



Economic and Social Council

Distr.: General
17 July 2012

Original: English

Tenth United Nations Conference on the Standardization of Geographical Names

New York, 31 July-9 August 2012

Item 11 (d) of the provisional agenda*

Toponymic data files and gazetteers: data standards and interoperability

Storage and display of special characters from aboriginal languages in Canadian geographical names

Summary**

A long-standing issue in the dissemination of aboriginal names relates to the ability to store and display special characters used in Canadian toponyms in aboriginal languages. Two groups of names have been affected. One group contains names in the Inuit language, written in a system known as syllabics. The adoption of a language code and other changes have now made it possible for the Canadian Geographical Names Data Base and the Geographical Names of Canada website to display the syllabics. The second group of names uses “hard-to-construct” characters (roman characters combined with diacritical marks or symbols). A solution for that group of toponyms was found through the use of the Unicode standard, a computer industry international standard allowing the encoding, display and retrieval of special characters, regardless of platform or language. Thanks to the adoption of the standard, the Canadian Geographical Names Data Base and the Geographical Names of Canada website are now also capable of displaying special hard-to-construct characters using diacritics that are found in many aboriginal language names.

* E/CONF.101/1.

** The full report was prepared by Heather Ross, a toponymist at the secretariat of the Geographical Names Board of Canada, and is available in the language of submission only from <http://unstats.un.org/unsd/geoinfo/UNGEGN/ungegnConf10.html>.

