

Information in Internet: publication, search and analysis

Anatoly Fedotov

Institute of Computational technologies, Novosibirsk, Russia

E-mail: fedotov@sbras.ru

This paper provides a brief review of an information publishing and transferring technologies, which influenced on the evolution of Internet. The author points out the following main periods of the evolution:

1. **The HTML** period (1990 – 1997) was characterized by the dramatic growth of the humans' interest to the new way of information interchanging. Mainstream technologies were oriented on the creating (CGI, SSI, PHP, Perl) and transferring (HTTP) human-readable documents (HTML, TeX, PostScript, PDF). Hence, a lot of developers made efforts to have “pretty-looking” documents (JavaScript, Flash). As a result, the only way to find a document with necessary information is to use global-search engines, which could find all the pages containing indicated keywords.
2. During the **XML** period (1997 – present time) understanding of the Internet have been changed. People started to realize that such a huge amount of data couldn't be managed manually. Therefore it needs to be well-ordered and well-structured for automated processing. A number of new technologies were developed. This paper, in particular, discusses the following: DublinCore initiative, web services, XML-based data schemas, protocols Z39.50, LDAP. Although the latter two technologies are not new and their concepts were developed in 80th, they took new life and were applied in new areas.
3. **The Semantic Web** period may be considered as a potential evolution for XML period. It presents a new level of the information representation comparing to XML. Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. It is based on the Resource Description Framework (**RDF**) technology, which integrates a variety of applications using XML syntax and URIs naming.

This paper also presents levels of the information formalization. Brief summary outlines information analysis tasks, which can be solved now, and directions of further IT-development.