

---

**Name of the proposed Centre of Excellence (CoE):**  
**Finnish Centre of Excellence for Algorithmic Data Analysis Research (Algodan)**

**Director:** Professor Esko Ukkonen

**Host organisations:** University of Helsinki, Helsinki University of Technology

**Senior researchers and their host organisations:**

Prof. Esko Ukkonen	University of Helsinki
Acad. Prof. Heikki Mannila	Helsinki University of Technology
Dr. Helena Ahonen-Myka	University of Helsinki
Dr. Jaakko Hollmén	Helsinki University of Technology
Dr. Patrik Hoyer	University of Helsinki
Dr. Aapo Hyvärinen	University of Helsinki
Prof. Jyrki Kivinen	University of Helsinki
Dr. Juha Kärkkäinen	University of Helsinki
Dr. Kjell Lemström	University of Helsinki
Dr. Juho Rousu	University of Helsinki
Dr. Marko Salmenkivi	University of Helsinki
Dr. Petteri Sevon	University of Helsinki
Prof. Hannu Toivonen	University of Helsinki
Dr. Roman Yangarber	University of Helsinki

**Applied funding for the years 2008 – 2010: 1 100 000 EUR**

**Abstract:**

The importance of data analysis in science and in industry is increasing continuously, as our ability to measure and store data grows. While data analysis is as old as science itself, the new methods of collecting raw data pose unprecedented challenges and opportunities to data analysis and to the algorithms of data analysis.

The Algorithmic Data Analysis (Algodan) Centre of Excellence develops new concepts, algorithms, principles, and frameworks for data analysis. The work combines strong basic research in computer science with interdisciplinary work in a wide variety of scientific disciplines and industrial problems.

The research of the Algodan CoE lies in the areas of combinatorial pattern matching, data mining, and machine learning. The work in Algodan is strongly interdisciplinary: we cooperate constantly with application experts in various application areas, formulating novel computational concepts and ways of attacking the scientific and industrial problems of the application areas. Developing new concepts and algorithms is an iterative process consisting of interacting extensively with the application experts, formulating computational concepts, analyzing the properties of the concepts, designing algorithms and analyzing their performance, implementing and experimenting with the algorithms, and applying the results in practice. The main application areas of the Algodan CoE are in biology, medicine, telecommunications, environmental studies, linguistics, and neuroscience.

The formulation of new computational concepts, their analysis, and the design of algorithms are some key ingredients that make the Algodan CoE unique. First, rather than concentrating on improvements to existing problems and methods, the CoE focuses on defining new tasks where significant impact can be made by introducing new concepts. Second, we emphasize the need for analyzing the performance of the algorithms, instead of just relying on heuristic approaches. Third, we use our strong background in algorithmic and probabilistic methods to guarantee that our algorithms perform well both in terms of modelling accuracy and robustness, and in terms of computational complexity and practical efficiency.

The research in Algodan is grouped under four interacting themes: sequence analysis, learning from and mining complex and heterogeneous data, discovery of hidden structure in high-dimensional data, and foundations of algorithmic data analysis. All these themes combine aspects of combinatorial pattern matching, data mining, and machine learning.

The host organizations of the Algodan CoE are University of Helsinki and Helsinki University of Technology. The CoE is in part a continuation of the “From Data to Knowledge” CoE, and consists of about 70 persons. The director of the Algodan CoE is Professor Esko Ukkonen and the vice-director is Academy Professor Heikki Mannila.