Use of metadata in Cross Language Information Retrieval for multinational counterdrug agencies

Laura Lucía Burkhart

Georgia Tech Research Institute

Atlanta, Georgia

United Counterdrug (UCD) Background

- Multinational forum to enhance policy development and operational coordination in the counternarcotics domain.

- Hosted by
  - United States Southern Command
  - Organization of American States InterAmerican Drug Abuse Control Commission (OAS/CICAD).

Georgia Tech Research Institute

United Counterdrug (UCD) Background

- ◆ **Multinational agencies**
  - ■ Operations
    - ● Air Interdiction
    - ● Maritime Interdiction
    - ● Customs
    - ● Police
  - ■ Policy
    - ● National Antidrug Councils
    - ● Foreign Ministries

- ◆ **US Agency Examples**
  - ■ Operations
    - ● JIATF-E
    - ● Coast Guard
    - ● Customs
    - ● DEA
  - ■ Policy
    - ● ONDCP
    - ● Department of State/INL

Georgia Tech Research Institute

United Counterdrug (UCD) Background

◆ Participating nations

- Argentina
- Bolivia
- Brazil
- Colombia
- Ecuador
- Guatemala
- Mexico
- Panama
- Paraguay

- Peru
- Trinidad & Tobago
- United States
- Venezuela
- Belize (observer)
- Chile (observer)
- Dominican Republic (observer)
- Uruguay (observer)

Georgia Tech Research Institute

Purpose

◆ Allow multinational agencies, regardless of their language, to exchange information by searching multilingual databases.

Challenges

- ◆ Language Barriers
  - ■ Three languages spoken among countries
  - ■ Different dialects spoken
  - ■ Glossary of terms
- ◆ Internet Access
  - ■ Most Latin American countries have slow internet access
  - ■ Access to large databases using a phone line is slow making a system ineffective.

Research Goals

- ◆ Develop methods for multinational agencies in a domain specific area to effectively exchange information in a timely manner regardless of language

Tasks

- Structure unstructured documents (metadata)

- Text analysis

- Keyword weighting techniques

- Generating and maintaining thesauri

- Classifying queries (metadata)

Current and future status

- Concept development stage

- Research will begin this fall

- Will prototype system in two or three countries including the US