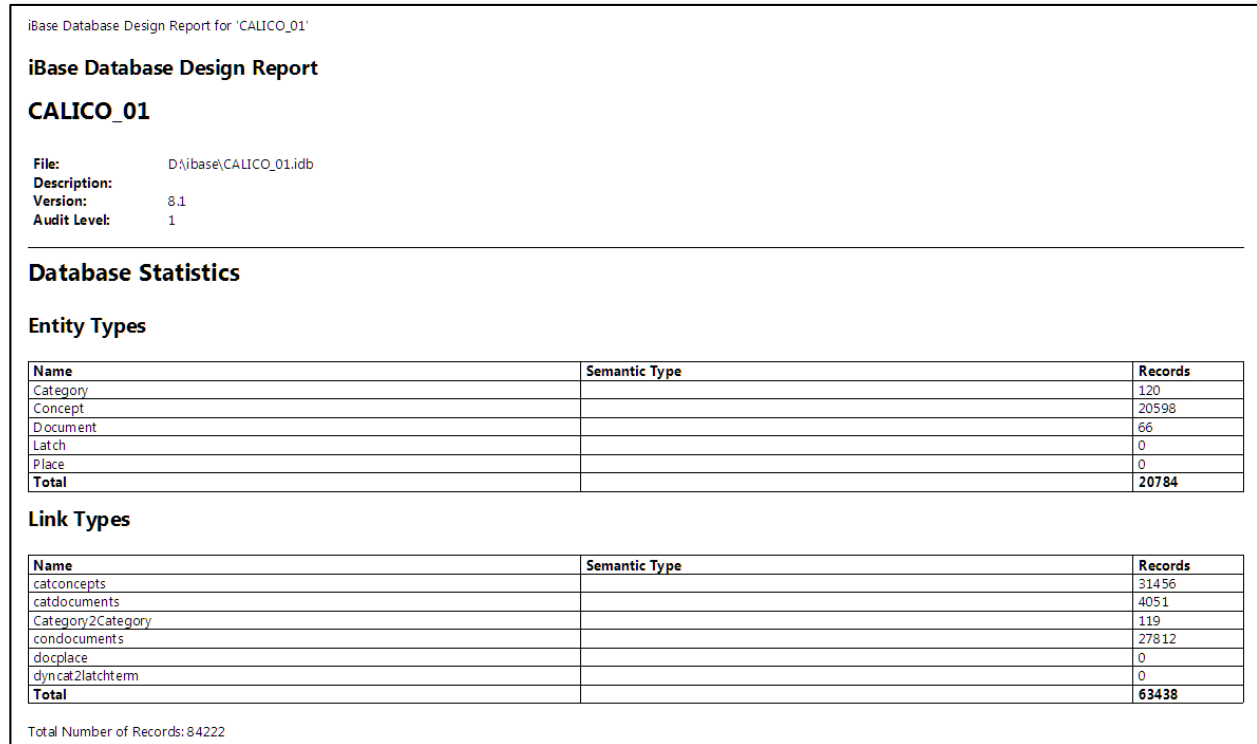


Figure 10. iBase Schema.

By using this link analysis database construct, it provides a robust way to construct queries but also a means to visualize the results of the document classifier. For example, the **figure 11** below shows a small subset of the categories and phrases found within just one document. The display shows the weighting as well as the number of times a particular phrase was latched in another document.

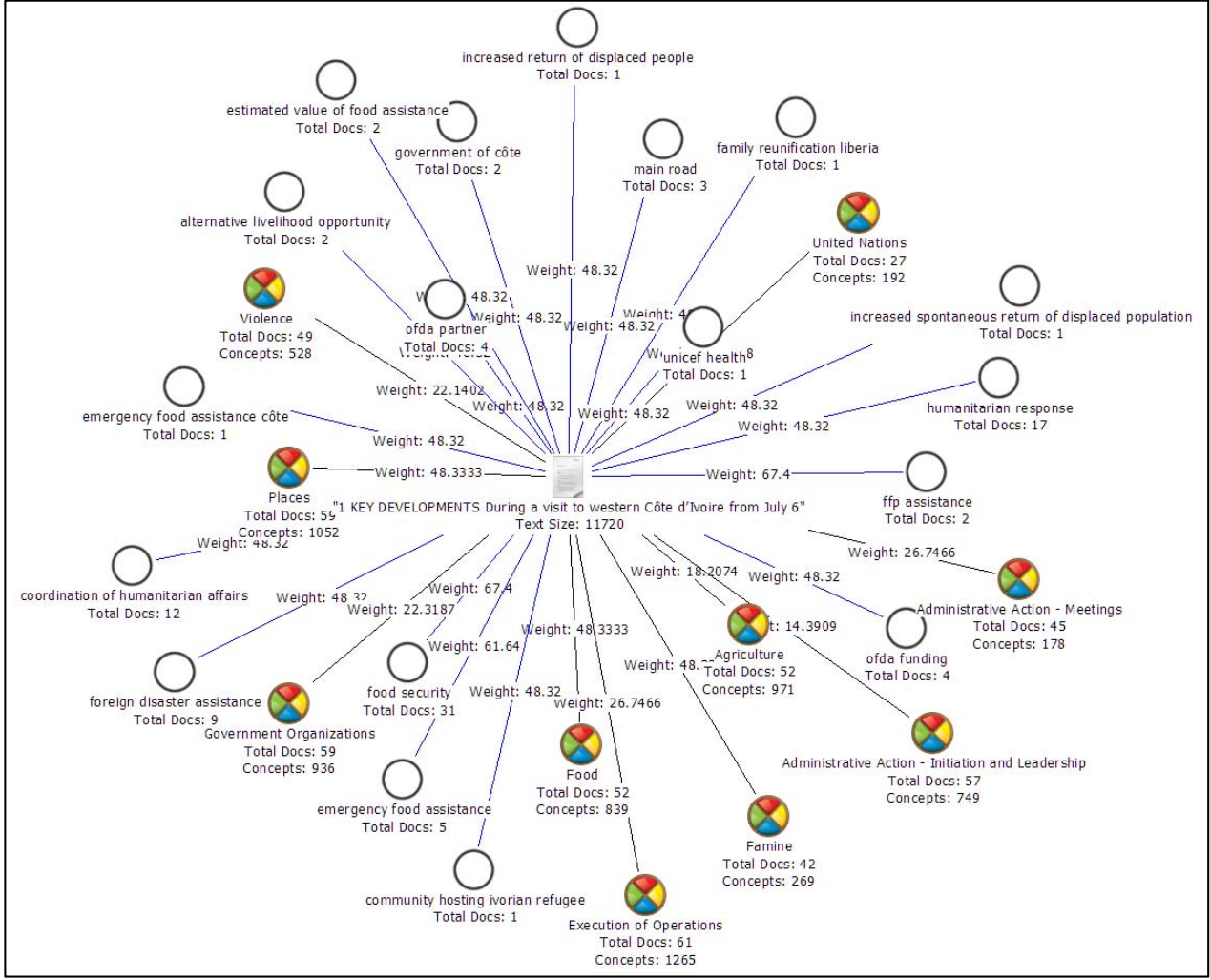


Figure 11. Subset of Categories and Phrases.

The following **figure 12** shows an example of a complex query within the IBase user interface. This allows a drag-and-drop construction of queries that may be immediately tested and refined. In this case, this query is actually using the result set from another query as indicated by the question mark balloon.

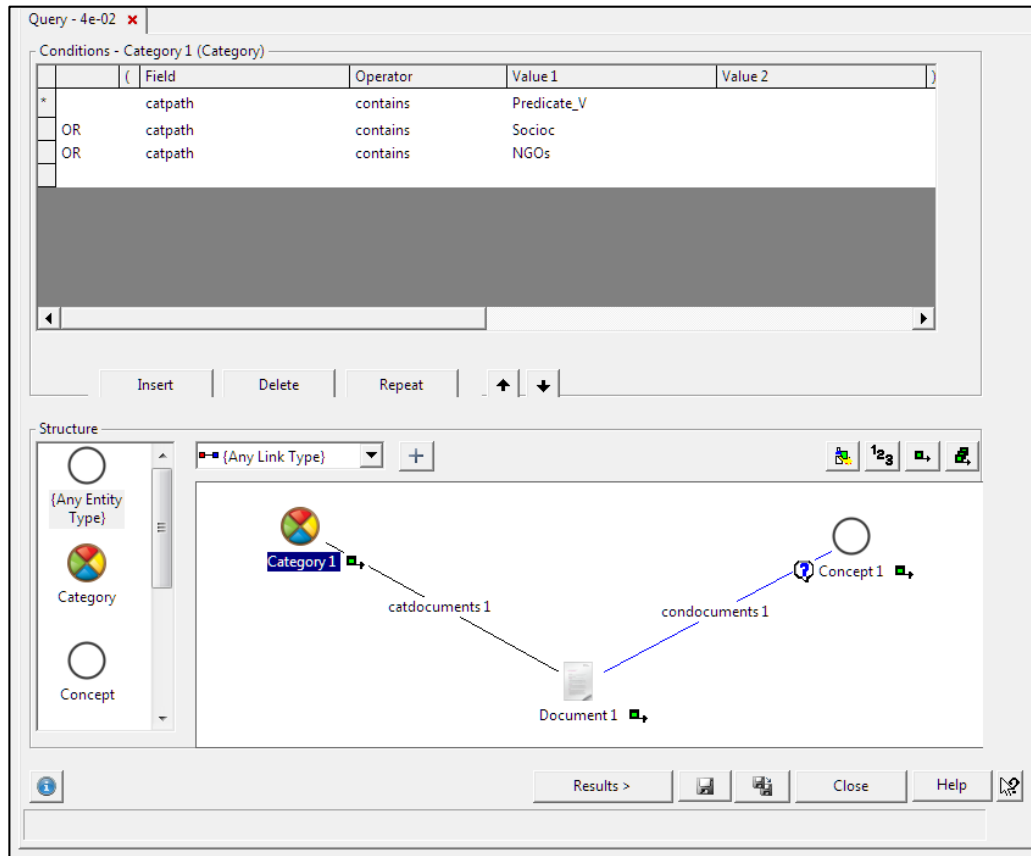


Figure 12. Example of a Complex Query within the IBase User Interface.

The result of the above query in a graphical view including the linking edges can be seen in the **figure 13** below. This complex view of relations is mildly interesting but not very useful for analysis. The typical result set from the query above is sent to a report as in the CALICO reference document "**4e2.pdf**" sans the interconnecting links/edges.

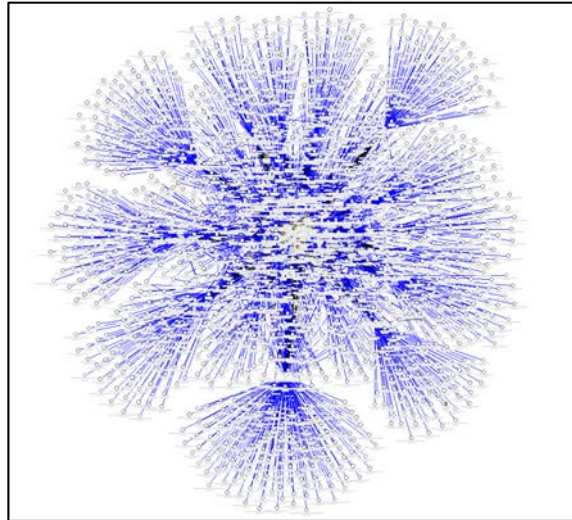


Figure 13. Query Graphical View.

Web Based Viewer (Semio Skyline 2.0)

A component of the Semio application is a web-based view of the classification results. This is a non-graphical interface that allows for multiple search strategies. Searching by categories and key words and multiple combinations of these provides fast access to the CALICO corpus. **Figure 14** below shows the user interface at the top level.

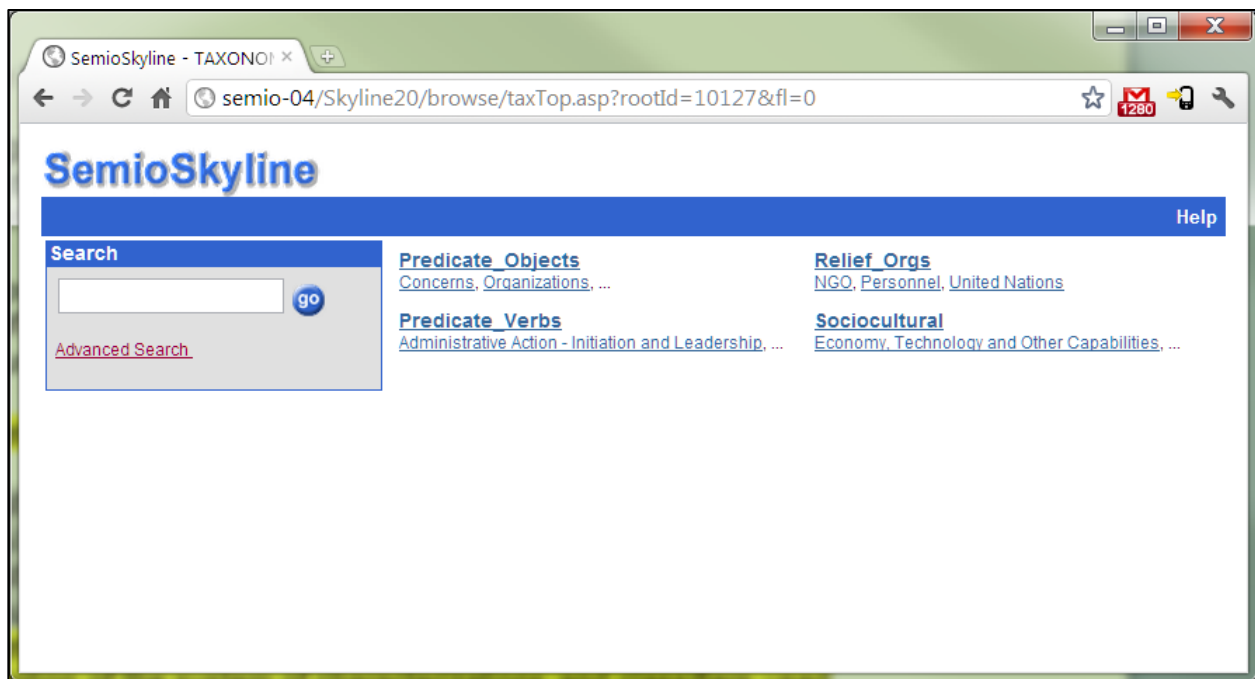


Figure 14. Web User Interface for Viewing Document Classifier Results.

Additionally, the search interface is shown that exposes the full-text indexing features of the interface. The type of interface is sometimes referred to as faceted search. This provides a search engine capability that may be constrained by facets (i.e., categories of information).

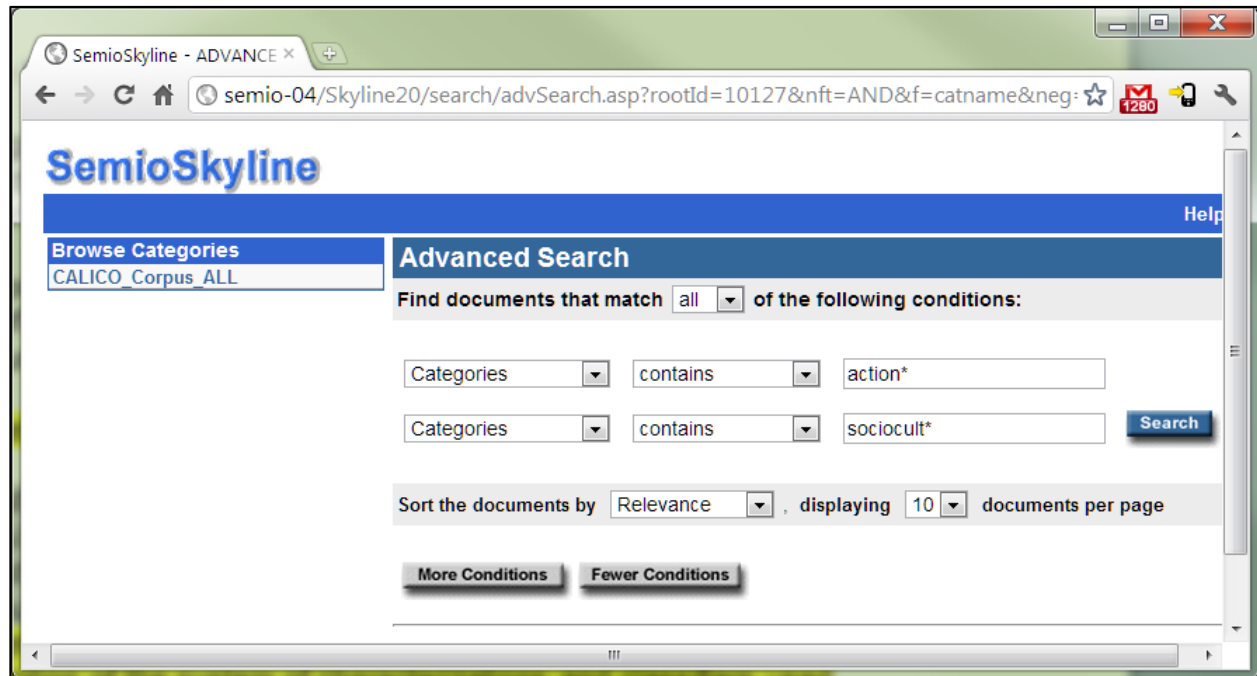


Figure 15. Advanced Search Interface for Viewing Document Classifier Results.

Finally, the document display interface was a useful tool to assess the performance of the classification engine. Shown below in **figure 16** is that interface.

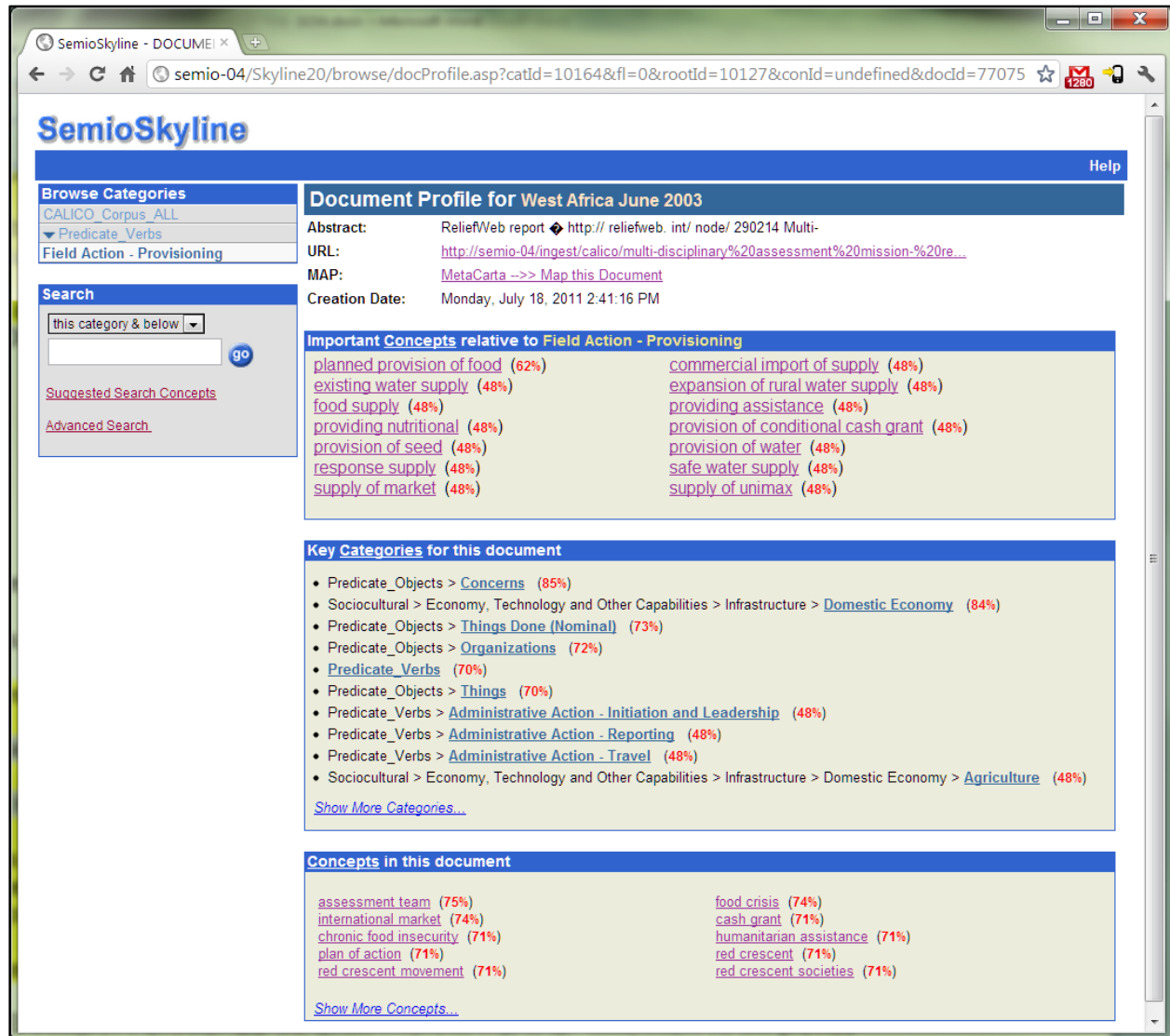


Figure 16. Web User Interface for Viewing Document Profile.

POS Tagger (VisualText – POSTagger v2.6.0.3)

Part-of-speech tagging was performed outside of the document pipeline described above. The results are included in the CALICO reference document section – files in the folder named “pos.” POS tagging was not a component or prerequisite for using the document classifier (Semio Tagger).

Word Frequency (MaxQDA v10 R180811 / IBase v8.5.7)

Word frequency was produced in two distinct manners. First, the MaxQDA tool was used to generate the word frequency from the entire corpus. This was followed by word frequency calculations that filter the words down to a set that were contained within the phrases extracted by the document classifier (i.e, Semio Tagger). The usefulness of the latter more constrained set of word frequencies seemed to better contextualize the results as they are a derivative of the taxonomy.

In summary, this section details the technical tools used with some added insights into the methodology. Again, the text-analytic methodology is very centered upon exploiting the results of a document classification engine. The exploitation takes place in another set of tools that are specialized in their ability to make very complex queries into a network of related nodes and edges. In this case, those nodes and edges are the documents, the phrases, and the categories. The edges are the weightings or fitness level of a document within a category, a concept's importance within the scope of a document, and finally what categories have latched to which concepts. These triads of relations are described above in **figure 4** [triad]. Section VI will contain the interpretation of the technical process described above.

Document Summarization (Copernic Summarizer v2.1)

This tool was used to create a machine based summary of each document. These summarized versions are not intrinsic to the overall text-analytic process. These summarized versions are within the CALICO document reference collection – filename “**CALOCO_txtsums.zip**”.

V. Conclusion from Analysis

The primary result of the use of the text analytic pipeline was to create a system that can be queried to answer the basic research questions posed in this project. This section presents answers to these research questions in two parts. Part A addresses how the system can be queried and provides sample answers to those queries. Part B summarizes the aggregate result of these queries to provide general answers to the basic research questions concerning a) who the relief organizations are that act, b) what actions they undertake, c) what objectives (concerns, goods and services) relief organizations have, and d) what sociocultural factors are germane to HA/DR activities executed by relief organizations.

A. Findings – Direct Results of Queries

This section will be broken down into distinct areas that strive to answer the research questions put forth within the original scope of work “Task 3” and portions of “Task 4”. In some cases, these tasks specified resulted in the creation of project artifacts that are provided as CALICO reference documents (e.g., POS tagging). These are mostly human-readable artifacts but in some cases machine readable only. Last item, there are also some machine readable versions of reports in Microsoft Access format that are provided for any post-project analysis.

POS Labeling, Word Frequency Analysis, Tagged Corpus

Task 3 parts 2 and 3 are addressed as project artifacts in the CALICO reference documents provided. These are subsets of the Semio Skyline web interface that was crawled into an operational file system equivalent; and, an XML dump file of the tagged corpus that include the tags applied by the document classifier – reference documents “**CALICO_WebCrawl.zip**”, and “**CALICO_ALL.cbf.xml**” respectively. The web crawl artifact can be browsed but not searched as it is a standalone set of files with no application functionality. The XML dump file can be further processed as a post-project activity by other researchers.

Task 3 part 1 references “*Part-of-speech labeling and word frequency analysis*”. The POS tagging output is represented also as a CALICO reference document – “**CALICO_pos.zip**”. The use of the POS results was not an intrinsic component of the text-analytic pipeline employed within this project as earlier indicated. POS results are offered as a stand-alone deliverable.

Word frequency analysis was performed in a simple mathematical sense and can be found in the CALICO reference documents “**WF_AllTexts.xlsx**”, “**WF_ConLatchedTop1k.xlsx**”, and “**WF_ALLCompare.xlsx**”. These artifacts contain the raw numbers of the frequency analysis. No statistical measures were applied to these results other than created by the tool used. It was thought to be more useful to compare the traditional frequency result set with another set where words that were constrained. This constrained set of words is derived from the phrases extracted by the document classifier based upon the taxonomies created for CALICO. **Figure 17** below a comparison using the top 14 most frequent words for both of these word frequency sets.

Frequency from latched phrases		Frequency derived from all text	
Word	Frequency	Word	Frequency
food	800	disaster	3885
government	701	children	2893
health	688	people	2862
community	669	food	2859
disaster	541	humanitarian	2848
assistance	535	africa	2822
response	531	more	2747
child	515	development	2683
risk	460	government	2632
water	381	health	2614
capacity	369	international	2534
management	359	management	2392
humanitarian	347	support	2321
conflict	324	assistance	2246

Figure 17. Word Frequency Analysis Partial Results.

The word frequencies shown above show the clear filtering impact of the latched phrase constraint. This is offered as an anecdotal item as time and scope did not permit doing further analysis. The 'latched phrase' word frequency table seemed more contextualized to the research questions of HA/DR and their activities and sociocultural elements.

Association Analysis

This section covers the essential research questions. These questions are derived from the scope-of-work sections 3.c.4 and 3.d.3.d though g. The overarching CALICO mission is to gain an *"...increasing understanding of indicators of cultural relevance within running text and of the connection between that relevant sociocultural information and the specific tasks described in the running text."* The guiding or top level aggregate research question is *"What do government and NGO reports on HA/DR activities tell us about the information needed to support effective interventions to ameliorate the human consequences of disruption and to enhance the effectiveness of indigenous governance?"* The more discrete individual research questions that setup the analysis are as follows:

- 4.d - A list of all predicates (tasks) carried out and the agents affected by the tasks by the authoring HA/DR organization for each document in the corpus.
- 4.e - A list of all sociocultural information cited in support of the individual HA/DR tasks carried out by the organization for each document in the corpus.

The above two research questions will be addressed individually followed by further technical discussion regarding the automated routines and technical tools implemented. The final portion of this section will be a discussion and synthesis of the analysis from the two questions above.

In addressing the first research question, it was critical to develop the required queries into the document classification data store – IBase. The query had to essentially intersect the three branches of the taxonomy as shown below in **figure 18**.

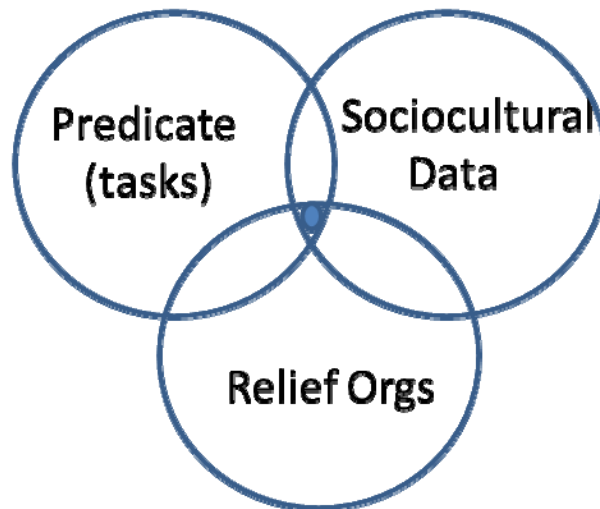


Figure 18. Venn Diagram of Query Intersection Construct.

The Venn diagram above describes the general framework of the query to be constructed. The detailed description of this query is described as:

- From all documents within corpus: (pseudo code query example)
 - Produce all documents, phrases, and categories where:
 - only categories the documents were highly classified within - per the Venn diagram
 - only phrases highly weighted (high fitness) within the scope of the document
 - only phrases that were latched into the above categories
 - only phrases that were found within 4 or more other documents

The pseudo code describes the query in terms of an intersection of taxonomic branches (i.e., Venn diagram) then further limited to a filtering of the strength of fitness of documents, categories, and how important a phrase is within the document. This is further filtered by showing only phrases that have been successfully latched into at least 4 documents corpus wide. **Figure 19** below shows the graphical construction of this query with the final “greater than 3” constraint upon “doccount”.

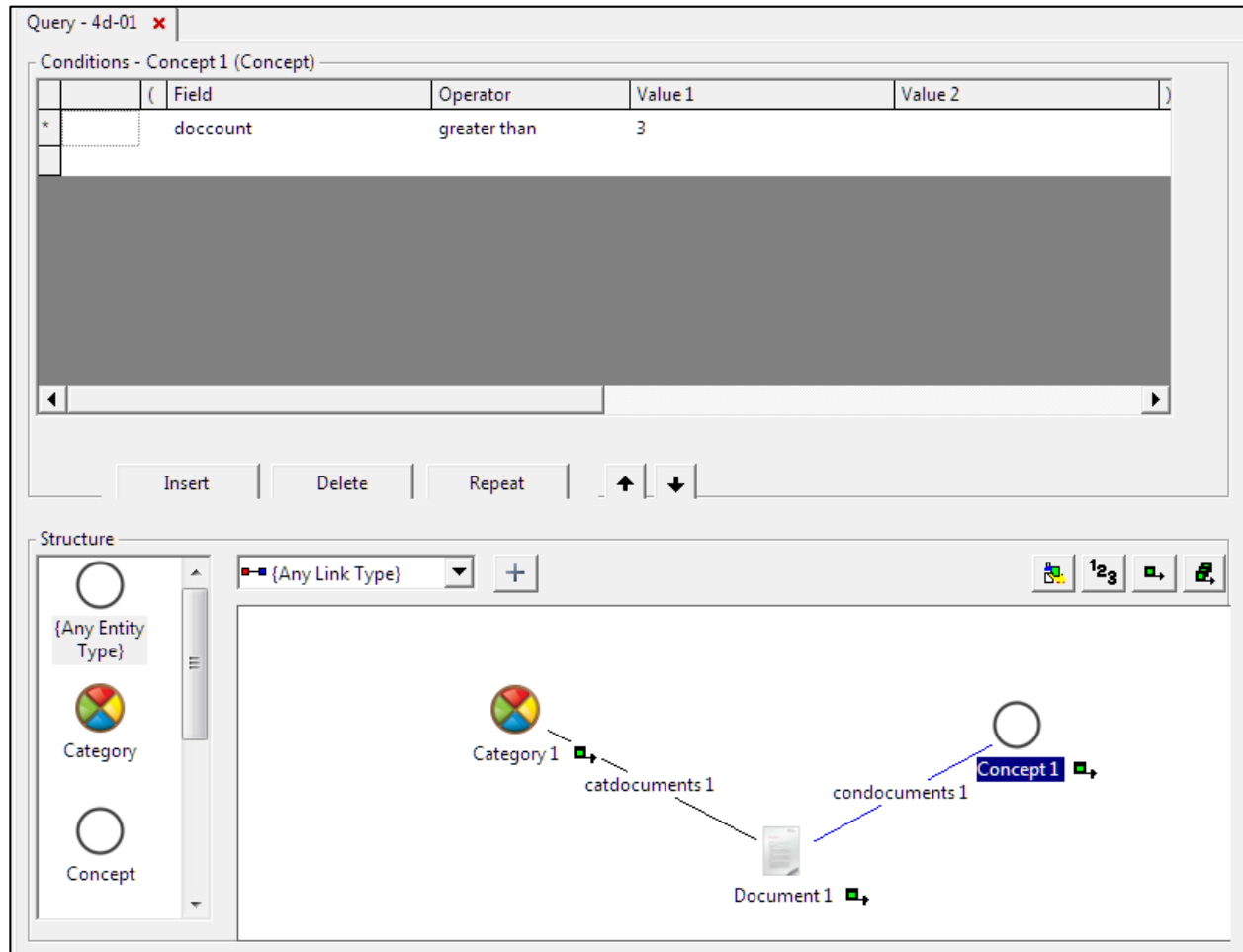


Figure 19. Query Builder User Interface.

The results from this query are found in PDF format in the CALICO reference documents collection – filename “4d.pdf”. This report style layout can also be seen in **figure 20** below.

Tasks and Agents Effectuated (for each document) [4.d]		CALICO 2011
shortabstract	rawtextsize	
1 KEY DEVELOPMENTS During a visit to western Côte d'Ivoire from July 6		11720
Strength of Taxonomic Classification (Fitness Level)		
weight	catname	cumdoccount
66.2258	Beneficiaries	66
51.7122	Things Done (Nominal)	66
48.3333	Assisting	59
48.3333	Businesses	49
48.3333	Community Village	63
48.3333	Coordinating	39
48.3333	Disease Morbidity	47
48.3333	Famine	42
48.3333	Food	52
48.3333	General Beneficiaries	58
48.3333	General Concerns	61
48.3333	General Organizations	61
48.3333	Macroeconomics	51
48.3333	NGO	60
48.3333	NGO functionaries	51
48.3333	NGOs	60
48.3333	Places	59
48.3333	Supportive Action - Augmenting	56
48.3333	US Government, Non-Military	21
Tasks & Agents Effectuated (Fitness & Frequency within Corpus)		
weight	conname	doccount
71.4	ofda assistance	4
67.4	food security	31
67.4	humanitarian assistance	36
67.4	market systems	4
61.64	emergency food assistance	5
61.64	health center	7
61.64	people displaced	10
61.64	security concern	7
48.32	assessment team	11
48.32	camp coordination	4
48.32	coordination of humanitarian affairs	12
48.32	department of state	12
48.32	displaced people	14
48.32	foreign disaster assistance	9
48.32	host community	7
48.32	humanitarian response	17
48.32	implementing partner	12
48.32	international committee of the red cross	17
48.32	ofda funding	4
48.32	ofda partner	4
48.32	response effort	5

Figure 20. Query Results in Report Layout.

The key components of this query report are: (column names from report snippet as shown above)

- shortabstract - title of the document as extracted by the extraction process
- rawtextsize - a normalized or text-only file size (size without graphics)
- weight - a measure of fitness of this document within the specific taxonomic class
- catname - category name (taxonomic class)
- cumdoccount - number of documents in the corpus assigned/latched into this category
- weight - a measure of importance of the phrase within this document
- conname - phrase/concept text
- doccount - number of documents that also contain this phrase/concept

The report layout is specific to answering the research question 4.d. This report is also found as a Microsoft Access database for post-project analysis in the CALICO reference documents collection – filename “**4d.mdb**”. The top tasks are consistently providing food security, water availability, and shelter needs. The top agents or organizations mentioned include the International Red Cross, the Red Crescent, and the World Food Program followed by host nation governments. Complete views of these results are found in the CALICO reference document “**4d.pdf**”.

A human, non-quantitative review of the output report begins to form a consistent impression that, instead of representing the fine-grained on the ground, tasks necessary to execute HA/DR activities, the reports in the corpus represent higher-level, administrative functions undertaken by relief organizations. Discrete recipes for HA/DR tasks and the agents that perform them in many cases are simply not covered within this particular corpus. Agents are nearly always conflated into the responding organization. In some cases, the agents will be named by title – such as engineer, physician, or medical staff.

The next research question to be addressed is 4.e – broken into 4.e.1 and 4.e.2. The question bears some review – “*A list of all sociocultural information cited in support of the individual HA/DR tasks carried out by the organization for each document in the corpus.*” The approach to answering this question resulted in two distinct query sets [4e-01, 4e-02]. The first query result [4e-01] is not a per document result set. It was created to illustrate which tasks and actors show up most frequently within certain classes (i.e., categories) of documents. Another Venn diagram is useful here to describe the query as shown below.

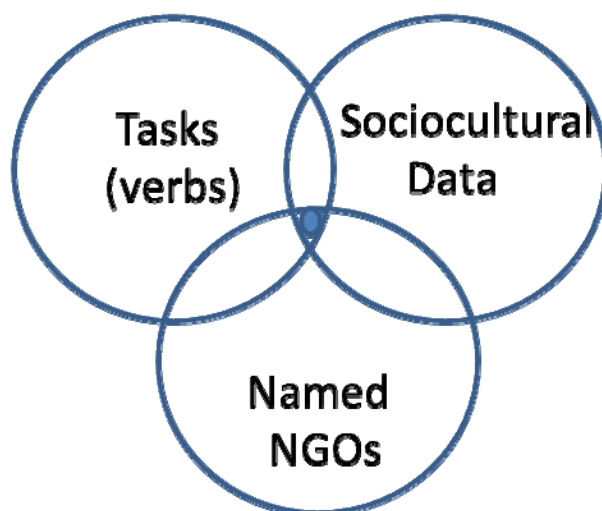


Figure 21. Venn Diagram of Query Intersection Construct.

The Venn diagram above describes the general framework of the query to be constructed. The detailed description of this query is described as:

- From all documents within corpus: (pseudo code query example)
 - Produce all categories and phrases where:
 - only categories are a taxonomic sub-branch of tasks (verbs)
 - only phrases that were:
 - products of latching into the Predicate_Verbs taxonomy branch
 - performed by a sub-query (note balloon in **figure 22**)
 - have also co-occurred in at least 2 documents within the corpus

The query builder interface for this query is shown below in **figure 22**.

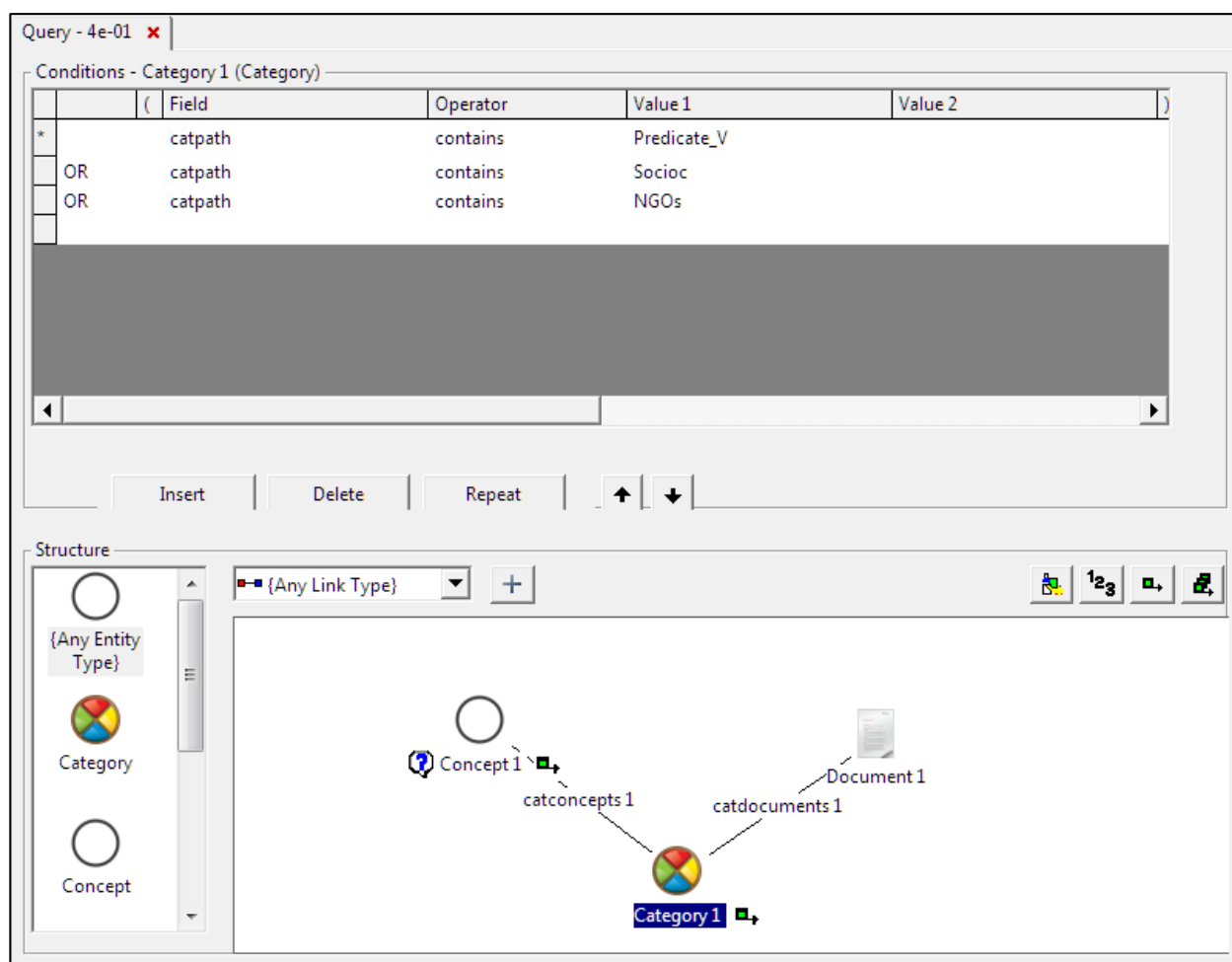


Figure 22. Query Builder User Interface.

The results from this query are also found in PDF format in the CALICO reference documents collection – filename “4e1.pdf” and “4e11.pdf” – the latter file eliminating the greater than 2 co-occurrence constraint. This report style layout can also be seen in **figure 23** below.

Sociocultural Task Categories and HA/DR Tasks (for each task area) [4.e.1]		CALICO 2011
catname	catpath	
Administrative Action - Oversight	CALICO_Corpus_ALL::Predicate_Verbs::Administrative Action - Oversight	
Sociocultural Tasks and Actors Linked to the HA/DR Tasks		
conname	doccount	
loss of life		7
monitor progress		5
loss of livelihood		4
community management of acute malnutrition		3
management of disaster		3
management of natural resource		3
management of severe acute malnutrition		3
administrative system		2
loss of family member		2
loss of home		2
loss of human life		2
loss of income		2
loss of property		2
maintenance of child		2
monitor ceasefires		2
monitor global hazard		2
monitor humanitarian condition		2
monitor implementing partners		2
monitor ongoing humanitarian program		2
stock management of drug		2

Figure 23. Query Results in Report Layout.

The key components of this query report are: (column names from report snippet as shown above)

- catname - category name (taxonomic class)
- catpath - full taxonomic path
- conname - phrase/concept text
- doccount - number of documents that also contain this phrase/concept

The report layout is somewhat specific to framing the research question 4.e. This report is also found as two Microsoft Access databases for post-project analysis in the CALICO reference documents collection – filename “4e1.mdb” and “4e11.mdb”.

This report displays what specific HA/DR tasks and actors are captured by specific taxonomic branches. This is more useful perhaps than the more limited view of tasks and actors on a document by document basis.

Similar to the above, the second query [4e.02] set assists in answering research question 4.e, but does provide the document-by-document reporting format. The Venn diagram is unchanged from query 4e-01 above. The detailed description of this query is described as:

- From all documents within corpus: (pseudo code query example)
 - Produce all documents, phrases, and categories where:
 - only categories the documents were highly classified within - per the Venn diagram
 - disregard weight or fitness for both categories and phrases
 - limit all phrases to those derived from the relief organization taxonomic branch

The query builder interface for this query is shown below in **figure 24**.

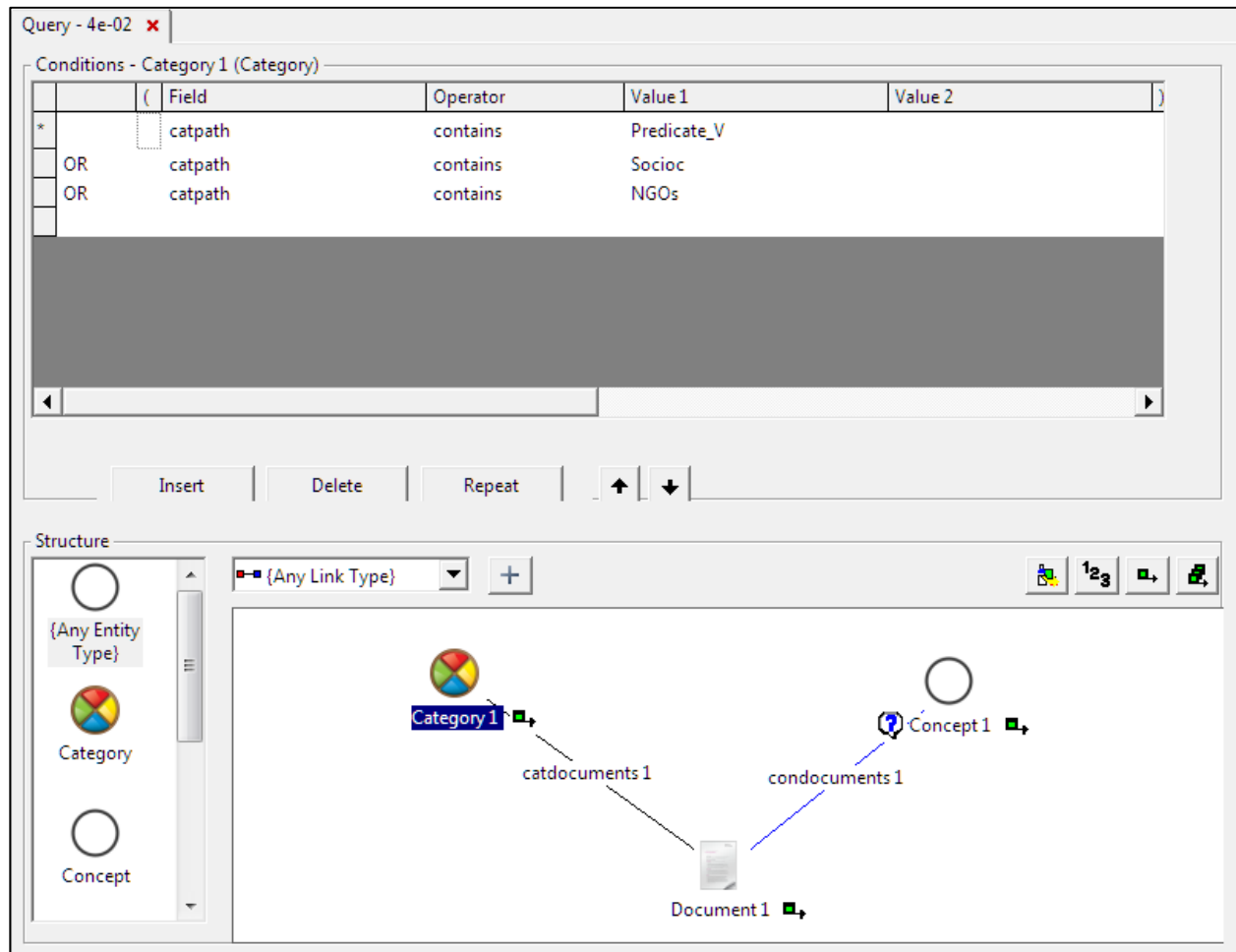


Figure 24. Query Builder User Interface.

The results from this query are also found in PDF format in the CALICO reference documents collection – filename **"4e2.pdf"**. This report style layout can also be seen in **figure 25** below with an example document from the corpus.

Intersect Sociocultural Information and HA/DR Tasks (for each document) [4.e.2]		CALICO 2011
shortabstract	rawtextsize	
? Hunger, disaster, hope: rethinking humanitarian action in Africa		58773
Strength of Taxonomic Classification (Fitness Level)		
weight	catname	cumdoccount
61.2926	Predicate_Verbs	66
58.5646	Domestic Economy	61
48.3333	Administrative Action - Reporting	51
48.3333	Agriculture	52
48.3333	Animal Husbandry	23
48.3333	Disease Morbidity	47
48.3333	Execution of Operations	61
48.3333	Field Action - Provisioning	55
48.3333	Field Action - Reconstruction	54
48.3333	Forecasting	23
48.3333	Macroeconomics	51
48.3333	NGOs	60
40.956	Administrative Action - Initiation and Leadership	57
40.8347	Supportive Action - Assistance	55
39.1185	Economic Risks	15
39.1185	Income	40
29.5916	Administrative Action - Oversight	50
29.5916	Financial Action	55
29.3919	Supportive Action - Augmenting	56
23.9795	Education	34
19.0652	Migration	22
17.2359	Education	51
17.2359	Field Action - Preparation	49
17.2359	Field Action - Security Rescue	35
17.2359	Research	50
NGOs Executing Tasks (Fitness & Frequency within Corpus)		
weight	conname	doccount
73.664	red crescent soci	1
73.664	red cross red crescent	5
71.4	international red cross	7
67.4	kenya red cross society	3
67.4	red crescent movement	8
67.4	red crescent societies box	1
61.64	gambia red cross society	1
61.64	international committee of the red cross	17
61.64	national red cross	4
48.32	13-kilometre kenya red cross society pipeline	1
48.32	bique red cross society	1
48.32	civil society capacity	1

Figure 25. Query Results in Report Layout.

The key components of this query report are: (column names from report snippet as shown above)

- shortabstract - title of the document as extracted by the extraction process
- rawtextsize - a normalized or text-only file size (size without graphics)
- weight - a measure of fitness of this document within the specific taxonomic class
- catname - category name (taxonomic class)
- cumdoccount - number of documents in the corpus assigned/latched into this category
- weight - a measure of importance of the phrase within this document
- conname - phrase/concept text
- doccount - number of documents that also contain this phrase/concept

The report layout is more specific to answering research question 4.e. This report is also found as a Microsoft Access database for post-project analysis in the CALICO reference documents collection – filename “4e2.mdb”.

This query [4e.02] report attempts to present a more focused view of sociocultural information within the context of HA/DR tasks specific to responding organizations – in this case limited to a set of named non-government organizations.

The corpus is largely a set of content authored for a Western target audience of interested parties, benefactors, and potential donors, and reflect those concerns. Specific sociocultural information germane to executing HA/DR activities on the ground are likely to be found in operational field notes (i.e., lessons learned from deployment – what not to ask/say/do etc.) from fielded staff and deployed lower-level management. The textual artifacts (i.e., corpus) that may be a closer to the perfect fit for these research questions would contain these facts, nuances, and lessons on how things really get accomplished in the field – some of which may be potentially distasteful information a disaster responder would not want publish.

Summary – Conclusions from Analysis

This section has covered the technical aspects of answering the research questions. A technical approach was developed to exploit a document classification engine. The results of this engine were then ingested into an object-relational construct (IBase) where complex queries were built to distill information supportive of the research questions. These queries use an approach that intersects distinct taxonomic branches to hone in on the distinct set of documents, agents, and actions performed by those agents.

The conclusions here are based upon quantitative information but tread also into some subjective interpretation. One concluding remark is that the nature of the corpus is such that it generally does not cover the detailed and granular natural language phrases connected to specific tasks and agents performing those tasks. Tasks and agents are rather conflated at a goal, title, and organizational level. Added discussion follows below regarding corpus selection and possible new research directions. The research questions are only partially satisfied by this text-analytic approach as a matter of content granularity of the corpus versus the technical analytical approach. This will be discussed further in the next section.

B. Findings – Aggregate Findings and Generalizations

The detailed analysis of specific queries is beyond the scope of the present work, as the queries actually generated extremely detailed reports on each document, which resulted in thousands of detailed answers per query. However, the results of the queries can be aggregated at the corpus level, providing a more general, and generalizable, view of what the corpus says regarding who executes what tasks and for what objectives. This section examines the results of the text analytic queries to address the research questions of this study from a more general perspective.

One measure of the importance of a concept in the corpus is the number of documents in which the concept occurs. Those concepts that occur across the corpus are probably more intrinsic to relief organization activity in HA/DR. The number of documents in which concepts occur will be used to answer the fundamental question, “Who does what for what purpose?” We also address the additional question concerning the role socio-cultural factors play in HA/DR activity in Africa.

The data for these general analyses are aggregated from the cumulative document counts (cumdoccount) from the queries described in section A.

1. Who Acts?

Specific relief organizations were, not surprisingly, mentioned in many documents, occurring in 61 of the 66 documents in the corpus. The revised taxonomy contains 4 high-level categories into which the various identified relief organizations were placed. These categories are: NGOs, UN, US non-military, and Personnel.

Relief Organization Type	Number of Documents Mentioned	PCT Documents
NGO	60	90.9
Foreign Government	59	89.4
US non-Military	21	31.9
Personnel	33	50.0

Figure 26. Identified Relief Organizations.

Not surprisingly, NGOs were mentioned in over 90% of the documents, since the subject of this study was NGO activity. Consistent with the hand-coded sample, Foreign government organizations were almost as frequently mentioned (89.4% of documents), revealing the collaborative nature of HA/DR activity; it is a close collaboration of NGOs and host nation (HN) governments. The UN was mentioned in over 40% of the documents, indicating the common role the United Nations plays in underwriting and supporting aid worldwide. The fact that U.S. non-military aid organizations were mentioned in nearly a third of the documents attests to the pervasive presence of U.S. aid activity on the African continent. U.S. organizations mentioned include USAID, USDA, USGS, and the US Department of Education. In the original hand-coded taxonomy, only 2 terms related to specific personnel were identified in the subjects of sentences. In contrast, specific personnel were mentioned in half of the documents, indicating that the corpus as a whole makes much more mention of the specific actors who execute functions in HA/DR activities. This is an area of the data that could certainly be mined at a finer level of granularity in order to identify the specific types of individuals U.S. forces need to engage when involved in HA/DR activities.

2. What Do They Do?

What relief organizations do is mostly captured in the verb phrases of sentence predicates, although additional actions are captured in the objects of sentences (see below). This section focuses on the action verbs immediately following relief organization subjects. After augmenting the original taxonomy, the revised taxonomy has 632 verbs and verb phrases that captured 7496 action concepts, distributed across 15 first-tier and 26 second-tier categories.

Relief Organization Activities (Verbs)			
First Category	Second Category	Number of Documents	PCT Docs
Execution of Operations		61	92.4
Administrative Action	Initiation and Leadership	57	86.4
Supportive Action	Augmenting	56	84.8
Field Action	Provisioning	55	83.3
Financial Action		55	83.3
Supportive Action	Assistance	55	83.3
Field Action	Reconstruction	54	81.8
Administrative Action	Reporting	51	77.3
Education		51	77.3
Research	Observation	51	77.3
Administrative Action	Oversight	50	75.8
Research	Research	50	75.8
Field Action	Preparation	49	74.2
Political Action		48	72.7
Administrative Action	Meetings	45	68.2
Field Action	Operations Beginning	43	65.2
Field Action	Security Rescue	35	53.0
Administrative Action	Travel	29	43.9
Success		27	40.9
Supportive Action	Advisory	24	36.4

Field Action	Mitigation	23	34.8
Research	Forecasting	23	34.8
Administrative Action	Personnel	13	19.7
Field Action	Operations Ending	12	18.2
Supportive Action	Replacement	3	4.5
Translation		2	3.0

Figure 27. Lists the First and second tier categories of relief organization activities in order of the number of documents in which their concepts were mentioned.

The activity type mentioned in the most documents (61 of 66 or 92.4%) is execution of operations. These types of activities refer to the direct involvement of a relief organization in running HA/DR operations.

Other very common categories of relief organization activity included a mix of administrative activities, support activities, education, research and financial actions, and field actions. Administrative activities included initiation and leadership (86.4% of documents), reporting (77.3% of documents), and oversight (75.8% of documents). Support activities are closely allied with administrative functions, and included augmenting (84.8 % of documents) and assisting (83.3% of documents) types of activities. Educational activities, such as training, were mentioned in 77.3% of documents. Research activities were often mentioned, including observation (77.3% of documents) and general research activity (75.8% of documents). The two primary field action activities included provisioning (83.3% of documents) and reconstruction (81.8% of documents). Financial activities were mentioned in 83.3% of the documents.

These primary categories of activities alone demonstrate that relief organizations engage in a variety of activities that range from on-the-ground field operations to high-level administrative functions. While relief organizations report on a range of activities, administrative functions appear to dominate, accompanied by closely allied supportive functions and education. This paints a general picture of relief organizations as primarily facilitators, organizers and coordinators, although they are not strangers to actual field operations.

3. On Behalf of Whom?

The targets of relief organization operations help to define how these organizations see their missions, and most important, the human factor that presumably defines their *raison d'être*. The general term used to refer to the people relief organizations exist to help its beneficiaries. We developed a taxonomy of beneficiaries from the sample hand-coded corpus. Then the beneficiary taxonomy was augmented as described above, and the full corpus was analyzed with this augmented taxonomy.

Beneficiary	Number of Documents	PCT Documents
Community Village	63	95.5
Government	59	89.4
General Beneficiaries	58	87.9
NGO functionaries	51	77.3
Businesses	49	74.2
Women	45	68.2
Children	40	60.6
Students	17	25.8
Farmers	13	19.7
Journalists	5	7.6

Figure 28. Beneficiary table.

The most commonly cited types of beneficiaries, across documents, are communities and villages (95.5% of documents). This reflects the small-scale focus of relief efforts. However, it is notable that the smallest social scale that is most in relief organization documents is the community and village, and not the individual people who actually receive aid. The second most cited beneficiary type are governments (mentioned in 89.4% of documents), in particular the governments of nations where aid is needed. This underscores the supportive role relief organizations play; relief organizations very much exist to assist sovereign nations to aid their own people. There were many abstract beneficiaries that could not be clearly defined that were assigned to the category of General Beneficiaries, and so this category was not very informative. The next most mentioned beneficiaries cited across documents were NGO functionaries, such as medical aid workers, field staff, etc., and they were mentioned in 77.3% of the documents. It is interesting that they were not mentioned as actors, but as beneficiaries of relief organization activity, indicating that a substantial amount of relief organization activity is directed toward supporting their own personnel. Businesses were mentioned as beneficiaries in nearly ¾ of the documents, reflecting the economic development focus of many relief organizations. Women and children were each mentioned frequently (in 68.2% and 60.6% of documents respectively), indicating the focus on these sorts of individuals for relief efforts. More minor categories of beneficiaries included farmers and journalists. The latter, while not often mentioned, indicate some level of media savvy on behalf of relief organizations and the use of the media to assist in accomplishing the goals of monitoring hazards and disseminating information.

4. What Are the Objectives of the Relief Organization?

a) Objectives

Relief organizations have a wide range of objectives, as evidenced by the initial hand-coding of documents. These objectives range from the general concerns that motivate relief organizations (poverty, natural disaster relief) to actions the relief organization would like to see executed, to the beneficiaries the organization would like to help, to actual products the organization would like to

deliver. In this section, these objectives will be divided into concerns, things done, and things, in order to highlight the most important types of objectives relief organizations report they have.

b) Concerns

The concerns of a relief organization encompass the largest category of sentence objects, and after augmenting the original taxonomy, captured 4683 concepts distributed across 15 categories.

Category of Concerns	Number of Documents	PCT Documents
General Concerns	61	92.4
Early Warning Risk Reduction	53	80.3
Poverty	49	74.2
Violence	49	74.2
Education	44	66.7
Famine	42	63.6
Financial	40	60.6
Farming	37	56.1
Health	36	54.5
Children	30	45.5
Governmental Capability	24	36.4
Coordination	22	33.3
Gender	21	31.8
Herding	12	18.2
Aid Organization Capability	10	15.2

Figure 29. Category of Concerns.

The category of general concerns is really a catch-all of myriad, unrelated concerns, and is not very informative. However, the other categories of concerns were of more specific types and indicate the relative importance of concerns relief organizations have. The concern mentioned in most documents is early warning risk reduction, which is mentioned in 50, or 80.3% of the documents. Early warning risk reduction refers to integrated systems at the local, regional and national level that provide information inputs to anticipate risks such as social unrest, agrarian disasters and economic hazards. The frequent

mention of such systems indicates the prominence of prevention and early warning in current relief organization concerns in Africa. Although, US DoD Civil Affairs forces may be present in a region prior to a disaster, this type of concern is not a common mission activity for active duty Civil Affairs forces and should be further researched to provide knowledge and increased understanding for the potential development of Civil Affairs doctrine, organizational behavior, and training.

Other commonly mentioned types of concerns include poverty and violence, each mentioned in 49 or 74.2% of the documents. Education, famine financial concerns, farming and health were less often, but still frequently mentioned, occurring in over one-half to two-thirds of all documents. The beneficiaries that were most frequently mentioned across documents were children, and women are implied in the gender category of concerns.

c) Actions (Things Done)

Relief organization actions were not only mentioned as verbs, but often nominalized and mentioned in objects of sentences. These represent actions that concern the relief organization, often actions the organization would like to see carried out. Therefore, nominalized verbs provide another window on what relief organizations do.

Category of Object Actions (Things Done)	Number of Documents	PCT Docs
Assisting	59	89.4
General Actions	55	83.3
Education	52	78.8
Analysis	51	77.3
Planning	45	68.2
Prevention	41	62.1
Coordinating	39	59.1
Meetings	39	59.1
Political Action	38	57.6
Provisioning	35	53.0
Financial	21	31.8
Medical Interventions	18	27.3
Oversight	18	27.3
Presentation	14	21.2

Figure 30. Category of Object Actions.

The types of actions mentioned as objectives are very similar to the types of actions represented by verbs, although there is a difference in emphasis. As with verbs, the objects focus mostly on administrative activities. Supportive actions, such as assisting, planning and coordinating occur in 60% to 90% of documents. Assisting is the most common type of action mentioned, and occurs in nearly all documents. Field actions, such as provisioning and medical interventions are mentioned in fewer

documents (53% and 27.3% respectively). Other types of activities mentioned include financial, education, and political action. The more frequent mention of administrative functions, and less frequent mention of field actions may reflect the more administrative and facilitating function relief organizations carry out, as well as the administrative purpose for writing the documents in the corpus in the first place.

d) Things

Ultimately, material goods and services are often the actual things that preserve life and livelihood, and are hardly neglected in relief organization reports.

Category of Material Things	Number of Documents	PCT Docs
Documents	54	81.8
Food	52	78.8
Shelter	45	68.2
Other Things	39	59.1
Medical Supplies	22	33.3
Money	22	33.3
Agricultural Things	13	19.7
Educational Supplies	13	19.7

Figure 31. Category of Material Things.

The category of other things is simply a catch-all of largely un-related material goods, and is not very informative. Two types of things, food and shelter, were often mentioned (in 78.8% and 68.2% of the documents respectively), reflecting the basic nature of aid required by people in need. Medical supplies and money were mentioned in only a third of the documents, and agricultural goods and educational supplies were mentioned in slightly less than 20% of the documents.

It is revealing that, by far, the most commonly mentioned artifact in relief organization reports are other reports. Documents were mentioned in 54, or 81.8% of the documents. This probably reflects the largely administrative role played by relief organizations, as well as the purpose in writing the report (documentation of their activities to accounting agencies), and their audience (other report-writing and consuming organizations).

5. Sociocultural Variables in HA/DR

The role of sociocultural factors is a largely unexamined dimension of relief organization activity. In order to capture this dimension, we developed a sociocultural taxonomy, leveraged from previous DoD projects. The taxonomy is very broad, encompassing all aspects of human social life, belief, demography, history and environment. The result of our initial hand-coding of a sample of documents indicated that relief organization reports mention only a fraction of possible socio-cultural factors, and the text analytics of the full corpus indicated the same.

Reflecting the apparent economic focus of relief organizations, nearly all socio-cultural concepts mentioned fall under the category of Economy, Technology and Other Capabilities, and these are often mentioned; 64 of 66 documents, or 97.0% of all documents mention these kinds of variables. The next most mentioned sociocultural category dealt with demographic factors, which were mentioned in 51 of the 66 documents, or 77.3%. The final category of sociocultural phenomena that emerged from the corpus was interests, which includes beliefs, norms of behavior and social identity variables. Interests were mentioned in 22, or 1/3 of the documents in the corpus. Each of these major categories of sociocultural variables will be described separately.

The most often mentioned economic variables by document concern domestic economy (92.4% of documents), or essentially, how people make a living. By far the most frequently mentioned types of domestic economic variables concern agriculture, mentioned in 78.8% of the documents. Finance, income (generally of the family) and education were the next most often mentioned domestic economy variables, occurring in 50% to 60% of documents. Animal husbandry was mentioned in little more than a third of the documents. Economic risks were mentioned in only 22.7% of documents, which may indicate that the risks so prominent in the frequent mentions of early warning risk reduction are probably more concerned with social unrest and natural disasters. Considering the importance of income (60.6% of documents), it is surprising that labor is mentioned in only 15.2% of documents.

Macroeconomic variables (prices, GDP, nation-level economic indicator) were mentioned in 77.3% of the documents, indicating the importance of these indicators as measures of well-being and possibly relief activity effectiveness. Almost 40% of documents mentioned transportation, indicating some interest in infrastructure.

Sociocultural variables - Economy, Technology and Other Capabilities			
Level 1	Level 2	Number of Documents	PCT Docs
Infrastructure		64	97.0
Infrastructure::Domestic Economy		61	92.4
Infrastructure::Domestic Economy	Agriculture	52	78.8
Infrastructure::Domestic Economy	Income	40	60.6
Infrastructure::Domestic Economy	Education	34	51.5
Infrastructure::Domestic Economy	Finance	33	50.0
Infrastructure::Domestic Economy	Animal Husbandry	23	34.8
Infrastructure::Domestic Economy	Economic Risks	15	22.7
Infrastructure::Domestic Economy	Labor	10	15.2

Economy			
Infrastructure::Domestic Economy	Fishing	6	9.1
Infrastructure::Macroeconomics		51	77.3
Infrastructure::Transportation		26	39.4

Figure 32. Sociocultural variables – Economy, Technology and Other Capabilities.

Demographic variables were the other major category of sociocultural variables mentioned in the documents, being mentioned in over ¼ of the documents. The most common demographic variable mentioned in the corpus concerned disease and morbidity, reflecting the health focus of many relief organizations. Mortality and migration were both mentioned in about 1/3 of the documents. A minor category of demographic variable was fertility, which was mentioned in only about 6% of the documents.

Sociocultural Variables - Demography		
Category	Number of Documents	PCT Docs
Disease Morbidity	47	71.2
Migration	22	33.3
Mortality	21	31.8
Fertility	4	6.1

Figure 33. Sociocultural Variables - Demography.

The only other major socio-cultural category mentioned in the corpus was interests, although it is only ever mentioned in a minority of documents. Local, traditional norms of behavior were referred to in 28.8% of the documents, and constituted the largest sub-category of interests mentioned in the corpus. This indicates some sensitivity to local culture by relief organizations. There are reputed to be over 3000 ethnic groups on the African continent, speaking a like number of languages.³ It is striking that identity variables are mentioned only once in the corpus, and then only in reference to religious identity (Muslim). While this may be an artifact of how the original taxonomy was augmented, the fact that more variables like these did not come to light indicates that key social identity factors may be ignored by relief organizations.

³ http://www.africamissions.org/africa/ethnic_group.html. Heine, Bernd, eds (2000). *African Languages: an Introduction*. Cambridge University Press.

Sociocultural Variables - Interests			
Level 1	Level 2	Number of Documents	PCT Docs
Motivating Factors, Religion, Ideology		22	33.3
Motivating Factors, Religion, Ideology	Uncodified Norms	19	28.8
Motivating Factors, Religion, Ideology	Materialism	4	6.1
Social Identity	Religious Identity	1	1.5
Social Identity		1	1.5

Figure 34. Sociocultural Variables – Interests.

The final category of socio-cultural variable mentioned does not deal with people per se, but with how data are gathered on them. While only 15.2% of the documents mention social science methods, they are at least given some attention by relief organizations.

C. Operational Analysis

The first thing that is clear from an operational viewpoint is the granularity of the corpus is somewhat restrictive of detailed knowledge extraction. The documents that are found on the Internet typically do not contain tacit knowledge from the leaders and knowledge workers on the ground during HA/DR planning or mission execution. In order to have much more informed content for tacit knowledge extraction and analysis, a project would have to be designed to get access to organizations personal journals, blogs, letters home, and various other recorded content relevant to the CALICO objectives. We feel this could be accomplished for the U.S. Military via coordination for such activity to take place from organizational portals, secret classified accounts, and or military emails.

This Contractor team was to answer the question: “What do government and NGO reports on HA/DR activities tell us about the information needed to support effective interventions to ameliorate the human consequences of disruption and to enhance the effectiveness of indigenous governance?”

The relief organizations that participate in such responses seek to provide relief in the form of money, food, or clothing to the people who have been affected by the disaster, are vulnerable to future disasters, and or are very poor. In the understanding of the content within all of the corpus documents, we must highlight the importance of understanding the objectives and strategies of the disaster responders and their organization as this impacts the timing, level of effort for some disasters and not others, and the methods for financing such responses. Our analysis provides the following 7 types of relief activities that are typically conducted as Humanitarian Assistance/ Disaster Response operations conducted by a combination of military forces and nonmilitary organizations that combine one or more of the elements of humanitarian operations, which include one or more elements of other types of operations such as foreign humanitarian assistance, nation assistance, support to insurgency, or support to counterinsurgency.

The analyzed corpus provides 7 types of relief activity which are similar, yet not the same as the current U.S. Taxonomy for Humanitarian Assistance, which include:

- Development (14 documents)
- Security (15 documents)
- Natural Disaster (6 documents)
- Famine (8 documents)
- Health Care (3 documents)
- Education (1 documents)
- Multiple (16 documents)

Our findings point to organizations classified according to their predominant influence on their operations.

- NGOs
- United Nations
- Foreign Governments
- US Government, Non-military
- Personnel

This leads us to the potential hypothesis that NGOs are the predominant actor in a HA/DR response in Africa. In addition, we believe a potential second potential hypothesis of the findings are foreign governments contribute to more than 70% of the resources required for HA/DR response activities. While this statement of work did not call for the analysis of the stated hypothesis, we feel this would benefit Civil Affairs analysis process for pre-mission training specific to overseas operations in foreign cultures. As stated before, this type of information is not common for active duty Civil Affairs forces and should be further researched to provide knowledge and increased understanding for the potential development of Civil Affairs doctrine, organizational behavior, and training.

A trend that we would like to highlight is that the results from this study clearly identify elements of the Social Cultural Taxonomy that are not being addressed by current DoD military. The following list of beneficiaries is not targeted by the military response organizations due to the restrictions on the Overseas Humanitarian Disaster Assistance and Civic Aid (OHDACA) funds. Further research and understanding of this phenomenon would be beneficial to the project and to DoD Civil Affairs forces along with a few other small elements of related finding within this project.

Beneficiaries

- Community Village
- Women
- Children
- Businesses
- General beneficiaries
- Students
- Professionals

- Government
- NGO Functionaries
- Farmers
- Organizations
 - General Organizations
 - Civil Society Organizations
 - Governmental Organizations
 - NGOs
 - Business Sector

Relief organizations perform more “type of actions” than active duty Civil Affairs forces for many reasons. First, relief organizations are more adaptable to meeting the needs of the HA/DR situation due to their worldwide source for funding that is not tied to limiting congressional title 10 funds. Relief organizations have a dynamic ability to build starfish networks of HA/DR capabilities with international, regional, and local experts and build on these relationships over time by having continuous presence instead of rotating personnel every six to 12 months. This correlates to a temporal and organizational behavior element that is not equivalent in active duty Civil Affairs forces. Additionally, a relief organization response to HA/DR is highly correlated to the global awareness of the human situation via the US and international media and this awareness drives the funding levels available for HA/DR responses. While initially this does not seem problematic, it is suspected to create a level of bias that is not commonly understood by the casual observer. The bias is created from the relationship of the financial donors within the donor community, their agendas, and how the relief organization performs actions on the ground, which support the donors’ agenda. This drives the type of reporting that justifies meeting the donor’s agenda and potentially shaping the probability of getting more in-kind funding. This bias is not present in all actors and their response to HA/DR, but it is present and must be considered when analyzing the difference in relief organizations response from that of the DoD Civil Affairs forces. The key point from an operational perspective is that understanding the dynamic structure of the various relief organizations must be considered prior to any final correlations of their actions within a HA/DR situation.

VI. Final Recommendations

The key take away points of this study concern what the corpus did not tell us, and what it in fact represented. The discrepancy between our original expectations and final result suggests directions for future research, and further hypotheses to test.

A. What the Corpus Did Not Tell Us

The original objective of the study was to catalogue the socio-cultural factors and actions important to on the ground HA/DR operations executed by relief organizations. The corpus available to us within the scope of this study touched upon some of these factors and actions, but mostly focused on higher-level administrative functions. This could be a reflection of what relief organizations actually do, or a function of the purpose the relief organizations had in writing publically available reports. If reports rich in the ground recipes for HA/DR activity exist, they may not be publically available. It is also possible that such information exists informally as tacit knowledge in the minds and memories of practitioners, as part of a loosely defined tradecraft.

B. What the Corpus Did Tell Us

The corpus did provide genuine insights into what relief organizations report publically about their actions. That information did contain frequent references to field activities, and so the original objective of this project was partially accomplished. However, the main emphasis that we can extract from the language of the documents is clearly on administrative functions, and in particular, on those functions intended to facilitate and assist local and host nation governments in providing aid. More research is clearly necessary to uncover the ground truth on what relief organizations do, but the language in these publically available reports should not be ignored, since relief organizations may be honestly reporting what, in fact, the bulk of their activities comprise. In summary, these activities include administration, both supportive and more directly managerial, education, and financial. Field operations are mostly focused on provisioning goods and providing health care. These activities are mostly carried out in partnerships between relief organizations and host nation governments, with support from the United Nations and the United States.

The corpus also provided us with insight into what concerns and objectives relief organizations have. In addition to traditional and expected objectives of helping children, alleviating poverty and violence, there was a clear emphasis on developing early warning systems to anticipate when aid and relief would be needed.

Finally, it is clear that the sociocultural factors relief organizations report as important to their operations are heavily biased toward economic variables. Other key sociocultural variables such as social organization, belief systems, and social identity are seldom mentioned or not at all, despite the fact that an extensive taxonomy of these factors was used in the text analytics of the corpus. Of the economic factors mentioned in the reports, those concerning domestic and household economy, and those concerning macroeconomic indicators dominated. This may reflect the use of aggregate measures

of economic well-being to gauge the economic well-being of households, which are two different sociocultural scales and may not consistently track with one another.

C. Challenges

This project faced several challenges. First, obtaining a corpus of documents that reflect the activities carried out by relief organizations was challenging. Relief organizations publish reports, but these occur at a high level in their organizations, possibly not representing the on the ground or even majority of activities relief organizations execute. A further challenge is that relief organizations increasingly publish through websites, which present less coherent narratives, and more disjointed, hyperlinked, intertextually referenced documents. Exactly how to treat this newer form of reporting is unclear.

Another challenge is analysis of the corpus itself. The queries developed for this study produced voluminous data, only the surface of which has been touched. It is possible, for instance, to produce detailed networks demonstrating how specific actions are related to specific actors, objectives, beneficiaries and resources for each document. This, however, would result in thousands of individual stories; how researchers would mine those stories and search for generalizations is not immediately clear, although mining the corpus at hand has much potential.

D. Hypotheses and Future Directions for Research

Several hypotheses emerge from this research, mostly concerning the sociocultural context in which relief organization reports are generated.

The first set of hypotheses concern the alternative purposes for which the reports are written. Alternative purposes would include:

- Reflecting actual activities and functions carried out by relief organizations
- Internal reporting and institutional control
- Reporting to benefactors

Of course, these purposes need not be mutually exclusive.

Future research could take several logical directions. These directions could include:

- Working more directly with several relief organizations to obtain internal documents that may reflect their on the ground activities more accurately
- Conducting interviews with relief organizations in order to uncover the nature of their activities
- Further mining of the corpus in order to uncover details that have not been addressed because of the limited scope of this analysis

While this scope of work did not call for the operational analysis of the correlation of U.S. Humanitarian Assistance with the content found in this analysis, we understand that further refinement and linkage to U.S. taxonomies for Humanitarian Assistance (HA) doctrine would be a worthy effort to provide greater clarity of the similarities and differences in what the relief organizations performed on the ground as

compared to the relief organizations performance during HA/DR. This could result with analysis that would inform Civil Affairs doctrine and future HA/DR performance objectives.

E. Operational Recommendations

While we were limited in scope on this task, we would like to recommend some future activities that may provide value to the future development of Civil Affairs doctrine, Tactics, Techniques, and Procedures (TTPs), and analytical methods for cultural analysis during pre-mission train-up.

The first recommendation is to conduct a survey of Relief workers and Civil Affairs forces based on the knowledge gained from this project. This should then be developed into a quantitative analysis of the dependent and independent variables based on a set of hypothesis developed to inform Civil Affairs Cultural Analysis methods derived from CALICO. During this research, we recommend the collection of data that will allow for the analysis of time order sequencing of the variables i.e. “democracy creates urbanization or urbanization creates democracy” or “ the presence of local NGOs in a host country occurs before disaster or host country disasters drive the presence of local NGOs.”

In addition, we recommend the external comparison of this research result with that of the previous research on Civil Affairs language used in doctrine. This could provide a sophisticated method for conducting a GAP analysis of the Civil Affairs organizational capability or simply serve as a validity check for the existing Civil Affairs doctrine, education and training methodologies currently in place.

For further development of this work, we recommend a more detailed analysis of these findings as they relate to location of the HA/DR response. This would be further specified as the geospatial and temporal analysis of HA/DR responses by actor and by action for the timeframe of 2000 -2011. In this analysis, we recommend comparing variables related to the African country specific UN Human Development Index (HDI), African Development Indicators, against the results of this research. In addition, we believe further visualization of this data with other socio cultural data i.e. Ethnicities, religion, political boundaries, economic (land use) will allow the CALICO results to better inform DoD culturally relevant operations. We envision this to further identify and describe the key factors that influence future responses to HA/DR as a function of host nation capacity. The language and taxonomy identified and used in this research will be further shaped with increased understanding of the reasoning associated with HA/DR responses to vulnerable populations as a function of host nation capacity. This future research will inform multiple aspects of the Doctrine, Organization, Training, Material, Leadership, and Professional components of the US Army and Marine Corps Civil Affairs forces.

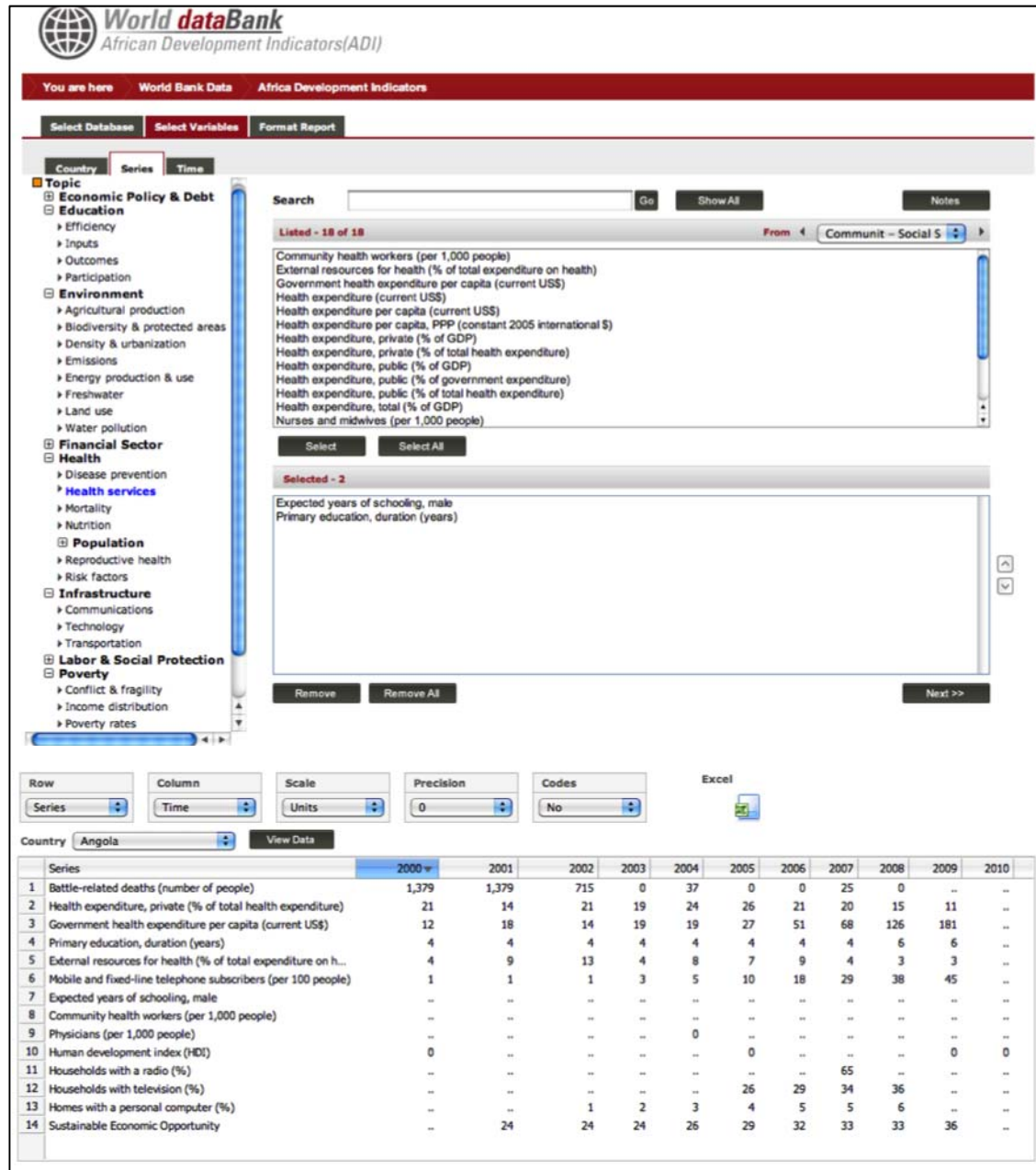


Figure 35. Recommended data types for future CALICO research to understand HA/DR indicators relevant to HA/DR responses (Source: World databank.com).

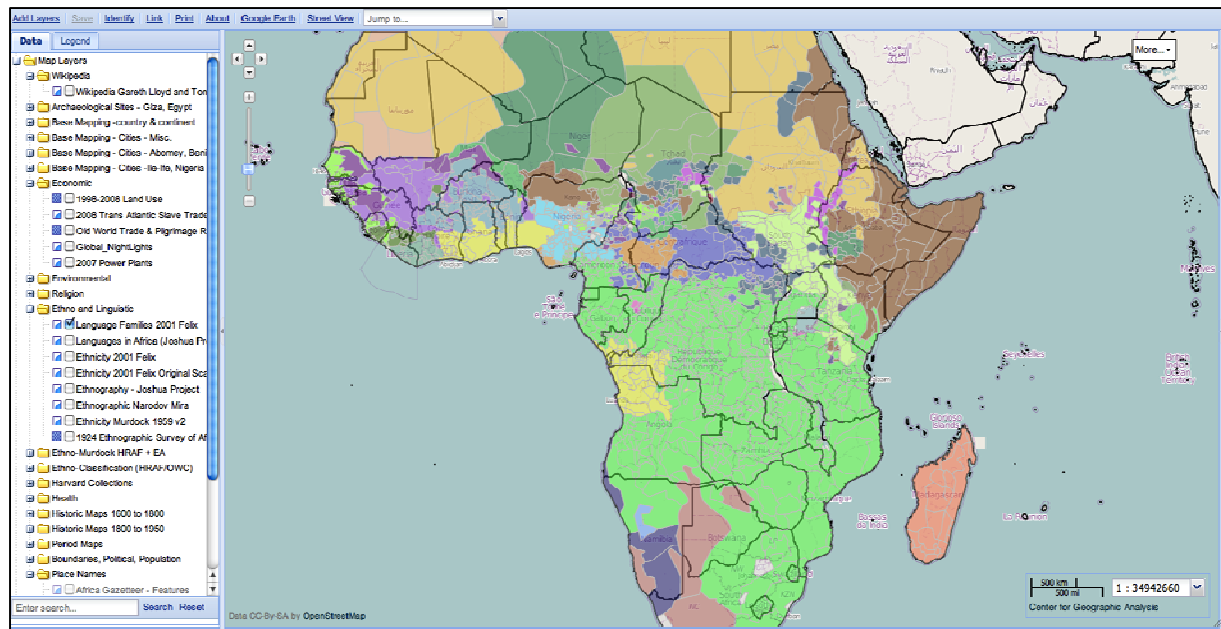


Figure 36. Multilayered Geospatial Analysis for visualizing research findings (source: worldmap.harvard.edu/africamap/).

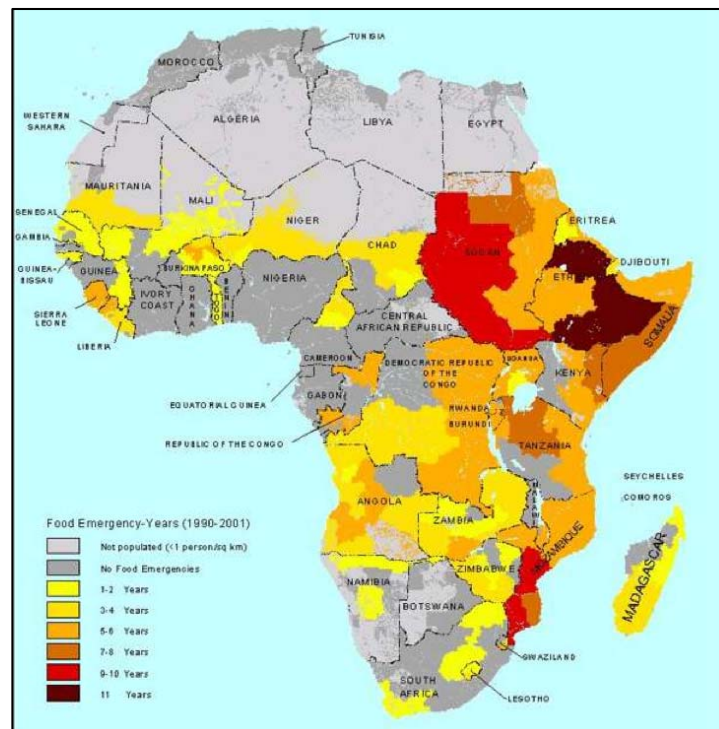


Figure 37. Example of Geospatial Mapping of food emergencies in Africa (source isciences.com).

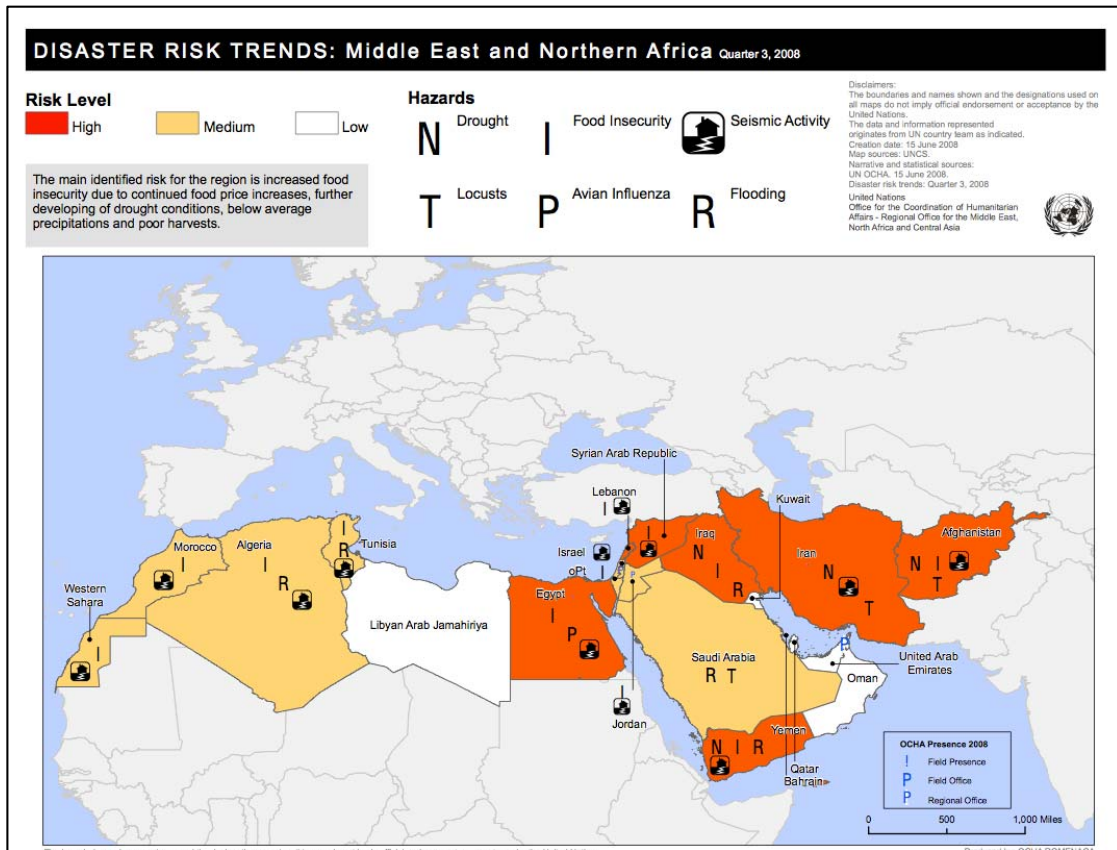


Figure 38. Example of Geospatial Mapping of Disaster Trends (source: ochaonline.un.org/MapCentre/).

VII. Appendix I: Evaluation of MAXQDA Qualitative Analysis Software for Coding and Taxonomy Generation

MAXQDA 10 is a text analysis tool that facilitates qualitative data analysis. This usage of the MAXQDA 10 software is to assess the compatibility and usability of the software in the CALICO project. At the time when MAXQDA 10 was acquired, a taxonomy for the entity extraction had already been created. This taxonomy was created using hard copies of the documents and was done by hand. MAXQDA 10 could have been used to create the taxonomy in a similar way, albeit electronically. Intrinsically, neither method, either by hand or MAXQDA 10, would have been significantly more advantageous. Both methods would have had methodological benefits and disadvantages. Subsequently, MAXQDA 10 was then used to apply the existing taxonomy to the CALICO documents.

A. MAXQDA 10 Training

The MAXQDA website⁴ offers multiple resources for training and assistance with the software. Two of the main resources are the MAXQDA 10 Reference Manual and Introduction PDFs, which are both available for download. The 91-page Introduction is a must-read for a self-guided student of the software. The Reference Manual, at 271 pages is not a front-to-back read but will be a fundamental source for a new user.

The MAXQDA provider partnered with a University of Florida anthropology professor to provide step-by-step video tutorials. According to the website, there will eventually be a 10-video series. However, the tutorial series is not complete- only four videos have been created and posted- and has not been updated in over a year. Currently, the series covers a basic introduction to MAXQDA functions including how to import texts, manage text attributes, and coding basics. For what it's worth, the existing video tutorials are helpful, step-by-step instructions for using MAXQDA.

The MAXQDA provider has also partnered with a U.S. private contractor to offer instruction and training at varying price points, levels of knowledge, and amount of time investment. Currently, the contractor offers a 1-hour free live webinar reviewing basic MAXQDA functions and possible applications.

A new user is recommended to “play” with MAXQDA 10. Working with one’s own real texts assists a user to become more accustomed to the MAXQDA interface and possible capabilities. Overall, learning MAXQDA 10 is not overly difficult. Ease of use increases as one is exposed to MAXQDA 10 over time. The ability to use and analyze data will be dependent on a user’s previous experience with content analysis and comfort level with software. As with other software, a new user will be limited in scope and probably complexity of analysis as analysis can be limited by knowledge of software capability.

⁴ <http://www.maxqda.com/>

Task	Approximate Time Required
Video Tutorials & Introductory Webinar	2 hours
Review MAXQDA 10 Manuals	1-2 hours & as needed
Familiarize with Basic Functions & Layout	1-2 hours

Figure 39. MAXQDA Training Time Requirements.

B. Coding Process

MAXQDA initial set-up is very straightforward and is not time intensive. MAXQDA 10 is self-contained, meaning that it keeps all of the files together within one folder. A user needs only to import his or her documents (.doc, .docx, .pdf, .rtf, etc) into MAXQDA's Document System. As long as the user is able to browse to the files easily, the user can import the files instantaneously.

The coding system can be entered by hand into the software's Code System function. MAXQDA 10 can work with the different file types. Older PDF versions (e.g. 2001's 1.4) seem a little clunky in MAXQDA 10. The newer versions respond faster to highlighting and coding. However, any difference is slight and does not affect use to any significant degree.

Inputting the coding system can be tedious. However, simple or straightforward coding systems are easy to handle. Larger, more complex systems require more attention to detail. MAXQDA 10 allows for hierarchical coding system so the user must maintain consistency during input. If a code is missed, the user can enter a new code in situ.

Using the text extraction taxonomies, two to three hours per document is required. However, this time estimation could be reduced with sufficient code system training. Additionally, a code system that is based on text segments larger than a word or words would decrease time. A code system made for automated software is not entirely appropriate for a human coder, even one using content analysis software. The act of coding a particular word, phrase, or sentence is simple in MAXQDA 10. If the code is already entered, the user simply highlights the word or phrase and applies the code via dragging and dropping the appropriate code (code to the highlighted phrase or vice versa). MAXQDA 10 allows multiple coding of words and phrases. Additionally, through its ability to "activate" and "deactivate" particular codes, an analyst can view the codes of interest without being hampered by extraneous or unrelated coding.

Task	Approximate Time Required
Creating New Project	Instantaneous
Import Documents	Instantaneous
Enter Code System	1 hour <i>But overall, this varies and depends on length of code system.</i>
Coding one CALICO document⁵	2-3 hours

Figure 40. MAXQDA User Time Requirements.

⁴ Using text extraction taxonomy

C. Analyzing Data Using MAXQDA 10

Instead of using MAXQDA 10 to create the taxonomies, they were instead applied to the CALICO documents to evaluate the software. Once the taxonomies were entered into MAXQDA's coding system, and the documents were coded, the software allows the user to "Retrieve coded segments"- "coded segments" refers to the coded words and phrases within the document.

The software's main window includes its Retrieved Segments pane that shows the coded words and the corresponding document and applied codes (**Figure 41**).

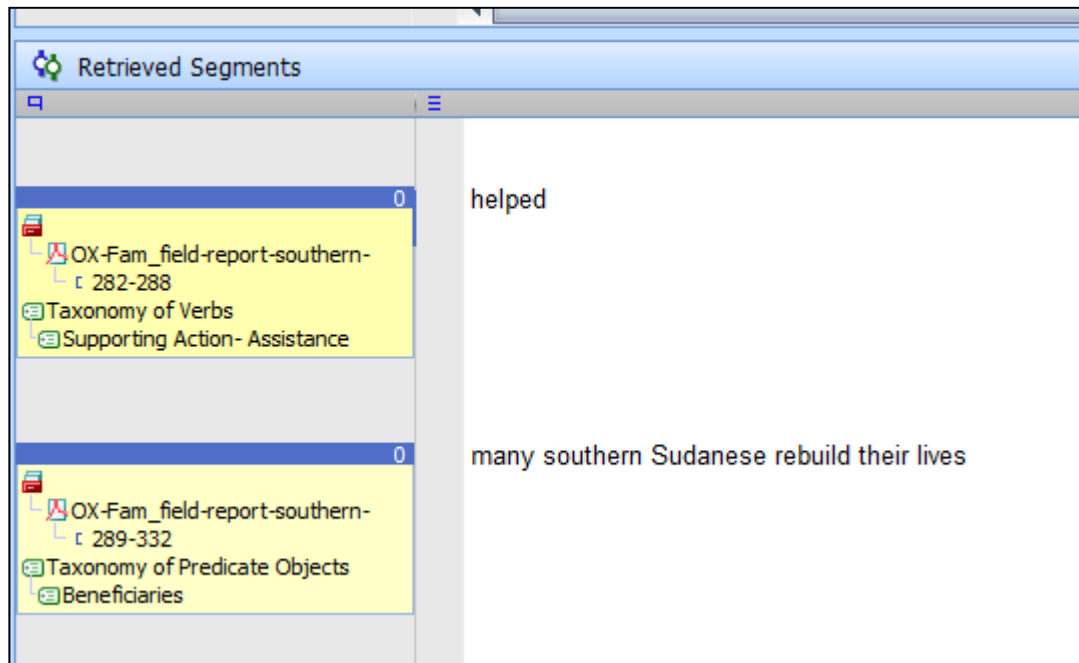


Figure 41. Example of MAXQDA 10 Retrieved Segments pane

Additionally, the Retrieved Segments can be displayed in spreadsheet form (**Figure 42**). The user can export this data to Excel as well as to an HTML spreadsheet. The user is able to control what data to work with at a time. For example, the user could work with a particular code within one document or compare that code across all documents simultaneously.

	Comment	Document	Code	Weight score	Preview	Doc
○		OX-Fam_field-report-south...	Taxonomy of ...	0	in its humanitarian assistance programs ...	
○		OX-Fam_field-report-south...	Relief Organiz...	0	the US	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	saved	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	hundreds of thousands of lives by pro...	
○		OX-Fam_field-report-south...	Relief Organiz...	0	the US	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	helped	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	spearhead the UN's OLS, a groundbreaki...	
○		OX-Fam_field-report-south...	Relief Organiz...	0	OLS	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	delivered	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	much-needed vaccines and foodstuffs ...	
○		OX-Fam_field-report-south...	Relief Organiz...	0	USAID	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	worked to expand	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	the mandate of its relief programs to i...	
○		OX-Fam_field-report-south...	Relief Organiz...	0	US	
○		OX-Fam_field-report-south...	Taxonomy of ...	0	help	

Figure 42. Example of MAXQDA 10 Retrieved Segments spreadsheet

MAXQDA 10 has several visualization tools that can aid an analyst. The Document Portrait tool displays a representation of the codes within a document according to the analyst-assigned color codes. This is a quick visualization of how often and where, within a document, particular codes appear and don't appear.

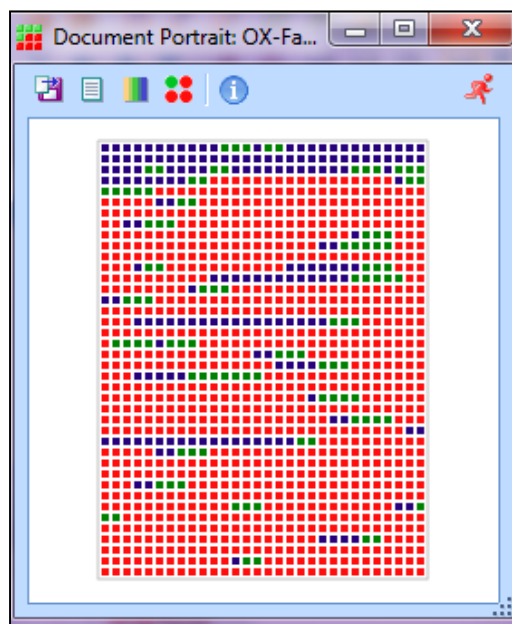


Figure 43. Example of a Document Portrait

The Codeline visualization tool can display what codes occur within the numbered paragraphs within a document. However, this function does not seem to work if the software cannot apply paragraph breaks to the document (in order to apply paragraph numbers) during import. Unfortunately, it does

not appear that MAXQDA 10 is able to apply paragraph breaks to PDFs and the CALICO documents that were tested were in PDF form.

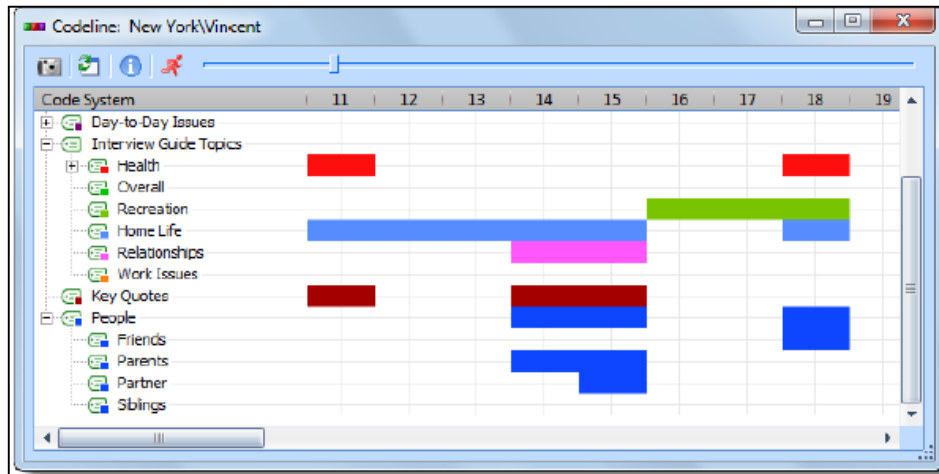


Figure 44. Example of a Codeline Visualization (taken from the MAXQDA 10 Reference Guide).

However, if possible, this function is a helpful visualization tool of applied codes. An analyst can see where codes occur in the document and how often, and what codes occur together.

MAXQDA 10 also has basic word frequency functions. In **Figure 45**, MAXQDA 10 creates a Tag Cloud from one document and generates a word frequency list.

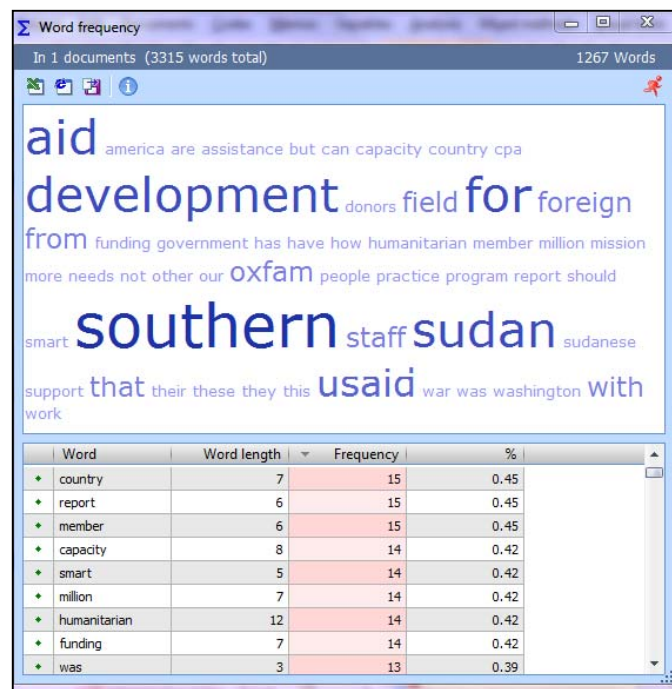


Figure 45. Example of a Tag Cloud and Word Frequency List

D. Comparison with Hand-Coding

The traditional approach to coding documents is to do so by hand, often marking words and phrases with color-coded markers, and then entering coded material into some sort of database. In this section, we address the advantages and disadvantages of hand-coding with the use of MAXQDA.

The main steps in hand-coding include: development of taxonomy, familiarization with taxonomy, the actual coding, and database entry. Time estimates for these activities are found in **figure 46**.

Task	Approximate Time Required
Taxonomy Development	5 – 10 hours
Familiarization with Taxonomy	4 – 8 hours
Coding	2-3 hours per document
Database Entry	1-2 hours per document

Figure 46. User Time Requirements.

Taxonomy development is a conceptual task that must be performed regardless of the coding method used. It is very difficult to estimate how long this process would take, since it is contingent upon the taxonomists' level of expertise and the size and complexity of the taxonomy under development. The use of software such as MAXQDA may speed this process by providing a more easily searchable electronic format for locating entities that the taxonomist must decide to classify. MAXQDA was not used for this purpose in this project, and so any quantitative comparison is not possible.

Familiarization with the taxonomy is another conceptual task that does not require a computer, and is also contingent upon a coder's understanding of the domain and the size and complexity of the taxonomy. The 4 – 8 hours estimate we offer here is based on the assumption that the coder has a background in the domain being coded.

It took 2-3 hours per document to code the texts in the sample, consistent with the estimates from using MAXQDA. The software adds no time to this task, and may speed it up marginally, due to the ease of paging through electronic documents versus hand-held ones.

Database entry is an area of potential timesaving by using qualitative analysis software such as MAXQDA. This process was extremely laborious when done by hand, taking at least 1-2 hours per document. Since the data are already coded in MAXQDA, this process is more or less instantaneous. The time savings in this particular project would, therefore have been at least 12 hours, or approximately two (2) person-days of labor.

MAXQDA provides data analytic functions that would be difficult and prohibitively time consuming, such as word frequency counts, so that the overall assessment is that MAXQDA is at least a meaningfully time-saving tool, and in the case of basic text analysis, possibly indispensable.

E. Advantages & Disadvantages of MAXQDA 10

1. Advantages

The MAXQDA 10 software keeps all project files together; avoiding the possible mishandling of data, different versions of documents, and overall organization.

Metadata on the documents can be used to add further dimensions to an analysis. For example, particular codes could be displayed only from documents relating to southern Sudan between 1998 and 2008.

MAXQDA 10 allows for what it calls in-vivo coding; in other words, creating new codes within a document. A word or phrase can be highlighted and a user can create a new code and apply it to this piece of text with one function. This allows a user to use MAXQDA 10 to keep track of relevant words and phrases and then use this data to create a taxonomy or coding system electronically rather than by hand.

The visualization tools and the creation of spreadsheets of the coded text in and of itself are advantageous and useful. Hand-coding documents is already a time-consuming task and upon completion, the analyst still must go back through the coded documents and create (either by hand or using the computer) an organized, and preferably searchable, list of the coded text. MAXQDA 10 creates this list during the initial coding process. The analyst then has the ability to use this spreadsheet of the data as needed thus saving time and eliminating repetitious work.

The comparison of hand-coding and MAXQDA labor requirements indicates that overall, MAXQDA provides at least some time-savings. However, if text analysis is to be done, MAXQDA provides basic capabilities that are prohibitively expensive to do by hand.

2. Disadvantages

A taxonomy of words (verbs and nouns) is not the best code system to be applied using MAXQDA 10 software. While the set-up of the taxonomy within MAXQDA 10 is straightforward, the act of coding can be cumbersome and time-consuming given a word search-like taxonomy.

As with any software, the user is limited by his or her knowledge of how to use the software, as well as possible limitations of the software itself. For example, MAXQDA 10 does not easily allow a user to highlight previously coded segments within the text. So when multiple codes are applied within a paragraph, the “coding stripe” can become quite full (see the grey box on the left side of **Figure 47**). There are nine codes applied to the paragraph. However, to see what text the codes were applied to, the user must click on the individual code in the coding stripe.

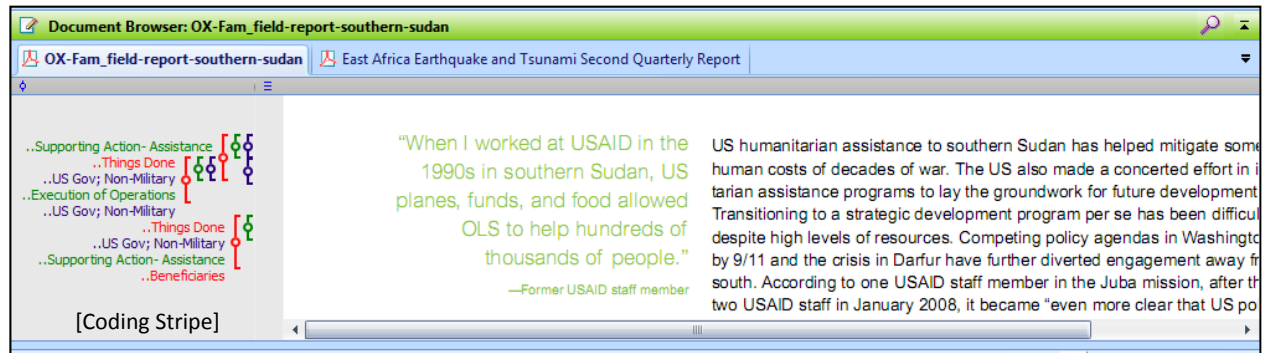


Figure 47. Example of Codes Applied in a Document

MAXQDA 10 supposedly has the option of highlighting coded sections of the text; however, this function does not seem to be reliable. The user can reliably highlight one coded section of the text at a time but no more (Figure 48).

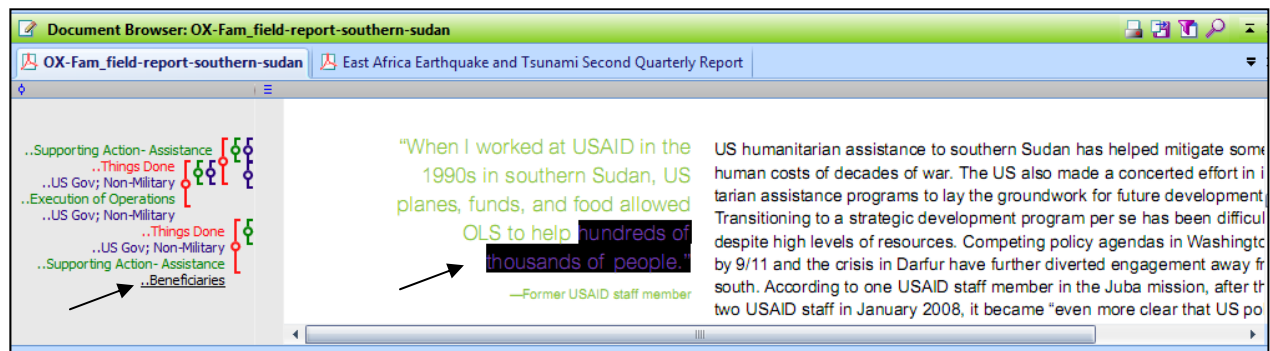


Figure 48. Single Code Highlighted in Text

Use of MAXQDA 10 will be limited to a user's knowledge and familiarity with the software. This is probably the main limiting factor. Once the user has gained a sense of control over the software, the existing software limitations may be easily bypassed or solved in another manner.

F. References

MAXQDA, Software for Qualitative Data Analysis. (1989-2011). VERBI Software. Consult. Sozialforschung GmbH, Berlin-Marburg-Amöneburg, Germany.

VERBI Software. (2011). *Reference manual for the text analysis software MAX QualitativeDataAnalysis 10: For Windows 2000, XP, Vista, and 7*. VERBI Software. Consult. Sozialforschung GmbH, Berlin-Marburg-Amöneburg, Germany.

VIII. Appendix II: Taxonomy of Relief Organizations in Subjects

Taxonomy of Relief Organizations

NGO

african network for prevention and protection of children against abuse and neglect
 anppcan
 africare's system of rice intensification (sri)
 sri
 better u foundation
 child rights ngo network
 crnn
 civil society
 civil society organizations
 civil society organization
 committee
 committees
 civil society organization
 cso
 delegations
 delegation
 donor agencies
 donor agency
 donor community
 red cross
 red crescent
 red cross red crescent
 rrcr
 east africa regional delegation
 federation
 federation dm officer
 federation of uganda employers
 federation secretariat working group
 german red cross
 high level tsunami group
 hltg
 hope after rape
 humanitarian partners
 humanitarian partner
 inter-governmental authority on government
 igad
 hospitals
 hospital
 kenya red cross
 mobil teams
 mobile team

national societies
national society headquarters
national union of plantation and agricultural workers (nupawu)
non-governmental organizations
non-governmental organization
ngo
ngos
non-state actors
non-state actor
nairobi peace initiative
npi-africa
ocbo
orphans community based organization
oxfam
oxfam america
partners
partner
partnership for peace
red cross puntland
red cross red crescent movement partners
red cross red crescent movement partner
red cross society of eritrea
save the children uk
scuk
save the children
seychelles red cross society
srcs
slum aid project
somali red crescent clinics
somali red crescent clinic
somali red crescent society
southern sudan agriculture revitalization program
staff of somaliland
stakeholders
stakeholder
sudan microfinance institution
sumi
system of rice intensification
sri
targeted supplementary feeding (tsf) programme
the operation
the project
tsunami steering committee
ucrnn
uganda child rights ngo network
uganda society for disabled children
usdc

uganda women lawyer association
 uganda youth development link
 village irrigated perimeters
 vips
 wash
 water, sanitation and hygiene

United Nations

unicef
 united nations children's fund
 wfp
 wfp hubs and spokes program
 world food program

Foreign Government

child labour unit
 child-based alternative non-formal education project (chance)
 disaster risk management and food security sector
 drmfss
 disaster risk management technical working group
 drmtwg
 dm department
 dm programmer
 early warning and response directorate
 erd
 ethiopian health and nutrition research institute
 ehni
 ethiopian humanitarian country team
 emergency health and nutrition taskforce
 empowering life long skills education
 else
 enhanced outreach strategy
 eos
 federal level emergency wash task force
 federal ministry of health
 fmoh
 federal wash task force
 fida (u)
 field extension agents
 food management and improvement project
 fmip
 government
 government agent
 health centers
 health center

health extension workers
hew
health posts
health post
health system
law reform commission
mfped
mglsd
ministries
ministry of agriculture (moa)
ministry of agriculture, animal industries and fisheries (maif)
ministry of finance, planning and economic development
ministry of gender, labour and social development
ministry of health
ministry of rural development and water resources (mrdwr)
ministry of water and energy
mobile health and nutrition teams
mhnts
mobile health and nutrition team
mhnt
mobilizing early response project (merp)
merp
mubende non-formal education
mnfe
multi-agency coordination (mac) group
national disaster prevention and preparedness committee
ndppc
office of the prime minister
opm
participatory poverty assessment program
police
regional disaster prevention and preparedness bureaus
dppbs
dppb
regional education bureaus
regional education bureaus
relevant agencies
relevant agency
reprioritization and reallocation of resources group
rrrg
sectoral taskforces
sectoral taskforce
the state

US Government, Non-Military

- us
- u.s.
- united states
- usaid
- usaid field mission

Personnel

- aid practitioners
- aid practitioner

IX. Appendix III: Taxonomy of Predicate Verbs

Taxonomy of Verbs (513)

Administrative Action (116)

- Initiation and Leadership (41)
- Oversight (8)
- Assessment (11)
- Reporting (20)
- Meetings (19)
- Travel (13)
- Personnel (4)

Execution of Operations (87)

Supporting Action (115)

- Augmenting (29)
- Assistance (62)
- Advisory (16)
- Replacement (8)

Field Action (94)

- Distribution (22)
- Reconstruction (12)
- Mitigation (16)
- Security Rescue (8)
- Preparation (16)
- Operations Beginning (12)
- Operations Ending (8)

Financial Action (4)

Political Action (25)

Success (8)

Observation (20)

Research (24)

Forecasting (8)

Education (8)

Translation (4)

Administrative Action – Initiation and Leadership

- begun
- began
- beginning
- begin
- formed
- forms
- form
- forming
- piloted
- pilot

pilots
 piloting
 planning
 plan
 plans
 planned
 planning to
 plan to
 plans to
 planned to
 lay
 established
 establish
 establishes
 establishing
 spearhead
 spearheads
 spearheaded
 spearheading
 started
 start
 starts
 starting
 initiate
 initiates
 initiated
 initiating
 introduced
 introduce
 introduces
 introducing

Administrative Action – Oversight

oversee
 oversees
 overseeing
 oversaw
 supervised
 supervise
 supervises
 supervising

Administrative Action - Assessment

assess
 assessing
 assesses
 review

reviews
 reviewed
 reviewing
 examines
 examine
 examining
 examined

Administrative Action – Reporting

presented
 present
 presents
 presenting
 released
 release
 releases
 releasing
 reporting on
 report on
 reports on
 reported on
 report
 reports
 reported
 reporting
 submit
 submits
 submitted
 submitting

Administrative Action – Meetings

attended
 attend
 attends
 attending
 chaired
 chair
 chairing
 hosted
 hosts
 host
 hosting
 held
 hold
 holds
 holding
 scheduled

schedule
schedules
scheduling

Administrative Action – Travel

travel
travels
travelling
travelled
visited
visit
visits
visiting
went to
go to
goes to
going to

Administrative Action - Personnel

recruit
recruits
recruited
recruiting

Execution of Operations

carried out
carrying out
carries out
carry out
handling
handles
handle
handled
operates
operate
operated
operating
making an effort
make and effort
makes and effort
made and effort
managing
manage
manages
managed
conduct
conducts

conducted
 conducting
 control
 controls
 controlled
 controlling
 covered
 cover
 covers
 covering
 diversify
 diversifies
 diversified
 diversifying
 engage
 engages
 engaged
 engaging
 tackle
 tackles
 tackled
 tackling
 addressing
 address
 addresses
 addressed
 adopted
 adopt
 adopting
 adopts
 requested
 request
 requests
 requesting
 select
 selects
 selected
 selecting
 put emphasis on
 puts emphasis on
 putting emphasis on
 decide
 deciding
 decided
 develop
 developed
 developing

offered
 offers
 offer
 offering
 organized
 organize
 organizing
 organized
 running
 run
 runs
 ran
 coordinated
 coordinate
 coordinating
 engages
 engaged
 engaging

Supportive Action – Augmenting

improve
 improves
 improved
 improving
 increase
 increases
 increased
 increasing
 revitalize
 revitalizes
 revitalized
 revitalizing
 boosting
 boost
 broaden
 broadens
 broadening
 enhancing
 enhance
 enhances
 enhanced
 expand
 expands
 expanded
 expanding
 modify
 modifies

modified
modifying

Supportive Action – Assistance

complimented
compliment
compliments
complimenting
helped
help
helps
helping
work in collaboration with
works in collaboration with
worked in collaboration with
working in collaboration with
work with
works with
worked with
working with
working in partnership with
work in partnership with
works in partnership with
worked in partnership with
supported
support
supports
supporting
contributed
contributes
contribute
contributing
empower
empowered
empowering
encouraged
encourage
encouraging
ensure
ensures
ensured
ensuring
insure
insures
insured
insuring
participated

participate
 participates
 participating
 play
 plays
 played
 playing
 played a role
 play a role
 plays a role
 playing a role
 take part
 takes part
 took part
 taking part
 participate in planning
 participates in planning
 participated in planning
 participating in planning

Supportive Action - Advisory

advise
 advises
 advised
 advising
 consulted
 consult
 consults
 consulting
 facilitating
 facilitate
 facilitates
 facilitated
 made recommendations
 make recommendations
 making recommendations
 makes recommendations

Supportive Action - Replacement

fill in
 fills in
 filled in
 filling in
 intervene
 intervenes
 intervened
 intervening

Field Action – Distribution

dispatched
 dispatch
 dispatches
 dispatching
 distributed
 distribute
 distributes
 distributing
 provided
 provide
 provides
 providing
 supplied
 supply
 supplies
 supplying
 dissemination of
 disseminate
 disseminating
 disseminates
 disseminated

Field Action – Reconstruction

rehabilitate
 rehabilitates
 rehabilitated
 rehabilitating
 renovation
 renovate
 renovates
 renovating
 building
 build
 builds
 built

Field Action – Mitigation

dealing with
 deal with
 deals with
 dealt with
 mitigate
 mitigates
 mitigated
 mitigating

minimizing
 minimize
 minimizes
 minimized
 preventing
 prevent
 prevents
 prevented

Field Action – Security Rescue

protect
 protects
 protected
 protecting
 saved
 save
 saves
 saving

Field Action - Preparation

pre-positioned
 pre-position
 pre-positions
 pre-positioning
 procured
 procure
 procures
 procuring
 screen
 screens
 screened
 screening
 targeting
 target
 targets
 targeted

Field Action – Operations Beginning

implement
 implements
 implemented
 implementing
 mobilized
 mobilized
 mobilizes
 mobilizing
 responding

respond
 responds
 responded

Field Action – Operations Ending

withdrawing
 withdraw
 withdraws
 withdrew
 end
 ends
 ended
 ending

Financial Action

fund
 funds
 funded
 funding

Political Action

advocates
 advocate
 advocated
 advocating
 push for
 pushes for
 pushed for
 pushing for
 raised
 raise
 raises
 raising
 ratified
 ratify
 ratifies
 ratifying
 urges
 urge
 urges
 urged
 urging
 promote
 promotes
 promoted
 promoting

Success

achieves
 achieve
 achieved
 achieving
 alleviate
 alleviates
 alleviated
 alleviating

Observation

identified
 identify
 identifies
 identifying
 recognized
 recognize
 recognized
 recognizing
 found
 find
 finds
 finding
 gather
 gathers
 gathered
 gathering
 monitoring
 monitor
 monitors
 monitored

Research

experimented
 experiment
 experiments
 experimenting
 investigate
 investigates
 investigated
 investigating
 map
 maps
 mapped
 mapping
 research
 researches

researches
researching
surveyed
survey
surveys
surveying
explores
explore
explored
exploring

Forecasting

anticipates
anticipate
anticipated
anticipating
forecast
forecasts
forecasted
forecasting

Education

training
train
trains
trained
teach
teaches
taught
teaching

Translation

translated
translate
translates
translating

X. Appendix IV: Taxonomy of Objects in Predicates

Taxonomy of Predicate Objects Categorized

Beneficiaries

Organizations

Things Done

Concerns

Things

Places

Beneficiaries (59)

Community Village (5)

Women (6)

Children (11)

Businesses (3)

General beneficiaries (10)

Students (3)

Professionals (2)

Government (3)

NGO Functionaries (9)

Farmers (7)

Organizations (64)

General Organizations (21)

Civil Society Organizations (6)

Governmental Organizations (33)

NGOs (2)

Business Sector (2)

Things Done (90)

Coordinating (7)

Assisting (5)

Oversight (5)

General Actions (14)

Education (4)

Provisioning (9)

Analysis (6)

Presentation (2)

Planning (3)

Prevention (2)

Political Action (2)

Meetings (11)

Medical Interventions (10)

Technical Support (5)

Financial (5)

Concerns (145)

- Children (6)
- Gender (2)
- Violence (11)
- Famine (9)
- Poverty (8)
- Aid Organization Capability (4)
- Government Capability (9)
- Coordination (16)
- Financial (4)
- Health (22)
- Farming (16)
- Herding (3)
- Early Warning Risk Reduction (18)
- Education (2)
- General Concerns (39)

Things (54)

- Documents (14)
- Medical Supplies (14)
- Food (6)
- Shelter (5)
- Educational Supplies (4)
- Money (3)
- Agricultural Things (2)
- Other Things (6)

Places (50)

Beneficiaries

- Community Village
- community members
- community
- communities
- villages
- village

- Women
- women
- females
- abused women
- sexually abused women
- rape victims

victims of rape

Children

children

child

street children

working children

working child

child sex workers

child sex worker

sexually abused children

sexually abused child

malnourished children

malnourished child

Businesses

individual retail institutions

retail institutions

retail institution

General Beneficiaries

southern sudanese

sudanese

members

member

the poor

people

disaster-affected populations

disaster-affected population

participants

beneficiaries

Students

students

learners

learner

Professionals

journalists

journalist

Government

officials

the official

an official

NGO functionaries

medical units
 medical unit
 health staff
 mobile teams
 mobile team
 team
 teams
 appropriate staff
 staff

Farmers
 small hold farmers
 small hold farmer
 rural smallholder farmers
 timbuktu smallholders
 smallholders
 sri participants
 lead farmers

Organizations

General Organizations
 development project
 committees
 committee
 child-led project
 partnerships
 partnership
 community-based first aid (cbfa) working group
 community-based first aid (cbfa)
 cbfa
 community-based first aid
 focal persons
 focal person
 cbfa programme
 cbfa program
 teams
 team
 local research stations
 local research station
 associations
 an association
 the association

Civil Society Organizations
 civil society organizations
 csos
 civil society organization

CSO

civil society networks

national societies

Government Organizations

local government programmes

local government programs

local government program

government programmes

government program

government programs

nutrition and early childhood development project

government

j pilot programme

poverty monitoring unit

districts

district

ministry of health

moh

ministry of education

moe

ministry of gender, labour and social development

mglsd

aids information center

aic

nationalized ewer program

nationalized early warning early response program

ewer program

early warning early response program

community development centers

community development center

policy makers

out patient therapeutic programme (otp)

federal ministry of health

fmoh

ethiopian health and nutrition research institute (ehnri) of public emergency management center

ethiopian health and nutrition research institute (ehnri)

ethiopian health and nutrition research institute

NGOs

somali red crescent

emergency health and nutrition taskforce

Business Sector

sumi

microfinance sector

Things Done (Nominal)

Coordinating
co-ordinated approaches
co-ordinated approach
coordinated approaches
coordinated approach
cross-sectoral coordination
sharing of early warning information
coordination

Assisting
participation
support
monitoring and support services
expand and broaden its cbfa programme
assistance

Oversight
revision
evaluation of the gaps in water and sanitation activities
needs assessment
evaluate sri methodology's application on other crops
case management

General Actions
government's modest efforts
government's modest effort
modest efforts
modest effort
effort
efforts
exercises
consultations to children under 5
consultations
surveillance
implementation of set strategies
early warning information
response plans
guidance

Education
child rights education, psychosocial support, direct legal support
education
training
a workshop on sexual reproductive health

Provisioning

role in providing for the orphans' daily needs
distributions of insecticide treated nets (itns)
insecticide treated nets (itns)
insecticide treated nets
itns
itn
spraying of ddt
storage, transportation and distribution of relief food in the region
distribution of relief food

Analysis

situation analysis of orphans
situation analysis
research
focus group discussions
focus group discussion
surveys

Presentation

presentation of childrens' views
presentation

Planning

contingency planning
planning
a plan

Prevention

violent conflict prevention
prevention

Political Action

advocacy strategies
advocacy

Meetings

workshop
coordination meeting
regular meetings
field coordination meetings
srcs health planning meeting
health planning meeting
planning meeting
meeting
somalia red crescent society planning meeting
meeting

meetings with partners

Medical Interventions

srh interventions (safe motherhood)
 campaigns against sexual harmful practice
 clinics
 essential health care services
 emergency nutrition responses
 sam
 sam treatments
 sam treatment
 support for delivery of routine health services
 delivery of routine health services

Technical Support

it and logistics support for key somali branches
 it and logistics support
 it support
 logistics support
 technical and financial support

Financial

microfinance
 the start-up of a small number of best-practice institutions
 best-practice institutions
 start up
 start-up

Concerns

Children

threats on the rights of children
 threats
 rights of children
 child well-being
 needs of children
 damaged schools

Gender

gender equality
 women's initiatives

Violence

peaceful resolution to the conflict
 peaceful resolution
 regional conflicts
 national call for peace
 issues and priorities for conflict prevention

conflict prevention
human costs of decades of war
human costs
stable areas
stable area
conflict-sensitivity

Famine
food security and nutrition
food security
nutrition
basic needs
targeted supplementary feeding (tsf) programme
severe acute malnutrition
sam
sam management services
cycle of relief-based food aid

Poverty
problem of poverty
poverty
poverty
social and economic status of women and their families
social and economic status
food and livelihood security
rural incomes
rural income

Aid Organization Capability
capacity of the unit
plans and commitments of all red cross red crescent members
revised plans
recovery programmes will be prioritized and financially supported

Governmental Capability
more effective and responsive government agencies
responsive government agencies
more effective government agencies
reforms and conditions
reforms
national societies' capacities for disaster preparedness
national societies' capacities
economic and social development of the state
needs of a country

Coordination
cooperation with un agencies, international organizations and donors

cooperation with un agencies
 cooperation with international organizations
 cooperation with donors
 interaction with their respective governments
 interaction with their government
 complimentary and in some cases subsidiary role
 complimentary role
 subsidiary role
 coordination mechanisms
 coordination mechanism
 multi-sectoral response
 strong links
 strong link
 links
 link

Financial

small business loan instrument
 small business loan instruments
 pooled funding mechanisms
 funding mechanisms

Health

cases of influenza a/h1n1
 influenza a/h1n1
 emergency water, sanitation and hygiene (wash)
 water, sanitation and hygiene (wash)
 water sanitation needs
 water sanitation
 outbreaks
 health services delivery system
 health services
 delivery system
 capacity of health system
 health system
 public health emergencies
 health emergencies
 awd outbreaks
 health related strategies
 health related strategy
 human suffering stemming from lack of adequate water supplies
 human suffering
 wash response plan
 wash coordination units
 health and nutrition

Farming

improved seed
agriculture-based livelihoods
resilient food security systems
natural resources
productivity of irrigated rice cultivation
irrigated rice cultivation
external inputs
farmer initiative
alternative systems organic fertilizer production
alternative systems
organic fertilizer production
fertilizer production
field preparation
irrigation regimes
irrigation regime
sustainability

Herding
animal health
commercial de-stocking
supplementary livestock feed

Early Warning Risk Reduction
early warning and early response
community-based risk reduction programme
risk reduction programme
risk reduction program
risk reduction
preparedness
emergency preparedness
comprehensive response
early warning, preparedness and mitigation measures
early warning measures
preparedness measures
mitigation measures
emergency response interventions
response intervention
emergency interventions
humanitarian response
response mechanisms
risks

Education
access to education
literacy lessons

General Concerns

mechanisms
 crime
 capacities
 stakeholders' reflective forums
 concerns
 factors
 progress of the tsunami recovery operation
 progress
 tight timetable
 existing and realistic needs
 realistic needs
 stakeholders input
 achievements
 groundwork for future development
 hundreds of thousands of lives
 un's ols
 mandate of its relief programs
 new branches
 new branch
 management information system
 political and economic contradictions
 political contradictions
 economic contradictions
 innovative efforts
 innovative effort
 economic and social development
 enhanced outreach strategy (eos)
 response
 duplication of efforts
 eos related activities
 enhanced outreach strategy
 community based intervention
 leading role in resource mobilization
 resource mobilization
 participation of relevant partners
 public relief
 active involvement of the malian government
 active involvement
 scale of intervention

Things

Documents
 un convention
 reports
 report
 convention
 guidelines

an action plan on child sexual abuse and exploitation
 leaflets on floods, cyclones and tsunamis
 a movement cooperation plan
 a website
 the website
 humanitarian requirements documents (hrd)
 hrd
 documents
 document

Medical Supplies
 clinic
 clinics
 a field health officer
 doses of meningitis vaccine
 doses
 h1n1 vaccine
 vaccine
 vaccines
 emergency health kits (ehks)
 ehk
 diarrheal disease kit (ddks)
 ddk
 essential drug kit (edks)
 edk

Food
 relief food
 food
 fortified blended food
 fbf
 corn soya blend
 csb

Shelter
 clinics structures
 clinic structures
 structure
 structures
 tents

Educational Supplies
 information-education-communication (iec) materials
 iec materials
 school-in-a-box kits
 educational materials

Money

two million sudanese pounds
sudanese pounds
usd

Agricultural Things

village irrigated perimeters (vips)
rotary weeders

Other Things

digital cameras
digital camera
renewable supplies
supplies
necessary equipment for emergency
necessary equipment

Places

districts
southern sudan
maldives
western sahara
mauritania
senegal
gambia
sierra leone
ghana
togo
burkina faso
mali
guinea
guinea bissau
benin
liberia
cote d'ivoire
ivory coast
mali
nigeria
niger
cameroon
chad
central african republic
sudan
eritrea
ethiopia
gabon

congo
democratic republic of congo
ganda
rwanda
burundi
tanzania
malawi
angola
zambia
namibia
zimbabwe
south africa
mozambique
lesotho
swaziland
kenya
somalia
morocco
algeria
libya
egypt
tunisia

XI. Appendix V: SMA Sociocultural Taxonomy

Intent: To provide a *generalizable* typology for military and intelligence analysts and planners to characterize sociocultural systems. These systems could include a military organization, a terrorist organization, a tribal society or a nation state.

Basis: This typology is based on comparative analysis of academic sociocultural typologies, sociocultural typologies produced for the U.S. military and intelligence community (mapHT, MCIA, CIA, ABCA, SMA Sudan, SMA DAPSE, NASIC, NSA, ACE), and standard approaches to examining levers of power (DIME) and its effects (PMESII).

It is informed by emergent ontology standards (Noy)

Academic roots: Morgan, Durkheim, Radcliffe-Brown, Malinowski, Steward, Hall, Flannery, Harris, Wilson, HRAF OCM

Caveats:

- Any specific problem will probably **only require a portion of the typology** (i.e. do not have to know everything about a sociocultural system to deal with an aspect of it).
- However, when dealing with a particular problem, analysts and planners should **at least examine all of the high level categories** and inquire whether they influence the problem at hand. This will prevent parochial examination of problems and mirror-imaging.
- This typology is **not intended to replace specific typologies** in use that work. However, it should be **consistent with other DoD, IC and academic typologies**.
- The SMA Sociocultural Typology can be used to **translate findings and concepts** from one typology to another.
- Sociocultural typologies should not only be used to examine target societies, but should be **turned back upon US** and its military for better understanding of our own reactions to others and our own COAs.
- Use of the SMA typology should be **iterative** to catch **dynamic changes** in the sociocultural system under investigation

The typology is organized in a hierarchy that includes increasingly detailed levels. The levels are 5 Fundamental Categories (**Interests, Capabilities, Context, Decision Making Psychology, Language**), 10 High Level Variables: Interests (**Motivating Factors Religion and Ideology, Social Identity, Objectives**), Capabilities (**Economy, Technology and Other Capabilities**), Context (**Roles/Life Cycle, Demography, Political and Social Organization, Environmental and Historical Context and Other Actors**), **Decision Making Psychology, Language**), and Subcategories added as necessary.

SMA SOCIOCULTURAL TYPOLOGY	
Fundamental Categories	High Level Variables
Interests	Motivating Factors, Religion and Ideology
	Social Identity
	Objectives
Capabilities	Economy, Technology and Other Capabilities
Context	Environmental (Physical and Political) and Historical Context
	Roles/Life Cycle
	Demography
	Political and Social Organization
Psychology	Decision Making Psychology
Language	Language

Figure 49. Typology Hierarchy.

A. Interrelatedness

A major reason why a broad and generalizable typology is required is to highlight the interrelatedness of different social phenomena. Failure to appreciate this interrelatedness leads to unanticipated nth order effects and missed opportunities for influencing a social system. **Figure 49** illustrates *primary* causal links among the High Level Variables of the SMA Sociocultural Typology. Also, some elements of one category may reasonably appear in another category as well, for example a perceived historical fact may be the basis for a social identity. Arguably, more links could be drawn, but focusing only on the most basic links demonstrates that influence in one part of the system will have impacts on others. Consequently, all key parts of a sociocultural system relevant to a particular situation must be considered for sound analysis and planning. For example the eigenvector centralities of all nodes in **Figure 50** are roughly the same; all are pretty essential.

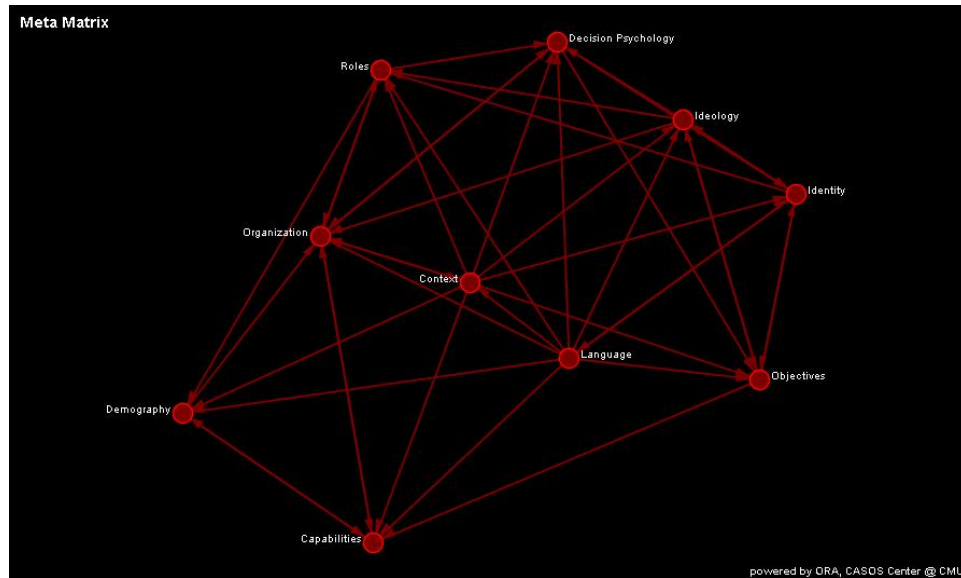


Figure 50. Basic Interrelatedness of High Level SMA Sociocultural Variables.

B. The 10 High Level SMA Typology Sociocultural System Variables

Environmental and Historical Context Other Actors (formerly Functional/Structural Environment Other Actors):

- Can be thought of as external factors that influence a system under study.

Physical Context: Climate, terrain, natural resources, arable land/pastures, water, security situation.

Other Actors: “No society is an Island” Often, relations with and influences from other societies are key factors influencing variables within a society, for example political/economic position in world. Influences range from political interference (Iran in Lebanon) to refugee populations (Darfurian in Chad, Iraqis in Jordan), to immigrants (Turks in Germany), to economic (Western capital intrusion into Third world societies, globalization), to cultural (spread of Western values and behaviors through globalization, spread of global salafist Jihadism). Note also important when considering potential deception by third parties. Ideally, the typology should be applied to all outside actors for a thorough understanding of them and how they may influence the target sociocultural system. **Historical Context:** All societies exist in time and their past influences nearly every aspect of social life. Historical context is also an important influence on ideologies, social identity and even concrete objective (e.g. reasserting dominance over an ethnic homeland).

Economy, Technology and Other Capabilities - includes:

- Production/Technology: “What people do for a living” jobs, productive activities (farming, horticulture, herding) PLUS technologies that people use (tools, weapons, implements) – indicators of capability
- Settlement/transportation/communications: “Where people live and how they get around/communicate” types and availability of housing, rural vs. urban settlement, road systems, communications – indicators of capability

- **Economic System:** “How production/exchange is organized at the social level” Includes: markets, barter systems, social division of labor (see Roles), industrial sectors, distribution of wealth and inequality

Demographics: “How people reproduce” Age/sex structure, age at marriage, availability of mates, marriage types (monogamy, polygamy); sexual behaviors, mate choice added since these vary by culture, are very important, and have an impact on demographic trends. Also includes mortality and morbidity (health and disease) patterns.

Roles/Life Cycle: “Functions and positions people play in groups” Social context (includes other categories) influences roles. Since a person's social roles typically change throughout life, they are included here. Effective operation in a culture requires knowing how roles change throughout the life cycle. Roles are functions carried out by people who hold particular statuses (status – man, function – hunt in a hunter-gatherer society; status – General, function – command in a military organization).

Political and Social Organization: “How are people in a society organized?” Huge category. Key for identifying decision units and constituencies Important: can include norms of behavior, habits, traditions, doctrines, rituals, and practices. Includes:

- Kinship (bilateral, patrilineal, matrilineal descent, kinship terminology) Influences how families are organized - often key primary alliances
- Sodalties: non-kin based social organization, county clubs, Rotary, etc.
- Political Parties
- Religious Organizations
- Military organizations

Motivating Factors, Religion and Ideology: Includes principles of leadership, political values (democracy, autocracy, communism), legal principles, military doctrine, religious dogma. Also includes basic existential and moral beliefs: Good/evil, Afterlife, moral principles such as honesty, Golden Rule, beliefs about proper place in social or natural world, cosmology; This is the symbolic realm of an actor's cognitive environment that provides aspirations – how they look at the world. Useful to separate into **codified** (i.e. military doctrine, church dogma, charters, legal codes) and **uncodified** (social norms, senses of right vs. wrong, morality). Codified motivating factors also include **rituals** and other scripted performances used to express motivating factors.

Social Identity – The constellation of factors brought together for self-identification or labeling by outsiders; these factors may include history, appearance, language, political objectives/ideology, geographic location, and any other element of the typology. The key here is that self-described identity may provide objectives and/or constrain actors to behave in certain ways. **Objectives:** Concrete goals that actors wish to achieve; often motivated by or justified by Motivating Factors, Identity, or even organizational structure (think of bureaucratic decision making and organizational culture).

Decision Making Psychology: Risk sensitivity, emotion, cognitive style, decision modes (typology), Maslow's hierarchy of needs, poliheuristic decision-making, bounded rationality, prospect theory, neuroscience. Influences range from more biological neuroscience to more social (emotional attachment to symbols) psychological considerations. Psychology permeates many aspects of other sociocultural variables.

Language – Symbolic communication is fundamental to nearly every aspect of human life and permeates all elements of the typology; Language includes not only traditional linguistic concepts such as knowledge of particular language/dialect used, grammar, lexicon, and phonetics, but also the socially appropriate use of language/dialect(s) in different social contexts.

XII. Appendix VI: Taxonomy of Socio-cultural Variables

Sociocultural Taxonomy

(216)

Interests (7)

Social Identity (1)

Religious Identity (1)

Motivating Factors, Religion, Ideology (6)

Materialism (1)

Uncodified Norms (5)

Economy, Technology and Other Capabilities (98)

Infrastructure (98)

Transportation (1)

Macroeconomics (12)

Domestic Economy (85)

Economic Risks (4)

Agriculture (54)

Animal Husbandry (13)

Finance (6)

Income (3)

Fishing (1)

Labor (2)

Transportation (1)

Education (1)

Context (106)

Environmental and Historical Context and Other Actors (43)

Geographic Factors (12)

Settlements (1)

Security Situation (30)

Roles and Lifecycle (3)

Demography (25)

Fertility (3)

Disease Morbidity (18)

Mortality (2)

Migration (2)

Political and Social Organization (35)

State (18)

Nonstate (14)

Social Fissures and Unrest (3)

Language and Symbolic Communication (1)

Social Science Methods (4)

Interests

Social Identity

Religious Identity

moslem

Motivating Factors, Religion, Ideology

Materialism

competition

Uncodified Norms

cultural beliefs

knowledge indigenous

religious events

community awareness

cultural/gender norms

Economy, Technology and Other Capabilities

Infrastructure

Transportation

roads

Macroeconomics

poverty

levels of poverty

economic inequities

inflationary pressures

food prices

private sector

price of cereals

gender equity

social and economic status

gdp

food prices

markets

Domestic Economy

Economic Risks

food crisis

insurance

diversified agricultural growth

credit institutions

Agriculture

agriculture

purchase feed

seasonal movement of laborers to commercial farms

crop

agro-pastoral

agricultural activities

small-holder farmers

chickpea

haricot bean

maize

potato

sorghum
 teff
 wheat
 seed
 cabbages
 carrots
 rice production
 smallholder
 farmers
 agriculture
 woodlots
 living fences
 windbreaks fields
 erosion control barriers
 yields
 intensification of production
 water
 seed
 fertilizer
 pesticides
 herbicides
 agricultural production
 rice farming
 farmers
 irrigation
 onions
 tomatoes
 beans
 backyard gardens
 yields
 wheat
 sorghum
 millet
 manure
 cash crop
 surplus produce
 weeding
 threshing
 plowing
 planting
 harvesting
 poultry
 water pumps

Animal Husbandry
 livestock
 pastoral

- pastoralists
- herding
- pastoralism
- pastoralists
- pastoral semi-nomads
- herding
- grazing grounds
- cattle
- pasture shortage
- de-stocking
- small ruminant
- Finance
 - cash income
 - credit
 - loans
 - profits
 - household properties
 - competition over scare water and pasture resources
- Income
 - income
 - livelihood
 - livelihoods
- Fishing
 - fishing
- Labor
 - housekeepers
 - labor
- Transportation
 - petrol
- Education
 - schools
- Informal Economy

Context

Environmental and Historical Context and Other Actors

Geographic Factors

- water
- ecological deterioration
- resource scarcity
- water shortage
- flooding
- drought

water supply
 rainfall
 soil infertility
 desertification
 climatic variability
 organic matter

Settlements

rural

Security Situation

conflict
 war
 political exclusion
 marginalization
 weakness of the state
 un-responsive policing
 illegal goods
 secure its borders
 corruption
 human rights abuses
 lack of political participation
 concentration of power and wealth
 militarization
 proliferation
 clashes between ethnic groups
 vigilante groups
 election fraud
 root causes
 conflict
 peace building
 national security
 police
 anti-stock theft unit
 organized crime
 nepotism
 cronyism
 internally displaced persons
 conflict
 displacement
 insecurity

Roles and Lifecycle

education
 married
 literacy

Demography

Fertility

- high fertility
- annual growth rate
- demographic balance

Disease Morbidity

- health
- morbidity
- female genital mutilation (fmg)
- health
- sanitary practices
- infectious disease
- latrine coverage
- hygiene
- sanitation
- safe water
- awd acute watery diarrhea
- measles
- malaria
- malnutrition
- pandemic influenza a/h1n1
- meningitis
- yellow fever
- malnutrition

Mortality

- mortality
- under five mortality

Migration

- migration
- population movement

Political and Social Organization

State

- state
- district
- sub-county, parish
- democracy
- land policy
- land ownership
- private property rights
- parliament
- police
- military
- academic institutions
- research institutions
- democracy

state government
ministries
citizens
government
woreda

Nonstate

village
ethnic tribe
household
ethnic groups
religious leaders
council of elders
chiefs
community
family
communities
regions
civil society
households
villagers

Social Fissures and Unrest

refugee
political and economic contradictions
discrimination

Language and Symbolic Communication

dialects

Social Science Methods

process tracing
discourse analysis
qualitative research techniques
focus groups

XIII. Appendix VII: CALICO Reference Documents

The following describes all CALICO project artifacts (folders, filenames, contents)

Folder Name	File Name	Contents
pdf	4d.pdf	Human readable version of query [4d-01]
	4e1.pdf	Human readable version of query [4e-01]
	4e11.pdf	Human readable version of query [4e-01] (no filtering)
	4e2.pdf	Human readable version of query [4e-02]
mdb (MS Access file format)	4d.mdb	Machine readable version of query 4d.pdf
	4e1.mdb	Machine readable version of query 4e1.pdf
	4e11.mdb	Machine readable version of query 4e11.pdf
	4e2.mdb	Machine readable version of query 4e2.pdf
misc	CALICO_WebCrawl.zip	A standalone version of the corpus user interface
	CALICO-stop-words.txt	The stop words used within CALICO
	IBase-design-report	A schema for the phrase/category/document database
	CALICO_txtsums.zip	Machine created document summaries of each document
pos	CALICO_pos.zip	Result from the part-of-speech tagger
tagged	CALICO_ALL.cbf.xml	The complete corpus tagged by the Semio Tagger engine
tax	CALICO_ALL_v2.txt	All CALICO taxonomies combined
	CALICO_Predicate_Objects_v2.txt.xml	- self described
	CALICO_Predicate_Verbs_v2.txt.xml	- self described
	CALICO_Relief_Orgs_v2.tax.xml	- self described
	CALICO_Sociocultural_v2.txt	- self described
	Taxonomy-Improvements.xlsx	- results of taxonomy improvement
wfreq	WF_AllTexts.xlsx	Word frequency for all words within corpus
	WF_ConLatchedTop1k.xlsx	Word frequency limited to latched phrases
	WF_ALLCompare.xlsx	Comparison of the two above

XIV. Appendix VIII: Bibliography of Corpus Documents

- Abiew, Francis Kofi. "From Civil Strife to Civic Society: NGO-Military Cooperation in Peace Operations." <http://www.peacedividendtrust.org/EIPdata/Library/Analysis%20and%20Reform%20of%20Peacekeeping%20Operations/CivMil%20and%20NGOs.pdf>.
- "Africa Regional Strategy for Disaster Risk Reduction." *Prevention Web*. http://www.preventionweb.net/files/4038_africaregionalstrategy1.pdf.
- Arthur, Peter. "ECOWAS and Regional Peacekeeping Integration in West Africa: Lessons for the Future." *Africa Today* 57, no. 2 (Winter 2010). DOI: 10.1353/at2010.0015.
- "Band Aids and Beyond, Tackling disasters in Ethiopia 25 years after the famine." *Oxfam International*. <http://www.oxfam.org/policy/band-aids-and-beyond>.
- "Best Practices & Innovations (BPI) Initiative, Agriculture & Rural Livelihoods, System of Rice Intensification (SRI), Africare." *Interaction*. <http://www.interaction.org/sites/default/files/2190/BPI%20Round%203%20-%20Africare%20-%20Updated.pdf>.
- Boettcher, William A. "Military Intervention Decisions Regarding Humanitarian Crises, Framing Induced Risk Behavior." *The Journal of Conflict Resolution* 48, no. 3 (June 2004): 331-355. doi: 10.1177/0022002704264271.
- Bruell, Abby. "Central African Republic, Chad & Sudan Member Activity Report, A Guide to Humanitarian and Development Efforts of InterAction Member Agencies in Central African Republic, Chad and Sudan." *InterAction*. http://www.interaction.org/sites/default/files/10-6-10%20Final%20Copy_0.pdf.
- Chang, Trish. "An Analysis of Humanitarian and Protection Operations for the Internally Displaced in Darfur." *Kofi Annan International Peace Keeping Training Center*. http://www.kaiptc.org/Publications/Occasional-Papers/Documents/no_18.aspx.
- "Cote D'Ivoire – Complex Emergency, Fact Sheet #15, Fiscal Year (FY) 2011." *USAID*. <http://reliefweb.int/sites/reliefweb.int/files/resources/07.15.11%20-%20USAID-OFDA%20Côte%20d'Ivoire%20Complex%20Emergency%20Fact%20Sheet%202315.pdf>.
- Cozens, Eleanor. "Education for Girls and Young Women in Liberia." *Children in Crisis*. <http://www.childrenincrisis.org/sites/default/files/uploads/press/RESOURCES%20%20LIBERIA%20GIRL'S%20EDUCATION%20PROJECT%20-%20EVALUATION%20REPORT.pdf>.
- "Doctors Without Borders/Medecins Sans Frontiers (MSF)." maic.jmu.edu/journal/3.3/profiles/DoctorsWithoutBorders.doc.
- "East & Central Africa DRR Strategy." *USAID*. http://www.usaid.gov/our_work/

humanitarian_assistance/disaster_assistance/regional/files/eca_drr_strategy2.pdf.

“East Africa Earthquake and Tsunamis Quarterly Report.” *International Federation of Red Cross and Red Crescent Societies*. <http://reliefweb.int/sites/reliefweb.int/files/resources/0C19C6C70C52E4D949257172002270B3-ifrc-tsunami2-17may.pdf>.

“East Africa Earthquake and Tsunamis Second Quarterly Report.” *International Federation of Red Cross and Red Crescent Societies*. <http://www.ifrc.org/docs/appeals/04/280457c.pdf>.

“Federation-wide Tsunami Semi-annual Report East Africa: Somalia and Seychelles.” *International Federation of Red Cross and Red Crescent Societies*. http://reliefweb.int/sites/reliefweb.int/files/resources/6D19B3D631F4174E8525751A006968C8-Full_Report.pdf.

Field, Tracy-Lynn. “International Disaster Response Law Research Report: Southern African Region.” *Institute of Foreign and Comparative Law University of South Africa*. http://www.ifrc.org/Global/Publications/IDRL/country%20studies/IDRL_sa_report.pdf.

Fielder, Dave. “Defining Command, Leadership, and Management Success Factors within Stability Operations.” *Peacekeeping & Stability Operations Institute*. <http://pksoi.army.mil/PKM/publications/papers/paperreview.cfm?paperID=21>.

Franks, Suzanne. “The Neglect of Africa and the Power of Aid.” *International Communication Gazette* 72, no. 1 (February 2010): 71-84. doi: 10.1177/1748048509350339.

Guang-qiu, Xu. “Global Governance: The Rise of Global Civil Society and China.” *Fudan Journal of the Humanities and Social Sciences* 4, no. 1 (2011). <http://en.ias.fudan.edu.cn/indexInfo.asp?ID=27&T=Announcements>.

“Horn of Africa - Drought, Fact Sheet #3, Fiscal Year (FY) 2011.” *USAID*. <http://reliefweb.int/sites/reliefweb.int/files/resources/07.21.11%20-%20USAID-DCHA%20Horn%20of%20Africa%20Drought%20Fact%20Sheet%20%233%20-%20FY%202011.pdf>.

“Humanitarian Assistance to Internally Displaced Persons in South Africa.” *University of the Witwatersrand*. <http://cormsa.org.za/wp-content/uploads/Research/Disaster/DisasterREvaluation.pdf>.

“Humanitarian Requirements-2011.” *ReliefWeb*. http://reliefweb.int/sites/reliefweb.int/files/resources/08146F0C4A95FC7C8525783100724B3D-Full_Report.pdf.

“Hunger, disaster, hope: rethinking humanitarian action in africa.” *International Federation of Red Cross and Red Crescent Societies*. [http://www.ifrc.org/Global/Publications/general/177700-BiA-Report-EN%20\(2\).pdf](http://www.ifrc.org/Global/Publications/general/177700-BiA-Report-EN%20(2).pdf).

“IFAD in Sudan.” *International Fund for Agricultural Development (IFAD)*.

<http://operations.ifad.org/web/ifad/operations/country/home/tags/sudan>.

“IFAD operations in Ghana.” *International Fund for Agricultural Development (IFAD)*.
<http://operations.ifad.org/web/ifad/operations/country/projects/tags/Ghana>.

“IFAD operations in Rwanda.” *International Fund for Agricultural Development (IFAD)*.
<http://operations.ifad.org/web/ifad/operations/country/projects/tags/rwanda>.

“Implementation of the UN Convention on the Rights of the Child in Uganda.” *Uganda Child Rights NGO Network*. [http://www.ucrnn.net/ucrnn/docs/1st%20Alternative%20Report%20to%20UNCRC%20\(Dec%202002\).pdf](http://www.ucrnn.net/ucrnn/docs/1st%20Alternative%20Report%20to%20UNCRC%20(Dec%202002).pdf).

“Indigenous Knowledge in Disaster Management in Africa.” *United Nations Environment Programme*.
<http://www.unep.org/IK/PDF/IndigenousBooklet.pdf>.

Iribarnegaray, Deanna. “Peacekeeping’s New Partnerships.”
<http://www.peacestudiesjournal.org.uk/dl/peacekeeping.PDF>.

Jarrell, Lynne, Bobbi Gray, Megan Gash, and Chris Dunford. “Human Faces of Microfinance Impact—What We Can Learn from Freedom from Hunger’s ‘Impact Story’ Methodology.” *Freedom from Hunger*. <http://www.ffhtechical.org/resources/research-reports/human-faces-microfinance-impact-what-we-can-learn-freedom-hunger-s-impact>.

Kut, George. “Mobilizing Early Response Project-Kenya, Kenya Assessment Summery Report.” *Nairobi Peace Initiative-Africa*. <http://gppac.net/uploads/File/Kenya%20MERP%20ASSMENT%20Report.doc>.

“Making Cash Count, Lessons from cash transfer schemes in east and southern Africa for supporting the most vulnerable children and households” *Save the Children*. <http://www.ipc-undp.org/publications/cct/africa/MakingCashCountfinal.pdf>.

“Multi-disciplinary assessment mission: Report on the Horn of Africa food crisis.” *ReliefWeb*.
http://reliefweb.int/sites/reliefweb.int/files/resources/2F16AE5C039F6747C125751C0050DFE2-Full_Report.pdf.

“Multiple Use Water Services for the Poor: Assessing the State of Knowledge.” *Winrock International*.
<http://www.winrockwater.org/docs/Executive%20Summary%20Multiple%20Use%20Water%20Services%20Final%20report%20feb%2008.pdf>.

“NGO Complementary Report on Implementation of the Optional Protocol to the UN Convention on the Rights of the Child on the Sale of Children, Child Prostitution and Child Pornography in Sierra Leone.” *UN Committee on the Rights of Children*. <http://resourcecentre.savethechildren.se/rb?q=cmis/browser&id=workspace://SpacesStore/ce748a08-6f66-484b-9805-4eeddb831b48/1.6>.

- "NGO Partner Survey 2010, Public Report." *Keystone Performance Surveys*.
http://www.keystoneaccountability.org/sites/default/files/Keystone%20partner%20survey%20Jan2011_0.pdf.
- Nkosi, Dr. Joanna and Dr. Morley Nkosi. "Reflections on International Humanitarian Interventions in Africa." *USA for Africa*. <http://www.usaforafrica.org/pdf/completereport.pdf>.
- "Office of U.S. Foreign Disaster Assistance Annual Report for Fiscal Year 2008." *USAID*.
http://www.usaid.gov/our_work/humanitarian_assistance/disaster_assistance/publications/annual_reports/pdf/AR2008.pdf.
- "Office of U.S. Foreign Disaster Assistance Annual Report for Fiscal Year 2009." *USAID*.
http://www.usaid.gov/our_work/humanitarian_assistance/disaster_assistance/publications/annual_reports/fy2009/annual_report_2009.pdf.
- "Office of U.S. Foreign Disaster Assistance, Annual Report for Fiscal Year 2009." *USAID*.
http://www.usaid.gov/our_work/humanitarian_assistance/disaster_assistance/publications/annual_reports/fy2009/annual_report_2009.pdf.
- Perry, Charles M. "Finding the Right Mix Disaster Diplomacy, National Security, and International Cooperation." <http://www.ifpa.org/pdf/TheRightMix.pdf>.
- "Programme of Action for the Implementation of the Africa Regional Strategy for Disaster Risk Reduction (2006-2010)." *African Union*. http://www.africa-union.org/Agriculture/Disaster_Risk_Reduction/Programme_of_Action.doc.
- "Report of the Second Ministerial Conference on Disaster Risk Reduction." *African Union*.
http://www.preventionweb.net/files/18733_englishreport.pdf.
- "SADC regional development strategy 2005." *Southern African Development Community*.
<http://www.sardc.net/editorial/sadctoday/documents/v7n6.pdf>.
- Schmeidl, Susanne and J. Craig Jenkins. "Measuring International Forced Migration, 1969-1999, The Global Forced Migration Project." *International Journal of Sociology* 33, no. 3 (Fall 2003): 70-87.
<http://www.jstor.org/pss/20628686>.
- "Smart Development in Practice, Field report from southern Sudan." *Oxfam*.
<http://www.oxfamamerica.org/files/field-report-southern-sudan.pdf>.
- Smock, David R. "Humanitarian Assistance and Conflict in Africa." *United States Institute of Peace*.
<http://www.usip.org/files/resources/pwks6.pdf>.
- "South, West, and North Africa – Disaster Risk Reduction (DRR) Programs." *USAID*.
http://www.usaid.gov/our_work/humanitarian_assistance/disaster_assistance/publications/pre_p_mit/mods/program_updates/SWAN_DRR_fs01_09302009.pdf.

- Spearin, Christopher. "Private, Armed and Humanitarian? States, NGOs, International Private Security Companies and Shifting Humanitarianism." *Security Dialogue* 39, no.4 (August 2008): 363-382. doi: 10.1177/0967010608094034
- Stoddard, Abby. "Beyond the Transatlantic Divide: Deepening Western Cooperation and Broadening Donor Engagement in International Humanitarian Assistance." *Global Public Policy Institute*. http://www.disastergovernance.net/fileadmin/gppi/Stoddard_2009_Transatlantic_Divide.pdf.
- "Testing Approaches for Increasing Skilled Care During Childbirth: Key Findings from Homabay and Migori Districts, Kenya." *Family Care International*. <http://www.familycareintl.org/UserFiles/File/SCI%20Kenya%20Report%20March%202008.pdf>.
- "The African Report on Child Wellbeing 2011, Budgeting for Children." *African Child Policy Forum* (2011). <http://resourcecentre.savethechildren.se/content/library/documents/african-report-child-wellbeing-2011-budgeting-children>.
- "The causes of conflict and the promotion of durable peace and sustainable development in Africa." *United Nations*. <http://www.un.org/ecosocdev/geninfo/afrec/sgreport/report.htm>.
- "The Inspection Panel Report and Recommendation On Request for Inspection LIBERIA: Development Forestry Sector Management Project (Trust Funds Nos. TF057090-LR; TF096154-LR; and, TF096170-LR)." *World Bank*. http://www.wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2011/03/31/000333038_20110401000026/Rendered/PDF/580110IPR0REVI1Disclosed03124120111.pdf.
- "Uganda Humanitarian Profile." *United Nations*. http://reliefweb.int/sites/reliefweb.int/files/resources/A7912A54E132F09AC125780E0046085B-Full_Report.pdf.
- "USAID Annual Report Kenya 2009." *USAID*. <http://kenya.usaid.gov/sites/default/files/USAID%20AnnualReport-2009%20Revised.pdf>.
- Utterwulghse, Steve. "Conflict Management in Complex Humanitarian Situations: Peacemaking and Peacebuilding Work with Angolan IDPs." *Journal of Refugee Studies* 17, no. 2. http://www.sfcg.org/Documents/DME/conflict_management_in_complex_humanitarian_situations.pdf.
- Weathington, Teresa. "InterAction Member Activity Report West Africa, A Guide to Humanitarian and Development Efforts of InterAction Member Agencies in Cote d'Ivoire, Guinea, Liberia, and Sierra Leone." *InterAction*. <http://reliefweb.int/sites/reliefweb.int/files/resources/F6B70EECA6B8F6D549256D630008B83E-inter-wes-30jun.pdf>.
- "WFP Strategic Plan 2008-2013." *United Nations*. <http://documents.wfp.org/stellent/groups/public/documents/communications/wfp228800.pdf>.
- Whittall, Jonathan. "'It's like talking to a brick wall' humanitarian diplomacy in the occupied Palestinian

territory." *Progress in Development Studies* 9, no. 1 (January 2009): 37-53. doi:10.1177/146499340800900104

"World Disaster Report 2010 Focus on Urban Risk." *International Federation of Red Cross and Red Crescent Societies*. <http://www.ifrc.org/Global/Publications/disasters/WDR/wdr2010/WDR2010-full.pdf>.

"World Food Programme 2009." *United Nations*. <http://documents.wfp.org/stellent/groups/public/documents/newsroom/wfp204445.pdf>.

"World Food Programme 2010." *United Nations*. <http://documents.wfp.org/stellent/groups/public/documents/communications/wfp220666.pdf>.

Yousafzi, Dr. Aisha and Karen Edwards. "Double Burden, A situation analysis of HIV/AIDS and young people with disabilities in Rwanda and Uganda." *Save the Children*. http://www.savethechildren.org.uk/en/docs/Double_Burden.pdf.

XV. Team Contact Information

T Hill & Associates, Inc.
Tiffany Hill, President
426 M St NW, Suite G
Washington, DC 20001
tiffany@thillassocaites.com
www.thillassociates.com

Millennium KI, LLC
Rick Myskey, President & CEO
102 Person St, Suite D
Fayetteville, NC 28301
rick.myskey@millenniumki.com
www.millenniumki.com

National Security Innovations, LLC
Dr. Lawrence A. Kuznar, Chief Cultural Sciences Officer
P.O. Box 263
Medway, MA 02053
lkuznar@nsiteam.com
www.nsiteam.com

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.				
1. REPORT DATE (DD-MM-YYYY) 10-2011		2. REPORT TYPE Final		3. DATES COVERED (From - To) June 2011 - October 2011
4. TITLE AND SUBTITLE CIVIL AFFAIRS LANGUAGE FOR INFORMING CULTURAL OPERATIONS (CALICO)			5a. CONTRACT NUMBER W9132T11P00476	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Hill, Tiffany; Myskey, Rick; and Kuznar, Lawrence A.			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) T Hill & Associates, Inc. Tiffany Hill, President 426 M St NW, Suite G Washington, DC 20001			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Engineer Research and Development Center Construction Engineering Research Laboratory PO Box 9005 Champaign, IL 61826-9005			10. SPONSOR/MONITOR'S ACRONYM(S) CEERD-CN-B	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT The report presents the results of content analysis of a corpus of documents from humanitarian and disaster relief (HA/DR) agencies and nongovernmental organizations (NGOs) active in Africa. The research was aimed at discovering what the corpus of HA/DR and NGO documents revealed about the sociocultural information required to support effective interventions to ameliorate the human consequences of disruption and to enhance the effectiveness of indigenous governance. The report details the selection of the corpus and content analysis of the corpus. It provides a taxonomy of relief activities and of the beneficiaries of those activities. Analysis of the corpus of reports of HA/DR agencies and NGOs suggests that those organizations focus on higher-level administrative functions, in particular, on those functions intended to facilitate and assist local and host nation governments in providing aid. To the extent that sociocultural factors and actions important to on-the-ground operations are important, the corpus reflects a strong bias toward economic variables.				
15. SUBJECT TERMS humanitarian assistance and disaster relief, nongovernmental organizations, content analysis, sociocultural factors, Africa				
16. SECURITY CLASSIFICATION OF: UNCLASSIFIED			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES 128
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified		
			19b. TELEPHONE NUMBER (include area code) 217-373-3432	