Knowledge Management in Environmental Decision Support Systems

Ulises Cortés[1], **Miquel Sànchez-Marrè**[1], **Joaquim Comas**[2], **Ignasi R-Roda**[2] and **Manel Poch**[2]

Abstract. In the last decade, Environmental Decision Support Systems (EDSS) has emerged as an effective software tool for supporting control decision-making to maximize the performance of a system and to minimize the negative impact of faults in environmental processes. Knowledge Management in Environmental Decision Support Systems appear as a necessity for users to place guarantees on the system's behavior but these methodologies are still under-explored. In this paper, a first approach to put forward a Knowledge Management Methodology by identifying the most relevant issues is introduced and discussed.

1 INTRODUCTION

The Environment is a system which is governed by complex laws of behavior, and the action of human societies on it has an important impact. For example, the increasing rhythm of industrialization, urbanization and population, which the earth has have faced this century, have forced society to consider whether human beings are changing the very conditions essential to life on earth. Often, there is a need to influence the dynamics of an environmental process and to bias its evolution into a desired direction. But environmental processes are not easy to model and our knowledge is still incomplete and uncertain.

A major research line in Artificial Intelligence community is concerned with building up systems that are able to help in the problem-solving tasks using both specific and general knowledge. Environmental Decision Support Systems (EDDS) [6], [7] and [21] are good examples of this interest.

In the last decade, EDSS came out as an adequate software tool for supporting control decision-making to maximize the performance of a system and to minimize the negative impact of faults. Particularly, because in environmental processes this implies a continuous intelligent monitoring system, and an increasing volume of data and, many times, a decreasing time horizon within control decision have to be made. Also, the global nature of Environmental problems creates the need for distribute the computation, and then again, the need for integrate and share the obtained partial results.

The real impact of those decisions on the Environment requires that they are Knowledge-Based. That is, an EDSS is not only capable of generating reasonable and understandable solutions, but also it should allow to incrementally extend the domain models it has. This has been shown crucial to accept these systems to work in real world domains [1] and [23]. EDSS are Environmental Sensitive Systems able to act to sustain Environment.

Knowledge Management in Environmental Decision Support Systems appear as a necessity for EDSS users to place guarantees on the system's behavior but these methodologies are still under-explored. Here we present a first approach to put forward a Knowledge Management Methodology by identifying the most relevant issues.

The organization of this paper is as follows: First, in §2, we approach the issue of the Exploitation of different sources of Knowledge. After we address the issue of Knowledge Management in EDSS as the key factor for high performance, then we touch on the issues related with the use of Intelligent Agents and the decision-making process. Finally, some conclusions are given.

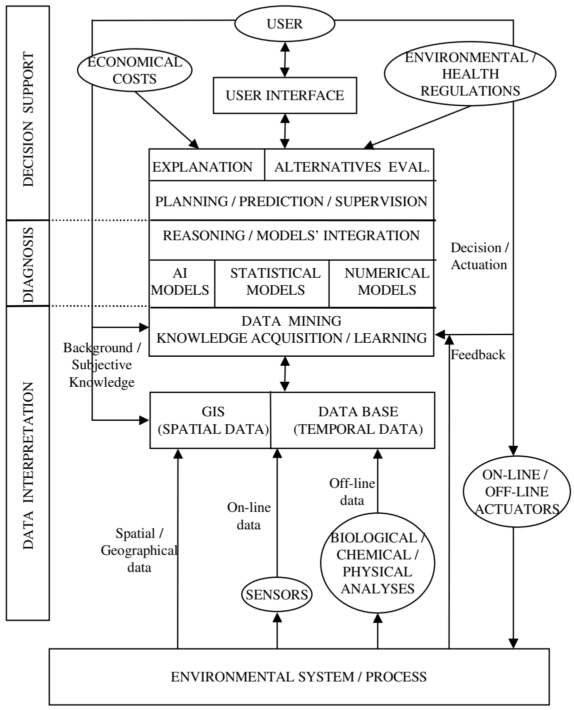

Figure 1. EDSS architecture

2 EXPLOITATION OF KNOWLEDGE RESOURCES

Our proposal of an architecture for Knowledge Management in EDSS has three levels: Data Gathering and Interpretation, Diagnosis or Prediction and, Decision Support (see figure 1). Those levels permit to capture the complex nature of environmental problems and specification of the interaction

1 Artificial Intelligence Section, Dept. of Software, Technical University of Catalonia, Campus Nord-Edifici C5, E-08034 Barcelona, Catalonia, EU, email: {ia | miquel}@lsi.upc.es

2 Environmental and Chemical Engineering Laboratory (LEQUIA), University of Girona, Campus de Montilivi, E-17071 Girona, Catalonia, EU,email: {quim|ignasi|manel}@lequia.udg.es

between different levels of reasoning involved. It is meant to be implemented as a multi-agent system where agents respond in a reasoned way to their goals, and events that occur in their environment. These agents have a specific set of conditions and associated goals, which indicate the events they should respond to.
This architecture stresses the problem of heterogeneous information and knowledge sources: the data, that EDSS needs to cope with, usually ranges from huge amounts of numeric streams arriving in real time from a variety of sensors, over visual information from video cameras, to quite informal messages such telephone calls: incomplete, erroneous or even noisy data may arrive. Another agent-based architecture designed to deal with Environmental issues is ARTEMIS that has five levels: Classification, Diagnosis, Prediction, Planning and, Action Selection [9].
Environmental Decision Support Systems use different Knowledge sources and this usually implies different ways to extract and combine information. The appropriate interpretation of this combination of information will result on the identification of the environmental events as for example:

a) Hazard identification.
b) Risk Assessment
c) Risk Evaluation
d) Intervention/Decision-making

where an EDSS play an strategic role in the decision-making chain helping to solve problems more efficiently.
Intelligent Access to Data (see figure 2) provides a wide selection of tools to exploit the various Knowledge resources that are available. The data space in Environmental problems is not very easy to define nor to the data easy to collect and to interpret.
Among those knowledge resources we can find: Conventional Data Bases as for example Object-Oriented Databases or Relational Databases that require specialized access languages and perfect match. Rule-based Knowledge Based Systems as for example, Expert Systems that presuppose a codified body of knowledge.

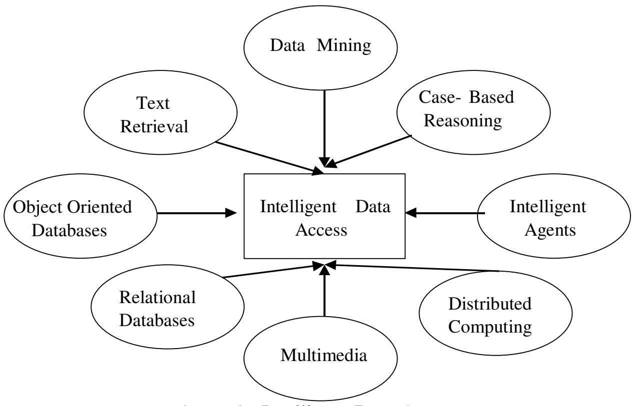

Figure 2. Intelligent Data Access

More elaborated methods for knowledge exploitation as Knowledge Discovery or Data Mining are used to find *hidden* relationships from compiled past cases.
Multimedia and Geographical Information Systems provide new sources of information that are increasingly introduced in EDSS but require automated ways for interpreting their contents [16]. For most of the Environmental problems there exists huge quantities of data about the process itself but the information about the causal or dependence relations among variables is not well-known or very rare, for example see [14].
Case-Based Reasoning allows fuzzy search and the use of past experiences to interpret and solve actual situations. CBR allows the systems to detect impasses. Also this technology allows the EDSS to learn giving access to experience as well as expertise [26]. In some cases the use of a CBR approach includes the extensive use of Ontologies to improve re-use of experiences and diminish impasse situations. What is still lacking is clear understanding of *how* to build Ontologies in systematic way.

3 KNOWLEDGE MANAGEMENT IN EDDS

Knowledge Management appears as a solution to integrate and exploit information coming from different sources, formats, granularity, etc to help in the decision-making process in EDSS, but this process is general enough to be applied to many other domains.
This Knowledge Management process has first to answer to two questions: *What kind of information must be integrated?* and *How to integrate that information?*
The first question is usually approached as a Knowledge Acquisition process and it has been recognized, in general, as a very slow and expensive process. This is also true in Environmental Problems. The availability of knowledge sources is scare and in many cases the Knowledge about an environmental process is considered as *strategic* for the enterprise that owns it and therefore not to be shared.
About *what* is necessary to integrate (see figure 3) the first step is related with the classification of the streams of Real Data coming from on-line sensors and/or other sources. This classification process is intended to reduce data complexity. The correlation step in a first phase is devoted to use the *indexation* produced by the classification process to create abstract types of objects as classes and/or prototypes that can be used in the reasoning towards the decision-making process. In the *correlation* phase, it is important to discover the existing relation between abstract objects and decisions. For example, to measure how the use of an abstract object impacts in the quality of the decisions.

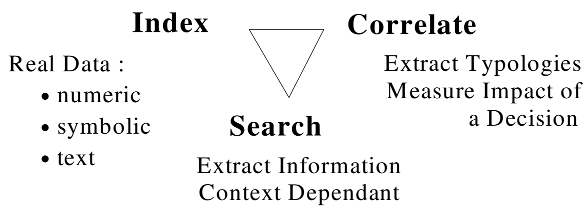

Figure 3. Integrating technologies: What?

In the *search* phase, the Knowledge Management System will *select* the relevant complementary information for the decision-making process. Also, in this phase, there is a task devoted to find and organize the context dependant information. We can identify this phase as a first step towards the construction of an Ontology [15].
This methodological approach suggests that in large knowledge-based systems should be organized (*integrated*) in order to facilitate knowledge acquisition and maintenance. This is clearly related with the notion of Knowledge-Level Model [19].

About *how* to integrate information, the processes are more complex and most of the involved tools and methods are necessarily specialized for each domain (see figure 4).

The integration of raw data is usually carried on by means of a kind of data structures ranging from Data Bases through Knowledge Bases. More elaborated structures as Ontologies are important as they include both data and knowledge. An Ontology provides the means for describing explicitly the conceptualization behind the knowledge represented in a knowledge base.

Ontologies are effective ways to unify terminology in a given domain. That is, they provide a means for common domain modeling.

All this integration will lead the EDSS towards the extensive use of domain models in the sense of *components of expertise* [29]. This integration of knowledge structures and methods describes the domain's *deep knowledge* level.

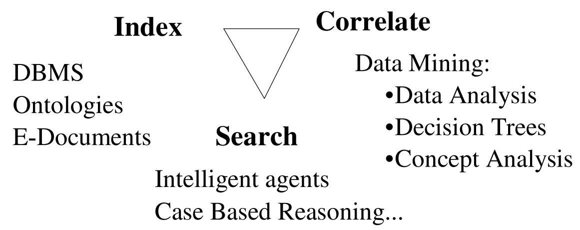

Figure 4. Integrating technologies: How?

In the integration phase the notion of task is an important one. A *task* is an abstract description of *what* is needed to do in order to achieve an objective.

Usually, tasks are decomposed following a top-down approach, and described in an and/or tree. Once tasks are decomposed they have to be distributed among agents. The general description of a task includes its name a short description, input and output, task structure, its control, preconditions and the required conditions and capabilities of the performer (an agent).

In the integration phase correlation takes profit of the existing data or knowledge structures. The use of unsupervised discovery methods as for example: Data Mining, Knowledge Discovery, etc helps to build new ways to interpret the domain.

The problem-solving methods apply domain knowledge to accomplish an intended *task*. In general, they perform two functions: divide a task into a number of sub-tasks or directly solve a sub-task. In either case, they can consult domain models; create or change intermediary problem-solving structures; perform actions to gather more data, for example, by querying the user or performing a measurement (see figure 2); and expand the problem solving situation model by adding or changing facts.

This integration and distributed problem solving is able to cope with the supervision of the different tasks of the system, to deal with any kind of data gathered from the process (quantitative and qualitative), and to include different types of knowledge co-existing in the domain: numerical, expert, experiential, and predictive.

The task of co-ordination among the different levels of abstraction is performed by an agent that is responsible for detecting interdependencies between agents' activities, this assures a correct distribution of tasks in the decision-making process.

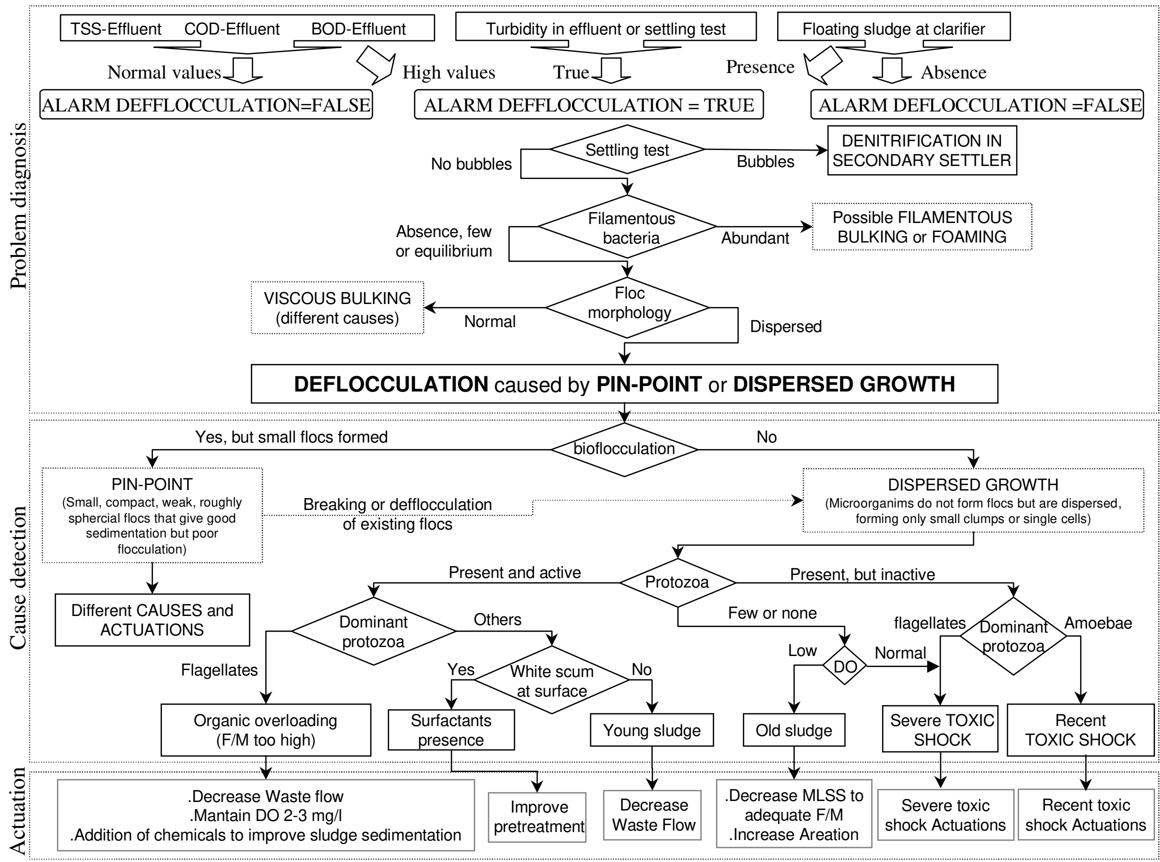

Figure 5. Decision tree to deal with deflocculation problem in wastewater treatment plants.

4 AGENTS AND THE DECISION-MAKING PROCESS

The decision making-process relies on a cycle that includes: *recognition* of the causes of an event (diagnosis) or the causes for a future event (prediction), *plan formation* to solve the incidences and, *execution* of the selected actions. Any agent that performs its activities in a changing world must model that world internally. Some agents model aspects of their world in a explicit way, so they can reason about it. In other multi-agent systems these models are often distributed thorough the architecture as not always the model is needed to identify local changes or to solve local problems [12].
In figure 5 we show the integration of information in the decision making-process performed by DAI-DEPUR [27]. Although it refers to a very specific case in the Wastewater Treatment process it shows the flow of information from sensors and laboratories (on the top) towards the Reasoning components. This figure shows different kinds of knowledge facilities that have to be used and coordinated using very precise communication protocols and, it gives a good taste of the complexity of these processes.
The modeling of different kinds of knowledge or domain theories into separate levels provides the architectures with additional modularity and independence that are always desirable in computational systems for achieving more general architectures.
The main reasons for interconnecting autonomous agents and expert systems, as shown in figure 6, are: to enable them to co-operate in solving problems, to share expertise, to allow parallel work, to take profit of modularity, to allow the system to be fault tolerant through redundancy, to catch and take profit from multiple viewpoint and knowledge/expertise of multiple experts, and to be *reusable.*
Many EDSS could be analyzed as Cooperative Information Systems [32] where agents are used to cope with the different sources of information and the different tools used for problem-solving (see figure 6). For EDSS to be able to co-operate they must have an intelligent interface that can cope with all types of requests and eventualities. Through this interface the system can communicate with other systems and reason about the information it contains. In this perspective, agents are given with meaningful units of work or tasks to be solved. Performing a task in many situations involves to initiate a communication with other agents.
The design of the communication protocols among the different agents is decisive in the system's performance. A communication protocol determines *how* the interchanges of illocutions among agents are structured. Agents may simply give orders to others and expect them to be executed, in other cases negotiation or more complex exchanges can take place. The communication channels through which information moves from one agent to another can differ in many ways as for example:

- Medium (shared physical environment vs. Internet)
- Addressing (broadcast, black-board, subject-based, personal communication, etc)
- Locality
- Persistence of the communication (How long?)

When something fails in this *integration* process then an impasse situation is produced. Failures can occur for example when an agent is unable to finish its task or to deliver the results of their task on time. As said before the use of specific Ontologies is one of possibilities to reduce the number of non solved impasses.

The plan formation requires de co-operation among the different agents or knowledge-based systems (see figure 6). Agents are autonomous in that they perform their tasks until they are required. We assume hierarchical coordination, for example: commands flow down from higher levels and status information flows bottom-up. In these systems each agent usually owns an *agenda* containing the actions to be performed and the agent's objectives. This agenda helps the agents to build plans and to accept tasks from the *manager***.**

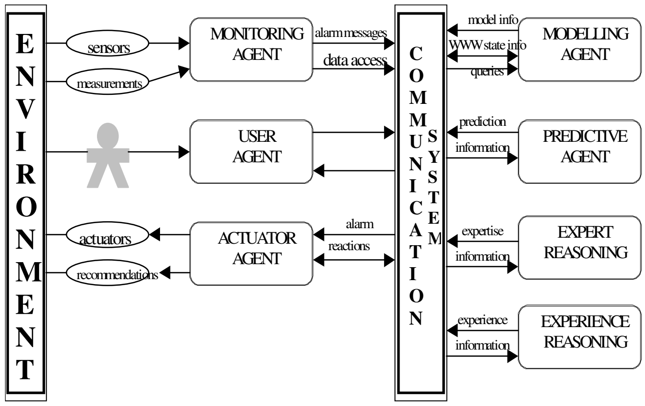

Figure 6. Cooperation among agents

The distribution of tasks can improve system performance. Durfee [Durfee, 1999] indicates that the combination of efforts brings:

- **Confidence**: Independent derived results can be used to corroborate each other, yielding a collective result that has a higher confidence of being correct.
- **Completeness**: The results from different subtasks can altogether cover a bigger portion of the overall task.
- **Precision**: To refine its own solution, an agent needs to know more about the solutions that others have formulated.
- **Timeliness**: Solving subtasks in parallel can yield an overall solution faster.

For EDSS the best mechanism for task distribution are:

- Multi-agent planning: i.e. planning agents have the responsibility for task assignment.
- Agency structure: i.e. agents have fixe responsibilities for particular tasks

Is in this phase when the reasoning components of each agent in the system have to perform its assigned task. We assume the bounded rationality [28] as an important characteristics of the agents involved.

5 CONCLUSIONS

Solutions for an efficient Knowledge Management in Environmental Decision Support Systems are to be developed. As we have shown one has to consider the great variety of data, and the strong dependencies of environmental processes to local constraints, such as weather conditions, climatic aspects, geographical positions, environmental and/or health law regulations, etc.
Methods, Models and Tasks are means to ensure desirable Knowledge-Based behavior and are crucial elements of the

Knowledge-Based Systems design. These represent the *components of expertise* needed to handle the problem. In figure 7, we advance a proposal to combine those components of expertise to build EDSS.

Advances in the area of Model-Based Reasoning are very promising [18] and specific models are to be developed for this field. Environmental problems require models of a great *generality*, *precision* (if possible) and *realism*.

Agent-based architectures introduce a powerful metaphor in the field of EDSS as agents integrate a collection of fuctionalities, achieved by the interplay about certain problem types and about the environment in which those agents operates. This allows the system to react to the Environment changes and situations and it is able to interact with other agents looking for an answer. The use of multi-agent systems to distribute the problem solving demands both group coherence and competence.

As stated the methodological use of Knowledge Management in Environmental Decision Support Systems is an open field of research, a first attempt is the one reported by Roda [24].

The Knowledge Management in EDSS research, from our point of view, has to deal with the following issues:

- Techniques for the performant integration of several data, knowledge and AI techniques,
- Improved knowledge acquisition methods,
- Ontologies or equivalent paradigms to represent better the knowledge over the Environmental system,
- Development of specific coordination models for Environmental problems
- Intelligent sharing and reuse of knowledge,
- Reuse of EDSS generic models and tasks,
- Security,
- Validation
  - How performance is measured?
  - How to measure the competence of an agent to perform a task? and,
  - How the measure the EDSS's performance?

In figure 7 you can see depicted an EDSS development process following the methodology proposed in [28]. The *collecting data, data analysis* and *model selection* phases clearly answer to the question *what kind of technologies are needed*?

The following *model implementation, model integration* and *EDSS evaluation* phases answer to the question *how to integrate those technologies*? In our opinion this represents an easy way to transport and reuse past experiences from one process to another. Also this methodology is general enough to be applied to other fields where complex processes exist.

We have to keep on mind that Environment is a critical domain where wrong management decisions may have disastrous economic and/or environmental consequences.

EDSS are important tools to understand our environment and to design the development our societies to be sustainable. EDSS are Environmental Sensitive Systems.

EDSS play a key role in the interaction of man and environment as they have to cope with the multidisciplinary nature and high complexity of environmental problems they result in a very fruitful field of work. And its main goal is to try to keep in good shape the limited resources of our planet.

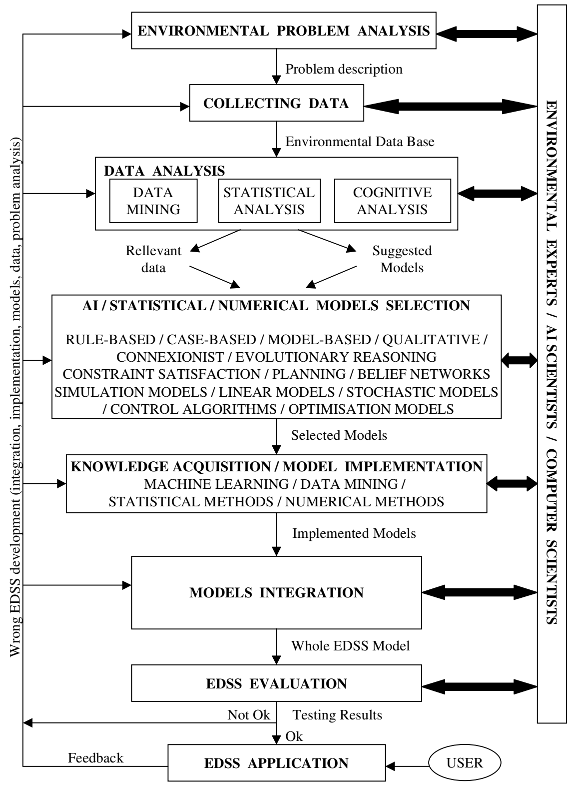

Figure 7. EDSS development process

ACKNOWLEDGEMENTS

We acknowledge the support of and of the Spanish CICyT project AMB97-889. The European Union project IST-1999-176101. The views in this paper are not necessarily those of A-TEAM consortium.

REFERENCES

[1] Avouris, N.M. 1995. Co-operating Knowledge-Based Systems for Environmental Decision-Support. *Knowledge-Based Systems*. 8(1):39-53.

[2] Binding Environmental Sciences and Artificial Intelligence. http://www.lsi.upc.es/~webia/besai.html

[3] Branting. L.K., Hastings, J.D., and Lockwood, J.A. 1997 Integrating cases and models for prediction in biological systems. *AI Applications* 11(1):29-48.

[4] Brilhante, V. 1999. Using Formal Metadata Descriptions for Automated Ecological Modeling. In Environmental Decision Support Systems and Artificial Intelligence. AAAI-Workshop. WS-99-07. Pp 90-95.

[5] Ceccaroni, L. 2000. Integration of rule-based expert system, a case-based reasoner and a ontological knowledge-base in the Wastewater Domain. Personal Communication.

[6] Cortés, U. , Sànchez-Marrè, M., Ceccaroni, L., R-Roda, I. & Poch, M. 2000. Artificial Intelligence and Environmental Decision Support Systems. *Applied Intelligence*. 13(1):225-239. 2000.

[7] Cortés, U. & Sànchez-Marrè, M. 1999. Binding Environmental Sciences and Artificial Intelligence in Environmental Modelling and Software. *Environmental Modelling & Software* 14(5): 335-337.

[8] Cortés, U., Sànchez-Marrè, M., R.-Roda, I. and Poch, M. 1999. Towards Environmental Decision Support Systems. In Workshop on Environmental Decision Support Systems. Cortés, U. & Sànchez-Marrè, M. (Eds). WS-99-07. AAAI Press.

[9] Cuena, J. & Ossowski, S. 1999. Distributed Models for Decision Support. In Weiss, G. (Ed) Multiagent Systems. MIT Press. 1999

[10] Durfee, E. 1999. Distributed Problem Solving and Planning. In Weiss, G. (Ed) Multiagent Systems. MIT Press. 1999

[11] Fedra, K. & Winkelbauer, L. 1999. A hybrid Expert System, GIS and Simulation Modelling for Environmental and Technological Risk Management. In Environmental Decision Support Systems and Artificial Intelligence. AAAI-Workshop. WS-99-07. Pp 1-7.

[12] Ferber, J. 1995. Les systèmes multi-agents: vers une intelligence collective. Intereditions. France.

[13] Gasser, L. 1991. Social Conceptions of Knowledge and Actions: DAI foundations and open systems semantics. *Artificial Intelligence* 47:107-138.

[14] Gimeno, J.M., Béjar, J., Sànchez-Marrè, M., Cortés, U. & R.-Roda, I.. 1998. Discovering and Modelling Process Change: an Application to Industrial Processes. Proc. of 2nd Int. Conference on the Practical Application of Knowledge Discovery and Data Mining (PADD 98), pp. 143-153. London, U.K.

[15] Gruber, T. 1993. A translation approach to portable ontology specifications. *Knowledge Acquisition*. Vol. 5. 199-220.

[16] Jones, E. and Roydhouse, A. 1995. Retrieving structured spatial information from large databases: a progress report .Procc. of IJCAI Workshop on Artificial Intelligence and the Environment, pp. 49-57, Montreal.

[17] Knowledge Sharing Effort http://www.cs.umbc.edu/kqml

[18] Lee, M.H. 1999. On models, modelling and the distinctive nature of model-based reasoning. *Ai*Communications 12(3):127-138

[19] Newell, A. 1982 The Knowledge Level. *Artificial Intelligence* 18: 87-118.

[20] Riaño, D., Sànchez-Marrè, M. & R-Roda, I. 1999. A wastewater Treatment Plant Multiagent System. Practical Application of Agents and Multiagent Systems. 3rd World Multiconference on Systemics, Cybernetics and Informatics (SCI'99).

[21] Rizzoli, A.E. & Young, W.Y. 1997. Delivering environmental decision support systems: software tools and techniques. *Environmental Modelling & Software*. 12(2-3):237-249.

[22] R-Roda, I., Comas, J., Colprim, J., Baeza, J., Sànchez-Marrè, M. & Cortés, U. 1999. A Multi-Paradigm Decision Support System to improve Wastewater Treatment Plant Operation. In Environmental Decision Support Systems and Artificial Intelligence. AAAI-Workshop. WS-99-07. Pp 68-73.

[23] R-Roda, I., Poch, M., Sànchez-Marrè, M. Cortés, U. & Lafuente, J. 1999. Case-Based Reasoning Systems: A Suitable Tool for the Supervision of Complex Processes. *Chemical Engineering Progress* 95(6):39-45.

[24] R.-Roda, I. 1998. Development of a protocol for the application of knowledge based systems to wastewater treatment plants management. Ph.D. Thesis. Universitat de Girona (In Catalan).

[25] M. Sànchez-Marrè, U. Cortés, I. R-Roda & M. Poch. 1999. Sustainable case learning for continuous domains. *Environmental Modelling & Software* 14(5): 349-357.

[26] Sànchez-Marrè M.; Cortés, U.; R.-Roda, I.; Poch, M. & Lafuente, J. 1997. Learning and adaptation in WWTP through case-based reasoning. *Microcomputers in Civil Engineering*, (Special issue: Machine Learning), 12(4):251-266.

[27] Sànchez-Marrè, M.; Cortés, U.; Lafuente, J.; R.-Roda, I. & Poch, M. 1996. DAI-DEPUR: a Distributed Architecture for Wastewater Treatment Plants Supervision. *Artificial Intelligence in Engineering*, 10(3):275-285.

[28] Simon, H. A. 1996. The Sciences of the Artificial. MIT Press. 3rd ed.

[29] Steels, L. 1990. Components of expertise. *AI* Magazine 11(2):28-49.

[30] Stephanopoulos, G. and Han, C. 1996. Intelligent Systems in process Engineering: a review. *Computers Chem. Engng*, 20(6/7):743-791.

[31] Thomson, A.J. 1997. Artificial Intelligence and environmental ethics. *AI applications* 11(1):69-73.

[32] Verharen, E.M., Dignum, F. and Bos, S. 1998. Implementation of a Cooperative Agent Architecture Based on the Language-Action Perspective. In Intelligent Agents IV. Singh, M.P., Rao, A., and Wooldrige, M. (Eds). pp31-44 LNAI 1365.

[33] Weiss, G (eds). 1999. Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. MIT Press.