Tracking and Summarizing News

“Name of the principal inventor: Kathleen Mc Keown

Automatic News Tracking System Consolidates Multiple News Sources: Columbia Newsblaster is a system to automatically track the day’s news. There are no human editors involved – everything you see on the main page is generated automatically, drawing on the sources listed on the left side of the screen.

Every night, the system crawls a series of Web sites, downloads articles, groups them together into ““clusters”“ about the same topic, and summarizes each cluster. The end result is a Web page that gives you a sense of what the major stories of the day are, so you don’t have to visit the pages of dozens of publications.

Newsblaster is from the Natural Language Processing group at Columbia University’s Department of Computer Science. It is designed to demonstrate the Group’s technologies for multidocument summarization, clustering, and text categorization, among others. It is funded under DARPA TIDES and KDD and has been operational online since September 2001. New Tracker Provides Multi-Document Summarization, Categorization Current and future enhancements include international perspectives, multilingual capability, and tracking events across days.

See also M03-003 & MS99/02/06C. ”

Inventors

Kathleen McKeown