The Age of Content and Knowledge Processing

Makoto Nagao

National Institute of Information and Communications Technology,
4-2-1 Nukui-Kitamachi, Koganei, Tokyo 184-8795 Japan

Internet has become an indispensable infra-structure on the earth with more and more high speed digital information services provided year by year. Correspondingly, a wide variety of content is created, stored and distributed. This includes not only text but also speeches, music, pictures and videos. On the Internet, an unbelievably large amount of content of every kind is accessible from anywhere in the world, and a smart, intelligent information retrieval system has become major priority. People are not satisfied with the current retrieval systems. In Japan there are research and development projects for new generation retrieval systems, where a central issue is how to guarantee the trustworthiness and the authenticity of the retrieved information. Another serious issue is that content on the Internet is not organized systematically, and is not necessarily suitable for everybody to use. Self organization of varieties of information into a knowledge system will be another important technology to be developed in the future. Digital libraries that provide reliable content with rich related information using relatively simple access methods act as a kind of standard reference to establish technologies to solve the issues above. Because the scope of information access is all over the world, multilingual machine translation is an indispensable technology. The paper will discuss these problems by surveying the state of the art of content and knowledge processing.