

# Next Generation Information Technologies and Systems

Proceedings of 4<sup>th</sup> International Workshop, NGITS'99, Zikhron-Yaakov, Israel, July 1999

edited by Ron Y. Pinter and Shalom Tsur  
Lecture Notes in Computer Science Vol. 1649  
Springer, Berlin Heidelberg, 1999, 325p.

ISBN 3-540-66225-1

**Keywords:** Data-Mining, Web Exploration, Query Optimization, Spatio-Temporal Data, Object-Oriented Databases, Dynamic Relationships, Deductive Database Languages, Multidatabases, Meta Information, Agent and Workflow Management Technology, Data Warehousing

## Introduction

The *Next Generation Information Technologies and Systems (NGITS)* workshop has been held in Israel every two years since 1993. The NGITS'99 clearly reflects the trend of many of the basic technologies in the traditional areas of database management systems being increasingly deployed in new systems and applications supported by a huge growth in computing speed, storage capacity, communications and world-wide web. In turn, these applications require serious improvements in data availability, information integrity and knowledge extraction. The workshop gathered many reputable members of the research community. Out of the thirty-four submissions, twenty-two (seventeen full-length papers and five short papers) were selected by the Program Committee to be presented at the workshop and to be included in the Proceedings. The papers were classified in five broad topics underlying the workshop sessions, as follows:

- Exploring the World Wide Web;
- Database Technology;
- Storage, Meta Information, Ontologies and Software Engineering;
- Agent and Workflow Management technology;
- Data Warehousing and Mining.

Two Invited Talks were delivered at NGITS'99:

**"Some Advances in Data-Mining Techniques"**, by J. D. Ullman (Stanford University). It tackled 3 main areas of interest in

the MIDAS research project at Stanford: a new algorithm for Web search which resulted in a commercially available search engine called Google, generalization of techniques such as "a-priori" (developed at IBM Research Center in Almaden) to allow "market-basket analysis" or "association-rule mining" (finding items that customers frequently buy together) and summarizing the knowledge of the Web in a form that resembles conventional relational data. Techniques developed by the MIDAS group to process very large amounts of data and efficiently detect items that are highly correlated but not very frequent, have been another point of the talk.

**"Workflow Management in the Internet Age"** by C. Mohan (INRIA Rocquencourt, France and the IBM Research Center at Almaden, San Jose, USA). Based on the author's experience as founder and leader of the EXOTICA workflow project at IBM Research Center and on his close collaboration with the IBM MQSeries Workflow and Lotus Notes product groups, the talk dealt with various directions for research and potential future extensions to the design and modeling of workflow management systems

The Proceedings include the papers selected for NGITS'99. They give a comprehensive image of advances in research and development in the areas of database and Web technologies as well as valuable insights into workflow management and data warehousing.

It is an interesting reading for both researchers and students in computer science but also for any open minded reader concerned with information technologies of the future.

In the following, an overview of the main topics covered by the sections of NGITS'99 is taken.

## 1. Web Exploration

The papers on Web exploration tackle several research directions, such as:

solutions and systems (in particular, the system QUEST) for querying and manipulating semantically tagged documents on the World-Wide Web. The documents are written in the markup language OHTML, a language which combines HTML and objects of the OEM data model;

design of new query languages for databases on the Web. The paper proposes a simple graph based language where both the query and its target are graphs;

strategies for filtering E-mail messages combining content-based and sociological filtering with user-stereotypes;

new methods for refining an initial set of search results, obtained through a meta-search engine, based on relevance feedback; the idea is to interactively obtain from the user a subset of relevant documents in an ongoing query, thereby a sample of the related vocabulary is provided. Terms acquired this way are combined with the terms initially in the query, in order to improve retrieval accuracy.

## 2. Database Technology

The Section gathered papers dealing with current trends in database research:

- solutions for optimizing object-oriented database queries, exploiting the potential invertibility of navigational operators in queries;
- modeling multidimensional arrays, heavily used in database applications ranging from GIS to OLAP; a cross-dimensional and application-independent algebra is presented, for high-level treatment of arbitrary arrays;
- dynamic relationships and their propagation in object-oriented databases;
- tracking moving objects using database technology – the project DOMINO;
- design of novel deductive languages to support important structurally object-oriented database features such as: object identity, complex objects, typing classes, class hierarchies, multiple property inheritance with overriding, conflict-handling and blocking and schema definitions in a uniform framework; a theoretical foundation is thus established for a practical deductive object-oriented database system for advanced applications;

- defining a formal model for multidatabases; the proposed model is based on formal assumptions of integrability, supports database heterogeneity and provides several degrees of freedom allowing it to model actual situations encountered in multidatabase applications;
- modeling temporal active rules and their application to various types of databases.

## 3. Storage, Meta Information, Ontologies and Software Engineering

The Section included a rather diverse set of papers, dealing with:

- cost effective jukebox storage via hybrid file-block caching;
- an ontological framework to capture semantic relationships between data sets spread over distributed heterogeneous scientific repositories and to allow the inference of valid combinations of scientific resources for the production of new data;
- global version management in federated systems; solutions are validated in the context of a federated turbine design environment and then generalized for a wider range of application domains;
- a language (the Object-Process language) as a natural, yet formal, way of specifying system analysis findings and design decisions, complementary to the graphic description through Object-Process Diagrams;
- a new methodology, requiring extremely low investment, to achieve reuse and knowledge sharing, by applying information retrieval techniques to object-oriented resources, while exploiting the object-oriented language semantics and characteristics.

## 4. Agent and Workflow Management Technology

The three papers included in this Section present:

- a component-based workflow system with dynamic modifications; a workflow process definition language, its graphical representation and a workflow process modification language are used to illustrate how the modification process is handled;
- a temporal reasoning approach to model workflow activities; the proposed logical



framework introduces a convention for representing causal laws and includes higher-order persistence and potential events;

- a model for dynamic configuration and enforcement of access control for mobile components; the enforcement mechanism extends Java 1.2 and allows mobile components to fulfill their "contracts" despite their potential relocation.

## 5. Data Warehousing and Mining

This last Section includes papers concerning:

- the MyView project, aiming at the integration of both structured and unstructured bibliographic information from a diversity of heterogeneous Internet repositories like electronic journals and traditional libraries;
- a tightly-coupled approach to integrating data mining with relational DBMSs;
- a new approach for morphological disambiguation to enable linguistic indexing for Hebrew search systems.

## 6. Conclusions

The NGITS'99 International Workshop constituted a good opportunity to exchange ideas between researchers in rapidly evolving areas as database, Web, and workflow technologies. The Proceedings of the workshop are a valuable reading for anyone interested in the current status and future trends in these different, yet converging, fields.

**Radu-Dan Bercaru**  
**Head of Database Systems Laboratory**  
**National Institute for Research and**  
**Development in Informatics**  
**Bucharest**