Transnational Digital Government Research Project Highlights

José A. B. Fortes (on behalf of the TDG team[1])
University of Florida, Advanced Computing and Information Systems Laboratory
P.O. Box 116200, 339 Larsen Hall, Gainesville, FL, 32611-6200
1 (352) 392-4964

fortes@ufl.edu

Transnational digital government relies on collaborative government processes that use information technology to address problems of regional or global nature. The broad context of the research done in this project is the ongoing process of transnational counter-narcotics cooperation among all Western Hemisphere governments whose goal is to reduce illicit drug production, traffic and consumption. The process is coordinated by CICAD – the Inter-American Drug Abuse Control Commission – which is a technical body of the Organization of American States (OAS), and composed of 34 member states. This project conducts research that has broad applicability to processes that require countries to monitor, share, analyze and integrate data and information on activities related to illegal drugs and other criminal activities. Broadly stated, research challenges on technology and tools for the collection, processing, exchange and integration of information needed by transnational digital government include:

- Spoken dialogue systems for data collection, training and learning;

- Data management and security techniques for role-based data sharing and filtering;

- Machine translation technology for sharing documents across different languages and countries;

- Middleware for transnational (heterogeneous) information grids that enable private, secure and dependable automation of collaboration processes and policies, and the delivery of computing services through Internet portals; and

- Network behavior modeling and optimization for delivery of acceptable quality of service.

The government-academia international partnership involved in the project consists of a team of researchers from seven universities (Carnegie Mellon U., North Carolina State U., U. of Belize, U. of Colorado, U. of Florida, U. of Massachusetts and Pontificia Universidad Católica Madre y Maestra (PUCMM)) and experts from agencies in three different countries (US, Belize and the Dominican Republic). Under the umbrella of the OAS, several ministries and agencies in the three countries are involved. These include OAS departments in Washington, D.C. (the Office of Science and Technology, the Office of Information Technology Services and CICAD's Inter-American Observatory on Drugs); the National Drug Abuse Control Council of Belize's Ministry of Health; and the National Drug Council of the Dominican Republic. The university researchers include experts on speech-based interfaces, machine translation, databases, information retrieval, Internet-computing, software requirements and networking.

The project focused its initial efforts on the research issues and implementation challenges faced by the design and development of a prototype of a system that allows countries to find, notify and receive accessible transnational information on dangerous individuals who cross borders. The prototype is a distributed software system that integrates a distributed query system; software for event creation, subscription and notification; conversational interfaces; machine translation engines; and virtualization middleware. Evolving versions of this prototype have been demonstrated in Belize, Washington, DC, and Dominican Republic. Copies have been deployed at universities in the USA, Belize and Dominican Republic universities and are currently being used for testing, evaluation and data collection purposes. Deployment in border stations in Belize and the Dominican Republic are scheduled for the second half of 2005. These prototypes are Web-accessible to the entire research team, and are being selectively used by staff from border control agencies, allowing the gathering of data for further development of machine translation and conversational interfaces, and providing feedback on user-interface and software requirements.

The project's unique features and early prototype development have engaged students and faculty in unique multinational, multilingual, multicultural educational experiences. In addition, they have triggered interested requests from other agencies as well as equipment grants to the participating countries. In spite of its success, the project also has had to face several changes and barriers to its progress. They are consequences of the international

---

[1] The TDG team consists of: A. I. Antón (anton@csc.ncsu.edu), O. Brooks (ndacc@btl.net), J. Carbonell (jgc@cs.cmu.edu), V. Cavalli-Sforza (violetta@cs.cmu.edu), R. Cole (cole@colorado.edu), R. Connolly (RConnolly@oas.org), L. de Brens (ldebrens@pucmmsti.edu.do), J. Fortes (fortes@ufl.edu), M. Herrera (consejo@codetel.net.do), C. McSweeney (cmcsween@ub.edu.bz), B. Piñeres (BPineres@oas.org), S. Su (su@cise.ufl.edu), D. Towsley (towsley@cs.umass.edu), J. Ventura (jventura@pucmmsti.edu.do), W. Ward (whw@colorado.edu).

nature of both the system under development and the research activities themselves. The common trait across the challenges is heterogeneity, i.e. there are significant differences in culture, infrastructure, policies and procedures across the involved institutions and countries. Information technology has also been successfully used by the research team to address some of these challenges. In particular, a web-based collaborative environment that includes collaboration forums and virtualized resources has enabled joint development of ideas and software across universities and countries.

Successive versions of the prototype have improved several aspects of the system. The texts used for training the machine translation and/or dialogue systems have been extended to include interviews with military and station officers responsible for immigration and security operations, dialogues and descriptions from research at border points, constructed dialogues produced by the OAS, and a significant collection of reports from project Aventinus, a system for retrieval of multilingual documents concerning international crime. Also developed was a corpus management system to facilitate the task of adding, classifying and selecting subparts of the corpus for training and testing purposes. Substantial effort has gone into making the system easier to use by others on both Windows and Linux platforms. So far, by using the prototype system in a number of scenarios, we have demonstrated that users/agencies in different countries, who use different languages, can 1) use either a form-based interface or a conversation-based interface to query for data stored in different countries, 2) define and subscribe to events of interests, and receive notification by emails, short messages through cellular phones, or remote program/process activations to enact processes to coordinate the activities of different agencies, 3) protect data security and privacy by specifying in export schemas those and only those data that users/agencies are willing to share with others, 4) define and enforce policies, regulations, security and privacy using rules that are automatically triggered by the occurrences of events, and 5) protect operational security by defining, and publishing only those sharable operations as Web services.

In summary, this transnational digital government project looks at unique challenges that either are absent, or are not acute, in national government activities. Individual countries differ in their languages, laws, regulations, cultures, administrative structures, resources, geopolitical characteristics and stages of technological development – all of which impact interactions among government agencies and impose unique requirements for software, language technologies and distributed processing. The project stresses the state of the art of distributed processing due to infrastructure heterogeneity, exposes socio-political constraints on scalability and sustainability of TDG solutions, and identifies software requirements of immigration processes.