

# SEMINAR

## SEMANTICS & SEARCH

April 28<sup>th</sup>, 2006

Location: F1

**09.15-** Opening of seminar  
**09.20** *Jon Atle Gulla, NTNU*

**09.20-** The Semantic Web and ontologies: what assumptions, what justifications?  
**09.50** *Karen Sparck-Jones, Cambridge*

**Abstract.** A great deal is written about the Semantic Web as if it already materially exists. There is also a widespread presumption that the Semantic Web requires an ontology, or some relatable set of ontologies, of a thoroughly logical kind. In my talk I will examine this presumption in relation to the manipulation of, and especially access to, information expressed in natural language. The realities of natural language information processing tasks show that aggressively formal ontologies are effective only for limited domains and communities, and that useful broad-cover, general-purpose ontologies have to be text-derived or at least text-endorsed, and will necessarily be soft and imperfectly logical.

**Short CV:** Karen Sparck Jones is emeritus Professor of Computers and Information at the Computer Laboratory, University of Cambridge. She has worked in automatic language and information processing research since the late fifties, and has many publications including nine books. She is a Fellow of the British Academy and of the American Association for Artificial Intelligence. She has received three awards for information retrieval research as well as, in 2004, the Association for Computational Linguistics' Lifetime Achievement Award. Her more recent research has been on information retrieval models and practice, on automatic summarising, and on system evaluation, where she is involved in international programmes.

**Publicity information:**

<http://www.cl.cam.ac.uk/~ksj>

**09.50-** Emerging Semantic Web Trends: Transparent and Trustworthy Applications  
**10.20** *Deborah L. McGuinness, Stanford*

**Abstract.** As web applications proliferate, more users (both people and agents) find themselves faced with decisions about when and why to trust application advice. In order to trust information obtained from arbitrary applications, users need to understand how the information was obtained and what it depended upon. Particularly in web applications that may use question answering systems that may be heuristic or incomplete or data that is either of unknown origin or may be out of date, it becomes more important to have information about how answers were obtained. Emerging web systems will return answers augmented with Meta information about how answers were obtained. In this talk, Deborah McGuinness will describe an approach that can improve trust in answers generated from web applications by making the answer process more transparent. The added information is aimed to provide users (humans or agents) with answers to questions of trust, reliability, recency, and applicability. While this is an area of active research, there are technologies and implementations that can be used today to increase application trustability. The talk will include descriptions of a few representative applications using this approach.

**Publicity information:**

<http://www.ksl.stanford.edu/people/dlm/publicity.html>

**10.20** Coffee Break (30 min)

**10.50-** XML full-text search and entity extraction  
**11.20** *Aleksander Øhrn, Fast Search & Transfer*

**Abstract:** I'll present some of the development activities done at Fast Search & Transfer where scalable search, XML technologies and entity extraction are brought together in order to form powerful search and discovery mechanisms. The combination of being able to express queries that involve semantically meaningful entities (and requirements on how they interrelate) plus the ability to analyze the returned result sets, enables some interesting applications. Some open problems will also be discussed.

**11.20-** Semantics for Personalized Search - An Example  
**11.50** *Per Gunnar Auran, Yahoo!*

**Abstract:** This presentation will introduce an example on how simple semantic relations can be used for a personalized search experience. The speech will discuss a prototype search application that was developed by Yahoo! Technologies Norway in 2004-2005 by a team lead by the author.

**Short CV:** Per Gunnar Auran is a Senior Research Scientist at Yahoo! Technologies Norway AS, where his main responsibility is search relevancy for Yahoo!'s vertical search platform, Vespa. He is the technical lead for research related to document analysis and ranking, query analysis and semantics, personalized and community search for Vespa.

From April 2000 to April 2003, Per Gunnar Auran was the R&D Manager of the Data Analysis Group, at Fast Search & Transfer ASA (FAST), focusing on web search data analysis and search relevancy for AllTheWeb, a leading web search engine of the time. Prior to that he was a research Scientist at the Norwegian Paper and Pulp Research Institute. He holds MSc and PhD degrees from the Norwegian Institute of Technology, specializing in engineering cybernetics.