Item 12 of the provisional agenda

Activities relating to the Working Group on Romanization Systems

Toponymic Database

Romanization Application (software)*

* Prepared by Naima Friha (Tunisia), Ing. MSc Geo-Informatics and Mohamed Ali Ilahy, Computer Sciences Ing, Tunisia
TOPONYMIC DATABASE
Romanization application (software)

Introduction

In the frame of updating its geospatial data, the CNT is setting up a topographic database built upon recent high resolution satellite images. A toponymic database is one subset of the topographic one. Geographic names are written in Arabic and Romanized according to Beirut Romanization system. A free software has been developed to perform automatically the transliteration from Arabic to Latin.

I – Toponymic Database

The following tables are defined based on the Beirut Romanization system

- Arabic letters table: contains all Arabic characters including specific ones based on the Tunisian dialect.
  Example: ن (na)

- Corresponding letters table: Each Arabic character has a corresponding Latin character
  Example: A /

N.B: In this table only syntactic rules are defined

- Layers table: contains all toponymic layers.
  Example:
  - ZH040: Layer of toponyms related to locations
  - BH170: Layer of toponyms related to Hydrography

- Area of coverage and projection system table: In this table are defined the coverage area and the projection system for one-time identification of each toponym and each toponymic layer.

- Geographic names table: In this table are defined the main characteristics of geographical names (Identifier, Layer, X and Y coordinates, French transcription toponym, Arabic transcription toponym, the Beirut transcription toponym. (The last field will be filled following the compilation).

- Saving table (Data security aspect): This table is implemented in order to save toponyms already processed and which may be restored when needed and also to keep track of processed databases.

- Toponymic Database Implementation
II – Developing an automated application which transliterates Arabic letters to Latin according to Beirut system

- **Definition of piles:** Two piles are stated for the processing: push and pull
  - First Pile FIFO: First In First Out: for Arabic characters
  - Second Pile FIFO: First In First Out: for Latin characters

- **Implementation of semantic rules:** For transliterating geographic names, functions are programmed to treat exceptions. Those functions are inserted in the compiler main code.

- **Example:**
  - **اﻟﺸﻤﺴﻴﺔ** Al Ash-shamsiyya A
  - **اﻟﻘﻤﺮﻳﺔ** Al Al qamariyya Al
  - **طﻮﻳﻠﺔ** ou ū

- **Development of a compiler:**
  - Setting up a verification procedure: This procedure checks Arabic letters and allows their transliteration in case of availability.
  - Creation of an Arabic keyboard

In this keyboard are defined all Arabic characters including diacritics and specific characters resulting from Tunisian dialect like the ﯾ (ga).

- **Compiler:** Schema showing how the compiler functions
**Problems**

1. Difficulties in exploiting Unicode Characters during the display
2. More than one Font management
3. Arabic Language treatment (specific characters), especially for the Tunisian dialect.

**Prospects/Future axes**
1. Creating a font including Arabic and special Latin characters to resolve the display problems
2. Securing a link between the software and the initial DataBase:
   Developing an integrated module (VBA Language) to ensure the interoperability
   between the cartographic DataBase and the Toponimic Database

Ms Naima FRIHA, Ing. MSc Geo-Informatics
Mr. Mohamed Ali ILAHY, Computer Sciences Ing

Centre National de Télédétection
Route de la Marsa El Aouina, BP 200 -1080 Tunis Cedex, Tunisie. Tél. : +216 71761333,
Fax : +216 71760890
E-m@il:cnt.dg@cnt.nat.tn,  www.cnt.nat.tn